

Plantwide modelling – anaerobic digestion of waste sludge from parent nutrient (N and P) removal systems

DS Ikumi¹ and GA Ekama¹

¹Water Research Group, Department of Civil Engineering, University of Cape Town, Rondebosch, 7700, Cape Town, South Africa

ABSTRACT

Wastewater treatment plant (WWTP) mathematical models are based on the behavioural patterns of microorganisms involved in the treatment process. These microorganisms are assumed incapable of thinking or planning but simply act according to the capabilities afforded to them by their surrounding conditions – hence different microorganisms pre-dominate different WWTP zones according to how well the conditions suit them. When waste activated sludge (WAS) from biological nutrient removal (BNR) activated sludge (AS) systems, containing phosphorus-accumulating organisms (PAOs), is fed to an anaerobic digester, there is a release of high quantities of metals, phosphorus (P) and nitrogen (N). The manner in which we model the release of these metals and nutrients significantly affects the accuracy of predicted anaerobic digestion (AD) outcomes. Previous studies of PAOs show that in the anaerobic zone of the AS system, they can form energy-rich poly-3-hydroxybutyrate (PHB) at the expense of their aerobically generated polyphosphate (PP). Thus, it is expected that the PAOs containing PP sent into an anaerobic digester with volatile fatty acids (VFAs) present, would utilize their PP reserves as they would in the anaerobic zone of an AS process ending up with formation and storage of some PHB. Ultimately, all the stored products of the PAO get released, since there is no alternating aerobic environment to cater for their growth. Since it has been established that the PP release in the AD occurs much faster than the PAO biomass hydrolysis rate, it is modelled as a separate process. Steps are presented in the development of this PP release mass-balanced stoichiometries that occur with AD of PAOs. By comparing outcomes from these proposed stoichiometries against measured experimental data, it is noticed that better predictions are obtained with acetate uptake for PHB formation than when modelling the AD PP release to occur with PAO death and hydrolysis.

Keywords: phosphorus removal, activated sludge, anaerobic digestion, plant-wide modelling

INTRODUCTION

The drive to recover phosphorus (P) and predict potential problems in wastewater treatment plant (WWTP) systems, such as struvite precipitation and pH changes in various unit processes, has led to extensive research into P removal in WWTPs. Phosphorus is removed from wastewater by transforming it from the dissolved liquid phase to the intracellular solid phase. Hence it was noted that for the AD of P-rich sludge, from biological excess P removal (EBPR) activated sludge (AS) systems, three-phase mixed weak acid/base chemistry is required because the release of biomass P or polyphosphate (PP) not only affects the system alkalinity but also can induce precipitation of minerals such as struvite (Van Rensburg et al., 2003; Harding et al., 2010). Batstone et al. (2012) outline the methods, available knowledge and the fundamental approaches towards the development of a physicochemical framework that supports the systematic development of such three-phase plant-wide models. This approach was applied by Lizarralde et al. (2015) in the development of a new general methodology for incorporating physicochemical and chemical transformations into multi-phase plant-wide models. Also, various research groups have worked both collaboratively and separately on related topics, towards the development of WWTP mathematical models that integrate bioprocess stoichiometry and physicochemical transformations, such that processes such as nutrient release and multiple mineral precipitations (MMP) can be included.

Related to this is the progress made on extending the Anaerobic Digestion Model No. 1 (ADM1; Batstone et al., 2002), for the inclusion of P removal and associated

processes. Flores-Alsina et al. (2016) presents ADM1 with an improved physicochemical description, P and sulphur (S) biotransformation, iron (Fe) reduction and multiple mineral precipitation (MMP) to show the effect of including the interactions amongst P, S, Fe and observing their potential effect on predicted model outcomes, such as total biogas production (CO₂, CH₄, H₂ and H₂S). The MMP component of this model was included according to the semi-mechanistic modelling approach of Kazadi Mbamba et al. (2015). Further extensions to ADM1 to accommodate P include studies by Wang et al. (2016), for including phosphorus-accumulating organisms (PAOs) uptake of various volatile fatty acids (VFAs) for the formation of poly-3-hydroxybutyrate (PHB), the effect of resultant PHB content on disintegration rate and the potential precipitation of the P (mostly released with the anaerobic PHB uptake). Bioprocess stoichiometry and physicochemical transformations are also included in the presentation of a new 'three-phase' (aqueous-gas-solid) plant-wide model that includes P, (PWM_SA) (Brouckaert et al., 2010; Ikumi et al., 2015) which includes compatible activated sludge (AS; ASM2-3P) and anaerobic digestion (AD; SDM3P) bioprocess model components and uses strict mass balance principles to track P through the unit processes of a WWTP (with recognition of its impact on the mutual interaction between the connected unit operations). The AD component of PWM_SA extends the AD model of Söttemann et al. (2006) by including multiple organic types (including WAS with PAOs), P release stoichiometry and multiple mineral precipitation. When including the behaviour of PAO biomass in AD, Ikumi (2011) concluded that most of the P release occurred with acetate uptake and PHB formation, which posed the question of whether there is any energy transfer between the AS and AD systems with PP release and, if so, how this energy transfer should be represented in AD. The significance of this energy

*Corresponding author, email: david.ikumi@uct.ac.za
Received 11 April 2018; accepted in revised form 26 June 2019

representation in AD systems is highlighted by the impact of the PP release process towards the determination of system pH and carbon dioxide partial pressure ($p\text{CO}_2$), which are required to be accurate for prediction of mineral precipitation potential and concentrations in the AD aqueous phase. This paper investigates the appropriate stoichiometry for this PP release process, using results from the investigation of an enhanced biological P removal (EBPR) AS system linked to an AD.

METHODS

In the study by Ikumi (2011), a University of Cape Town (UCT) NDEBPR system was operated and fed 600 mgCOD/L settled wastewater together with additives (i.e., 200 mgCOD/L acetate and 40 mgP/L from di-potassium hydrogen phosphate (KH_2PO_4)). This UCT AS system was operated at steady state, at a 10-day system sludge age (SRT) and its WAS was anaerobically digested in a separate flow-through anaerobic digestion (AD) system operated successively at different solids retention times (SRTs) (Table 2). The experimental set-up of Ikumi (2011) is given in Fig. 1. The results from this experimental set-up are used to examine the transfer of P to the AD and the release of poly-phosphate (PP) in the AD fed the NDEBPR WAS.

AD volume was always 16 L apart from 60 d AD which was operated using small 5 L AD volume reactors, The NDEBPR WAS taken from AS system was 5.7 L/d at a concentration of about 9 gCOD/L. The required volume of this WAS was fed to the AD without thickening.

Model implementation

A three-phase plant-wide model (Ikumi et al., 2015) was developed for simulating the biological processes to track and predict the output of materials (chemical oxygen demand (COD), carbon (C), hydrogen (H), oxygen (O), nitrogen (N), phosphorus (P), magnesium (Mg), potassium (K) and calcium (Ca)) along the unit processes of a WWTP and tailored to conform to the laboratory AD and AD connected systems. The model comprises 3 sub-models, integrated for simulation of the entire WWTP comprising different unit operations (i.e., UCT NDEBPR AS system linked to an AD or anoxic-aerobic digestion (AAD) for WAS stabilisation). The sub-model components used to simulate the above experimental set-up include:

- The ionic speciation model (Brouckaert et al., 2010). This model includes the pairing of ionic components and

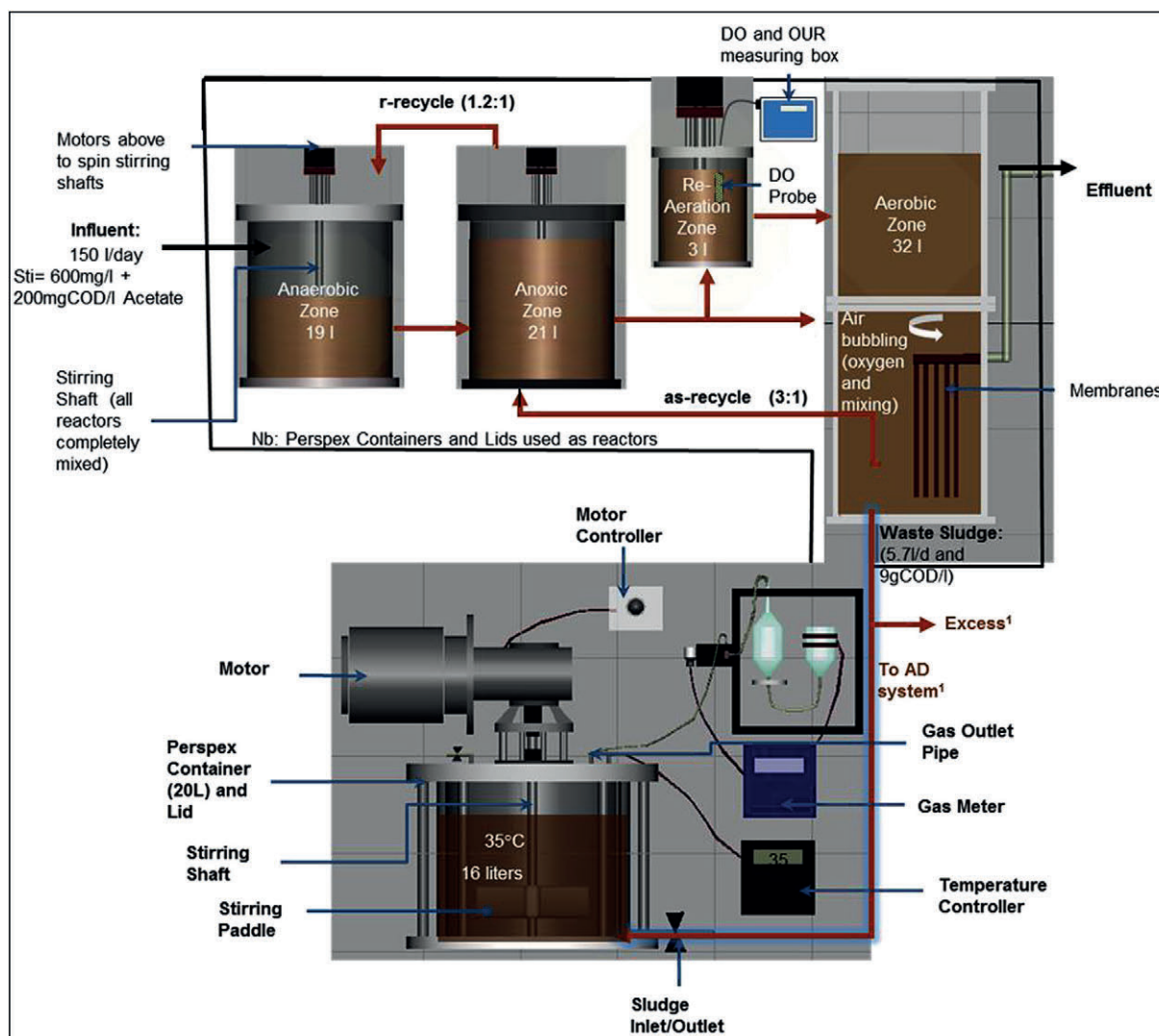


Figure 1. Experimental set-up used to carry out research investigation

inter-phase transfers of component species.

- The ASM2-3P model: This is the Activated Sludge Model No. 2 (ASM2, Henze et al., 1995), modified to include the ionic speciation model (Brouckaert et al., 2010), the inorganic settleable solids (ISS) model of Ekama and Wentzel (2004) and including multiple mineral precipitation according to Musvoto et al. (2000a,b).
- The SDM3P Model: This is the University of Cape Town Sludge Digestion Model 1 (UCTSDM1 (Sötemann et al., 2005)), modified to include the hydrolysis of multiple organic sludge types (primary sludge (PS), NDEBPR WAS and PS-WAS blends), the Ekama and Wentzel (2004) ISS model and the Brouckaert et al. (2010) speciation model which facilitates ionic speciation and multiple mineral precipitation.

The model evaluation procedure presented by Ikumi et al. (2015) was based on the BIOMATH protocol developed by Vanrolleghem et al. (2003). The complete description of the model (including components, processes and parameters) is reported in a previous publication of Ikumi et al. (2015). However, the new mass balanced stoichiometries of PP release (with the uptake of PHB and death of PAOs) that are used towards extension and modification of this model are evaluated below using the data from Ikumi (2011).

PAO behaviour in phosphorus removal unit processes

During the anaerobic phase of EBPR, the electrogenic release of PP occurs in conjunction with the production of protons to generate a proton motive force that energises the uptake of acetate to form poly-3-hydroxybutyrate (PHB) (Van Veen et al., 1994). Noting that PHB is more of a reduced compound than acetate, it has been proposed that the reducing power for its synthesis is mainly from intracellularly stored glycogen (Mino

et al., 1994; Smolders et al. 1995; Pereira et al., 1996) and also possibly from nicotinamide-adenine dinucleotide (NADH₂) formed by the oxidation of acetate to carbon dioxide (CO₂) (Comeau et al., 1985; Wentzel, 1988). However, during AD treatment of waste sludge (containing PAOs) from the aerobic reactor of the AS system the question arises: are the PAOs capable of carrying out the same P-release mechanisms in the AD as in the anaerobic reactor of the NDEBPR AS system? Essentially, microorganisms are assumed incapable of planning their actions based on a WWTP system configuration, and are expected to act according to the capabilities afforded to them by their surrounding environmental conditions. Thus, PAOs containing polyphosphate (PP) sent into an anaerobic digester with volatile fatty acids (VFAs) present, would utilize their PP reserves as they would in the anaerobic zone of an AS process. The PAOs are also expected to behave similarly in the absence of electron acceptors as regulated by the intracellular adenosine triphosphate (ATP), NADH₂ and acetyl-CoA levels (or related ratios). This suggests that there is some potential for energy transfer from the AS to the AD via PP when WAS is treated in AD systems. However, the quantity of PP used as an energy source is known to vary according to the balance between energy generation and consumption in the cell (Smolders et al., 1995; Mino et al., 1994), a mechanism that is linked to the system pH – Smolders et al. (1995) observed changing P/acetate ratio (0.25 to 0.75) for varying pH ranges (5.5 to 8.5). However, the effect of pH on the EBPR stoichiometry is yet to be included in the current models.

After anaerobic P release and PHB uptake, the PAOs require a terminal electron acceptor (oxygen, usually supplied in the aerobic reactor of AS system) to utilise this PHB for growth and energy generation metabolism. However, oxygen is not available in AD and, in this case, the PAOs continue to release over time all remaining stored PP and PHB. From anaerobic batch tests on the NDEBPR WAS, Harding et al. (2010) noted that it took 5 to 8 days to release practically all the PP under anaerobic conditions. Whether this is faster than the death rate of the PAOs in the AD is uncertain, but it is significantly faster than AS biomass hydrolysis rate in the AD. Therefore, in the AD, both the PP release with PHB formation (as in the anaerobic reactor of the NDEBPR system) and PP release with the eventual death of PAOs takes place.

Ikumi (2011) explains the steps for the development of a generalised stoichiometry to model the release of PP, with the uptake of acetate (HAc), in the AD system. This includes:

- The use of enzymes to degrade PP (via hydrolysis), in the presence of adenosine diphosphate (ADP), for the release of phosphate and production of adenosine triphosphate (ATP, formed as an intermediate for utilization in energy-consuming processes and to drive some enzyme-controlled reactions) (Cole and Hughes, 1964; Smolders et al., 1995).
 $(\text{polyphosphate})_n + \text{ADP} \rightarrow \text{ATP} + (\text{polyphosphate})_{n-1}$
 Then $\text{ATP} + \text{H}_2\text{O} \rightarrow \text{ADP} + \text{H}_2\text{PO}_4^-$, ultimately:

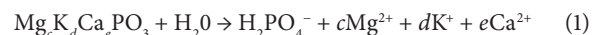


Table 1. Design and operating parameters for UCT AS system

Parameter	Value
Sludge age (d)	10
Influent COD (mg/L)	600+200 ^a
Influent flow (L/d)	150
Waste flow (L/d) (from aerobic reactor)	5.74
Volume (L)/Mass fractions: Anaerobic	19; 0.133
Volume (L)/Mass fractions: Anoxic	21/0.275
Volume (L)/Mass fractions: Aerobic	35/0.592
Recycle ratios: a (aerobic to anoxic)	2.8–3.4 ^b
Recycle ratios: s (from settling tank)	---
Recycle ratios: r (anoxic to anaerobic)	1.1–1.2 ^b
HRT – nominal/actual (h): Anaerobic	3.04/1.41
HRT – nominal/actual (h): Anoxic	3.36/0.61
HRT – nominal/actual (h): Aerobic	5.6/1.27

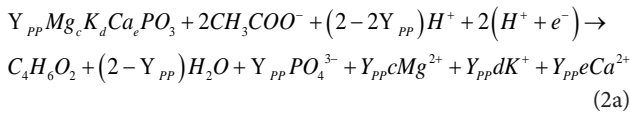
^aDosed 200 mg COD/L sodium acetate from 2.8 to 3.4 and s recycle varied from 1.1 to 1.2 during its operation

^bThe MBR UCT a recycle varied

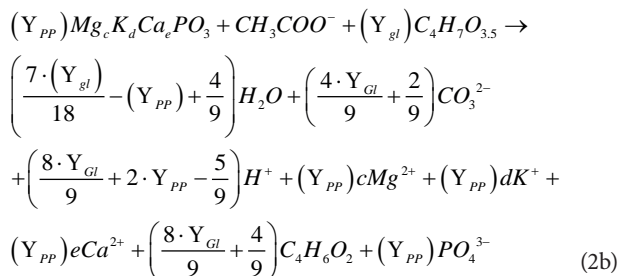
Table 2. Design and operating parameters for AD system

Test period	1	2	3	4	5
Period dates	1 Feb–08 Apr	9 Apr–04 Jul	1 Feb–12 Jun	5 Jul–28 Aug	28 Aug–2 Nov
Period duration	68	87	133	55	66
AD sludge age	18 d	40 d	60 d	25 d	10 d
Flow (L/d)	0.89	0.4	0.08	0.64	1.6
Flux (gCOD/d)	8	3.6	0.72	5.76	14.4

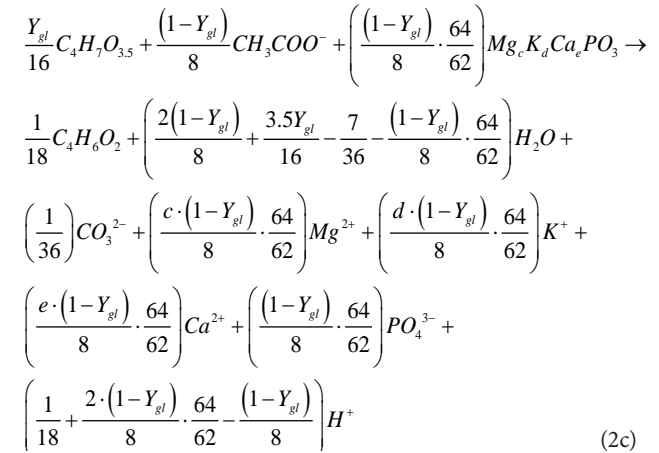
(ii) Anaerobic uptake of acetate (using the produced ATP, with energy) and its conversion to acetyl-CoA, which is subsequently converted to PHB with the use of NADH₂ (formed as NADH-H⁺, when NAD⁺ molecules take up two electrons and two H⁺ atoms), as reported by Smolders et al. (1995) and also confirmed by Mino et al. (1994). Equation 2a provides a general reaction for anaerobic PP release:



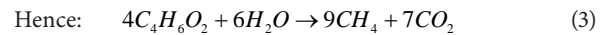
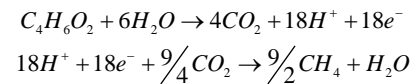
Where Y_{PP} is the mols of P released per mol of PHB formed. The uptake of H⁺ ions together with the electrons, during PHB formation, allows the PP release process to moderate anaerobic pH changes that occur with the release of orthophosphate (OP). The working AD pH is usually close to 7; hence mostly dissociated acetate is used in the AD PHB formation process. According to Wentzel et al. (1990), at a pH maintained at ≥ 7.25 , 2 mg COD acetate is taken up per 1 mg P of PP released. Therefore, 2/64 mol of acetate is taken up per 1/31 mol of PP (hence for 1 mol of acetate taken up 64/(2 × 31) mol of P in PP is released). Since, in the tricarboxylic acid (TCA) cycle, 2 mol of acetate are used to form 1 mol PHB, 2 × (64/62) mol of P are released in the process. This is close to the observation of Smolders et al. (1995), where at a pH of 7 about 0.5 mol of P in PP are used to provide 0.5 mol ATP to form 1 mol of C in PHB. Hence, for 4 C mol of PHB (i.e., 1 mol PHB), 2 mol ATP are provided by the PP released (which is 2 mol P of PP). However, apart from the possibility of the acetate (anaerobically available) being used as the source of some of the H⁺ and e⁻ for the PHB formation, Smolders et al. (1995) report that the main source of these electrons (and H⁺) is the aerobically generated glycogen (i.e., to act as a reducing agent). Since acetate uptake is linked to P release (for energy generation) and glycogen consumption (which maintains the redox balance), the quantity of glycogen available in PAOs becomes an important factor during the anaerobic PP release. With the PP release process taken as the reverse of its aerobic manufacture process, the energy content of PP could be determined to quantify the equivalent quantity of glycogen that is available for use in place of this energy. Aerobically, the PHB is used as the electron donor towards (i) anabolism of new PAO biomass, (ii) manufacture of glycogen, (iii) catabolism for the generation of energy, and (iv) uptake of P to form polyphosphate. Because both the PAO biomass and the glycogen stored in their cells are essentially organic components of PAO, we have to account for the quantity of glycogen that makes up part of the PAO organic mass in our models. This is such that $Y_{biomass} + Y_{glycogen} = Y_{PAO} = 0.66$, where $Y_{glycogen}$ and $Y_{biomass}$ are the yield coefficients for PHB converted to glycogen and biomass of PAOs, respectively. Smolders et al. (1995) estimated the yield value of PAO biomass (without the glycogen) to be about 13% lower than that of usual heterotrophic biomass. However, for the equivalent donating capacity of 2e⁻ and 2H⁺, an extra 1/8 mol of glycogen (C₄H₇O_{3.5}) is utilised in the process and less alkalinity gets generated.



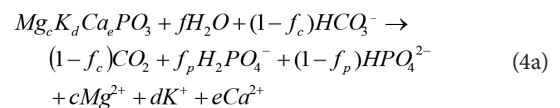
Where Y_{gl} is the mols of glycogen used per mol of PHB formed, noting that the electrons of glycogen represent energy in PP transferred from the AS to AD. Since the molar electron donating capacity for glycogen, acetate and PHB is 16, 8 and 18, respectively, and 1 mol of acetate uptake releases 64/62 mol PP, then $(1 - Y_{gl})/8$ mol acetate uptake results in $(1 - Y_{gl}) \cdot (64/62)$ mol PP release. This is used in formulating the PP release and PHB formation stoichiometry below:



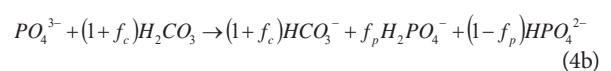
(iii) In the AD environment, the PAOs are not modelled to compete with the AD biomass for the acetate, but after an initial rapid uptake of acetate and release of PP, leave the remaining acetate to be degraded by AD biomass and give up the PHB with their hydrolysis. Thus, the PHB produced from the reaction of Eq. 2 is later broken down in AD, to form carbon dioxide together with some H⁺ and e⁻, as shown (Eq. 3). The products of this breakdown contribute (as with other biodegradable organics) towards AD biomass growth and energy (methane) generation:



(iv) In the AD, the remaining PP that doesn't get released initially with PHB formation also eventually gets hydrolysed (as shown in Eq. 1a), since the dead PAO biomass cannot hold on to it further. In the aerobic digester, Vogts et al. (2014) observed the PP release with the death/endogenous respiration of PAO biomass to be slow (at a similar rate to the endogenous respiration process). However, in the AD the remaining PP is degraded faster than in the aerobic digester. Harding et al. (2010) provide a steady-state stoichiometric equation for this process:



Where the f value fractionates the total phosphates to HPO₄²⁻ and H₂PO₄⁻, according to pH, as shown by Harding et al. (2010), i.e.:



It was decided that both the PP release mechanisms in AD, i.e., with PHB formation (Eq. 2a, b and c) and with PAO

death (Eq. 1), together with PHB disintegration (Eq. 3), be entered into the dynamic model because all have a possibility of occurrence. However, this may result in complexities for the calibration of their kinetics.

Ikumi (2011) accounts that the final outcome, for an AD system operated at steady state, is equivalent to the PP hydrolysis occurring with the death of PAOs, as was reported by Harding et al. (2010) (see Eqs 4a and 4b). This is because in the AD model, the PHB formed eventually gets completely degraded, hence all the COD removed in AD is ultimately destined for conversion to an inevitable quantity of AD biomass and biogas (some electrons go towards CH_4), irrespective of the defined form of the biomass (i.e., the active PAOs are modelled with or without inclusion of glycogen). The discrepancies observed during the calibration raised a few questions regarding the current stoichiometric model: Mainly (i) why is less glycogen produced than consumed? And (ii) is the energy in PP all represented by glycogen (organic material that is measurable as COD) or is there a possibility of it not being measurable as COD in the anaerobic reactor (i.e., NADH_2 formed aerobically from PHB degradation)? The breakdown of an organic component such as glycogen or acetate, to act as the reducing agent in the anaerobic phase, results in CO_2 production. With the partial pressure of CO_2 maintained according to Henry's law expression (Loewenthal et al., 1994),

CO_2 gets dissolved into the aqueous phase resulting in a lower pH prediction (see results section below). Is there a possibility of lower CO_2 solubility in the AD to stabilise the pH accordingly? This may be important to explore as the pH prediction is significant for AD of P-rich sludge, due to the potential for the occurrence of further processes such as mineral precipitation.

The flow diagram (Fig. 2) provides tracking of PP in the aerobic and anaerobic environments. In aerobic P uptake, ATP is required to make the PP, indicating that PP contains energy, which came from the electron transport chain in aerobic conditions. This 'energy' in PP is equivalent to electrons donated from PHB to O_2 in the catabolic PP uptake process – instead of being lost as heat this energy is used in linking orthophosphate to the PP chain (Kortstee et al., 2000). The electrons used here are accounted for by the oxygen used in the aerobic process. The fraction of COD in PHB used to store PP is defined by a parameter Y_{PHA} . For a plant-wide model, the COD of PHB used to form PP (that is released in the anaerobic digester) is accounted for by deducting from the flux of oxygen used aerobically (FO_2), the Y_{PHA} fraction of the PHB used for aerobic storage of PP. In the anaerobic reactor (or digester) this energy (released with PP hydrolysis) is used to reclaim electrons from the oxidation of water during the PP release process (Eq. 1), to HAC to make 1 PHB ($\text{C}_4\text{H}_6\text{O}_2 = 8e^-$).

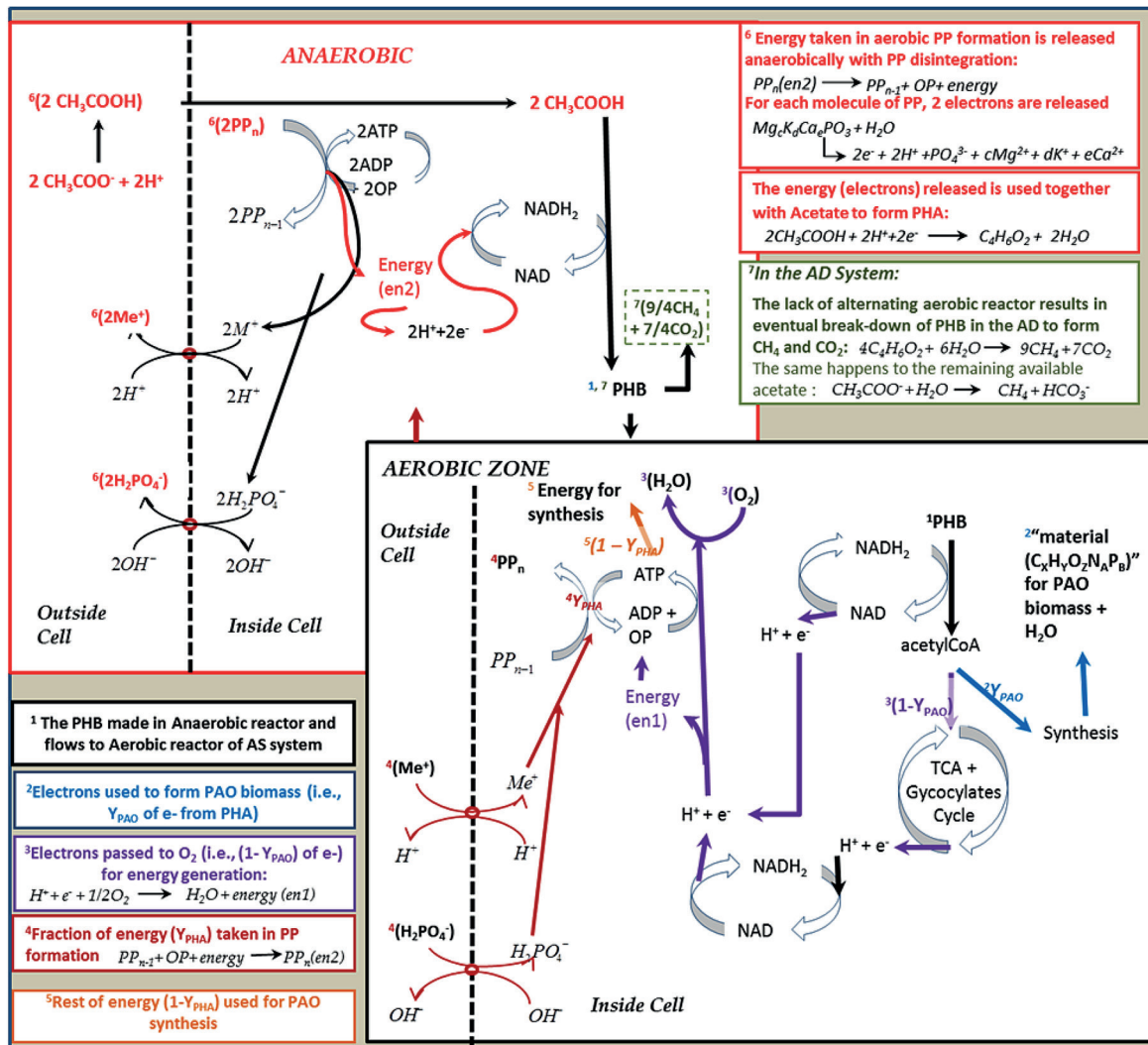


Figure 2. Schematic representation of the aerobic PHB utilisation and anaerobic PP release for a AS-AD linked WWTP

EXPERIMENTAL AND MODELLING RESULTS

Figure 3 shows the graphs resulting from the progressive validation of the three-phase (aqueous-gas-solid) ASM2-3P model, through a comparison of model simulation results to data generated from the experimental investigation by Mebrahtu et al. (2007), Ikumi (2011) and Vogts et al. (2014), who operated the UCT AS system from which WAS was abstracted and fed to the AD of the experimental set-up (see Fig. 1).

As exhibited in Figs 3a and b, the ASM2-3P models the behaviour of the parent UCT NDEBPR system well by predicting values close to experimental observation for: mixed liquor COD (includes all biomass and other waste organics), ISS (includes PP and reactor-accumulated influent ISS; mineral precipitation was not observed here, hence the model correctly excludes precipitates from the ISS, which, incidentally, also validates the ISS model of Ekama and Wentzel (2004) that is included in the extension of the ASM2-3P model), OP (for organics and P removal shown in Fig. 3a), total Kjeldahl nitrogen (TKN) and FSA (for N removal shown in Fig. 3b). This is important for plant-wide modelling because the PP is the main source of orthophosphate (OP) and metals (Mg, Ca and K) in the subsequent WAS treatment systems (AD or aerobic digestion), and hence determines the extent to which precipitation and other weak acid/base processes take place (Ikumi et al., 2015).

Figure 4 shows a comparison between the results simulated by the three-phase AD dynamic model (SDM3P) and AD data measured with changing AD sludge age (10, 18, 25, 40 and 60d) for the AD of the NDEBPR WAS. Also, two different modelling

approaches are compared here (i.e. using simulated predictions of Model 1, where P release occurs with PHB formation by PAOs as shown in Eq. 2c and Model 2, where P release occurs without PHB formation, as shown in Eq. 1).

Apart from COD removal (where all simulated results match measured ones well), the results predicted by the SDM3P simulation model (Fig. 4a to f) do not yet have a good correspondence to those measured (as mentioned earlier, this has resulted in various questions regarding the approach to be taken in modelling the process). In all cases, the FSA release shows a similar trend to COD removal, indicating its steady release as the biodegradable organics of the sludge is hydrolysed (the kinetics of hydrolysis initially having been calibrated by Ikumi et al., 2014). The biodegradable organic (i.e., the active ordinary heterotrophic organism (OHO) and phosphorus accumulating organism (PAO) biomass) sludge hydrolysis, results in the release of organically bound N, which is in the non-ionic NH_3 form and is in the form of non-reference species for the ammonia weak acid/base system. On release, the NH_3 picks up an H^+ ion from the bulk liquid which is supplied (in part) by the dissolved CO_2 (H_2CO_3^*) forming HCO_3^- i.e.:



However, for AD of EBPR WAS the total alkalinity generation not only depends on the release of organically bound N but also on how PP is released.

The initial PP release in the AD occurs with the uptake of acetate and results in increased alkalinity by the addition of H_2PO_4^- (see Eq. 2c). The eventual PAO death and hydrolysis in

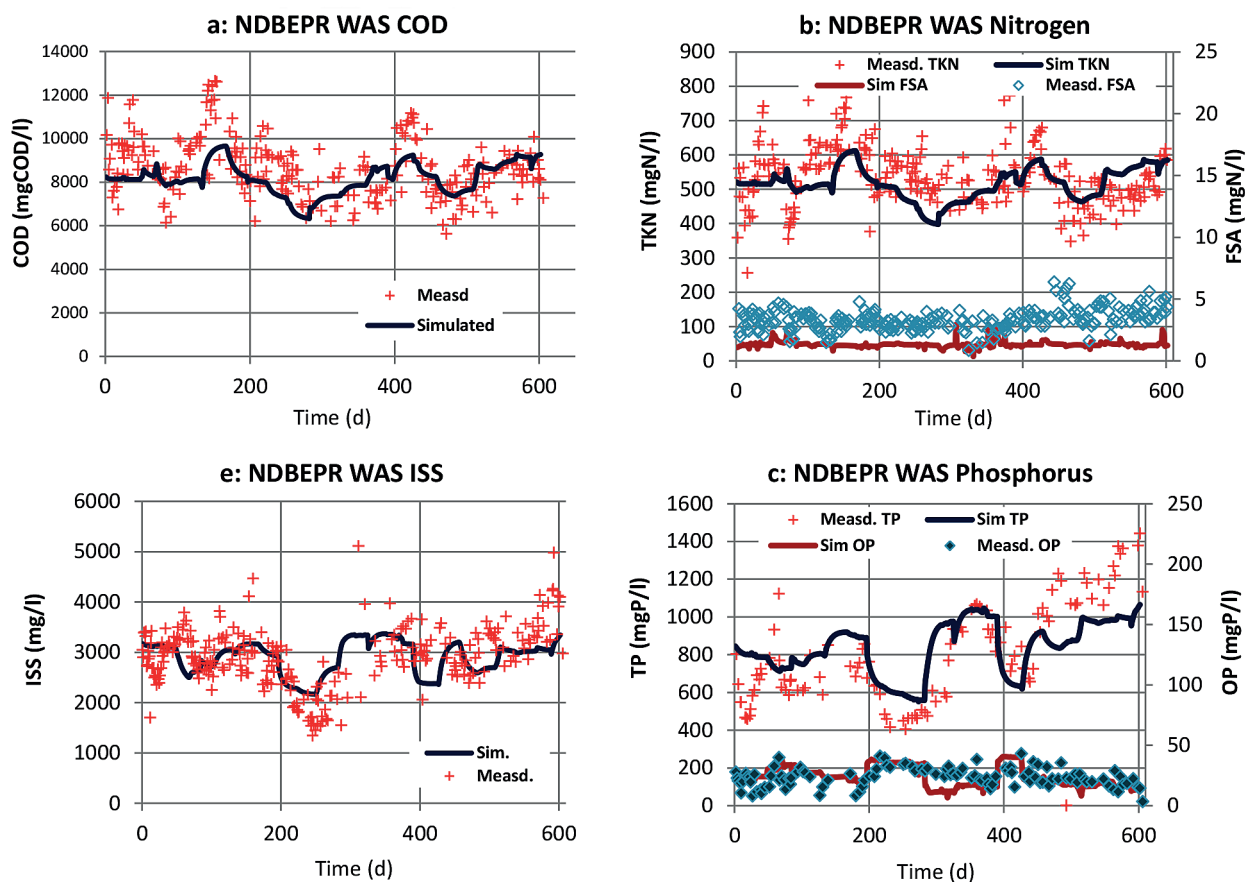


Figure 3. Simulated results for the parent NDEBPR AS system that is linked to the AD system in the plant-wide experimental set-up

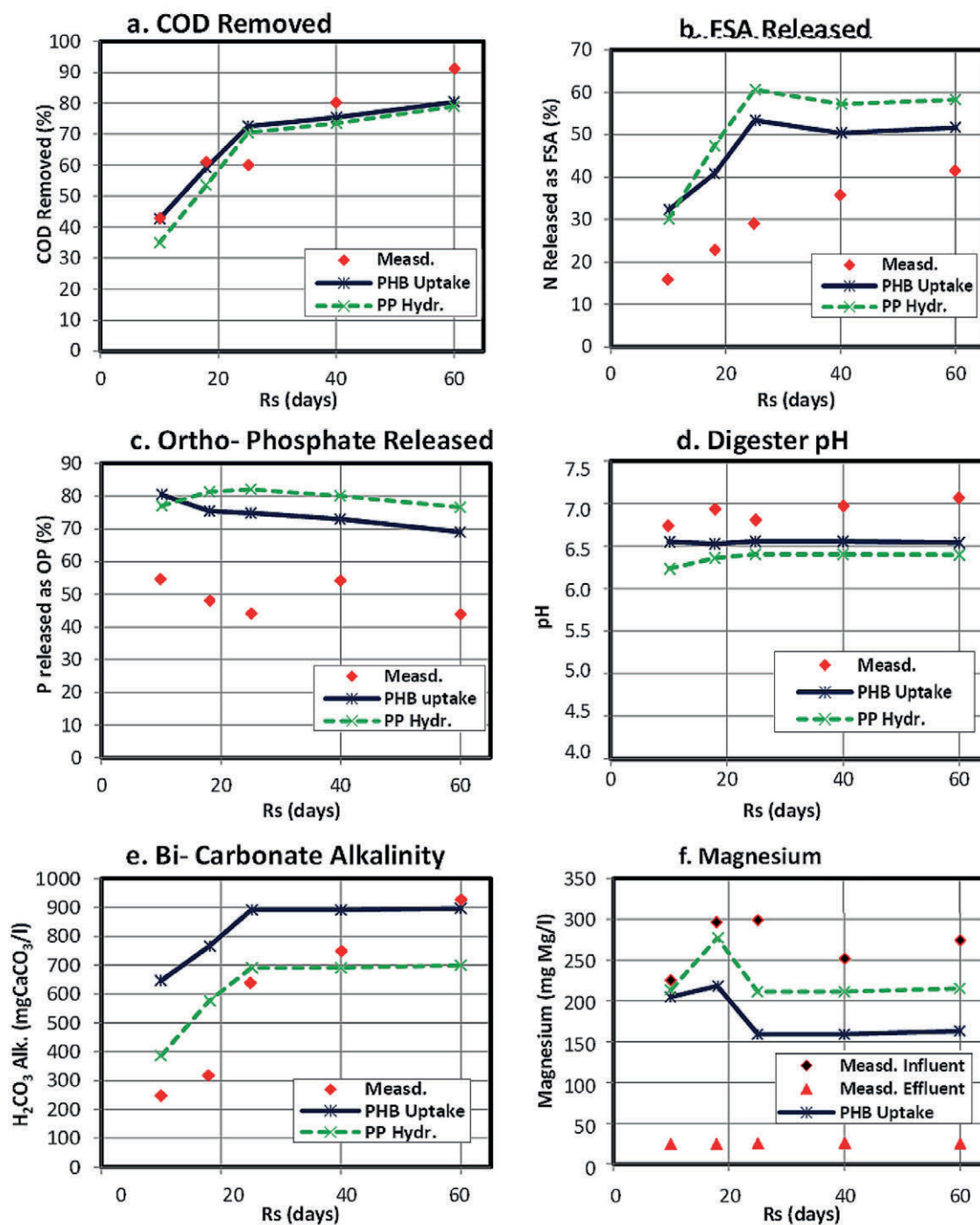
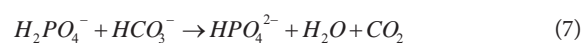


Figure 4 (a to f). Simulated results for the AD system treating WAS taken from the aerobic reactor of the NDEBPR AS system. The model predictions include simulations where PP is hydrolysed without any PHB formation (PP Hydr.; see Eq. 1) and where PHB formation occurs with an organic substance (i.e., glycogen or acetate – both would yield the same result) used as the reducing agent (see Eq. 2c).

the AD (due to the lack of terminal electron acceptor, oxygen, to allow for their growth) results in the subsequent release of the remaining PP and other stored PAO components such as PHB. This release of stored PHB and PP is also modelled to occur independently at the PAO hydrolysis because there is not enough information to know the manner by which disintegration of the stored PP and PHB occurs after exposure of the biomass cytoplasm.

The PP release upon PAO hydrolysis (see Eq. 3a) results in, depending on the charge/proton balance requirements, some of the

$H_2PO_4^-$ species becoming HPO_4^{2-} species while keeping the H_3PO_4 alkalinity constant. This release of a proton from $H_2PO_4^-$ results in increased p_{CO_2} and decreased pH – a proton is added to HCO_3^- producing H_2CO_3 , which then gets released to the atmosphere as CO_2 when the Henry's equilibrium constant is exceeded.



Since with PP release, there is lowered pH and no net alkalinity change, the reaction releases $2H^+$ with each PO_4^{3-} ; however the PO_4^{3-}

is approximately equally divided between HPO_4^{2-} and H_2PO_4^- at pH 7, so it effectively releases 1H^+ per PO_4^{3-} to the solution.

Connected to (and hence at the same rate as) hydrolysis of 'dead' PAO (and OHO) biomass, organic P is released as H_3PO_4 . Because this is the reference species for the OP weak acid/base system, this organic P release does not cause a change in total alkalinity. However, in the AD pH range 7 to 8, the H_3PO_4 species lose 1 or 2 H^+ to become H_2PO_4^- and HPO_4^{2-} species. As shown in Eq. 7 this H^+ reacts with HCO_3^- to become H_2O and CO_2 . So while the total alkalinity does not change, the species that represent it do, and the CO_2 that would have been retained in the aqueous phase as HCO_3^- now exits the AD as gas, which increases the ρCO_2 (the CH_4 gas production remains unchanged because it is fixed by the COD of the biodegraded organics).

Therefore, accurate prediction of total alkalinity, ρCO_2 and pH requires accurate estimation of the N and P content of the PAO (and OHO) biomass, the PP quantity stored and the correct rates of PP release (both with PHB uptake and after PAO death) and biomass organics hydrolysis. If a limited quantity of acetate is generated from partial hydrolysis (e.g., due to low AD sludge age) and subsequent fermentation of biodegradable organics, the PP release with PHA uptake may be interrupted, allowing more PP hydrolysis with PAO death.

The rapid breakdown of PP yields high concentrations of P, Mg^{2+} , K^+ and C^{2+} species in the AD mixed liquor already at short sludge ages (< 10 d; see Fig. 4c). This, together with the increased release of ammonia with sludge age, results in struvite ($\text{MgNH}_4\text{PO}_4 \cdot 6\text{H}_2\text{O}$) precipitation potential. This struvite precipitation removes PO_4^{3-} species from the aqueous phase, hence decreasing the total alkalinity and resulting in re-speciation of the inorganic carbon and other weak acid/base systems to cause a further decrease in ρCO_2 and AD pH (Loewenthal et al., 1994).



The measurements of total and soluble K in the influent and effluent suggest the complete (or almost complete) release of PP for all residence times; in that the effluent dissolved K is very close to the effluent total K. Although the K from the released PP does not precipitate, the Mg has the potential

for precipitate formation – this is not accounted for well in the model simulations of Fig. 4. The influent soluble Mg concentrations are around 80 mg/L, while the effluent concentrations are remarkably constant at around 25 mg/L – this in spite of the Mg released from PP. The constancy suggests equilibrium with a precipitate, but none of the Mg precipitates in the model appears to be able to reduce the Mg to that level, due to low pH prediction.

As can be noticed in Figs 4a–f, both the simulation of PP release without considering PHB formation (i.e., using stoichiometry from Eq. 1) and with PHB formation, where an organic compound is utilised as a reducing agent (see Eq. 2c), results in underprediction of the system pH, and consequent miscalculation of precipitation potential (low pH does not support the formation of struvite and other mineral precipitates). This miscalculation of precipitation potential (due to low pH prediction) results in the poorly predicted AD effluent FSA, OP and metals (Mg, K and Ca) in the aqueous phase. If the PP release is modelled according to Eq. 2a (with the e^- and H^+ sourced from NADH_2 formed aerobically with PHB degradation in the parent NDEBPR AS system), better predictions for pH are obtained due to improved accuracy of calculated precipitation, AD FSA, OP and metals (see Fig. 7a–d). This suggests the uptake of H^+ ions is an aspect requiring further investigation. It has been observed that the quantity of PP released to energise PHB formation is dependent on the system pH, rather than vice versa. With PHB formation maintained constant, an increased reactor pH would result in more PP degraded (Smolders, 1995). Moreover, greater certainty of predicted ionic strength would be a valuable addition to the calibration exercise (lacking in this dataset) – due to the increased concentration of metals and phosphates released. Ionic strength and ion pairing affect mineral precipitation calculation much more strongly than pH calculation (Tait et al., 2012) especially for a trivalent ion like PO_4^{3-} .

Apart from the polyphosphate release stoichiometry, other possible contributions to explore for the under-predicted pH from Fig. 4d include:

- Formation of precipitates in the feed: All the precipitation reactions in the models are favoured by higher pH values, but lower the pH when they occur in the anaerobic reactor. Conversely, if precipitates are present in the feed, they will

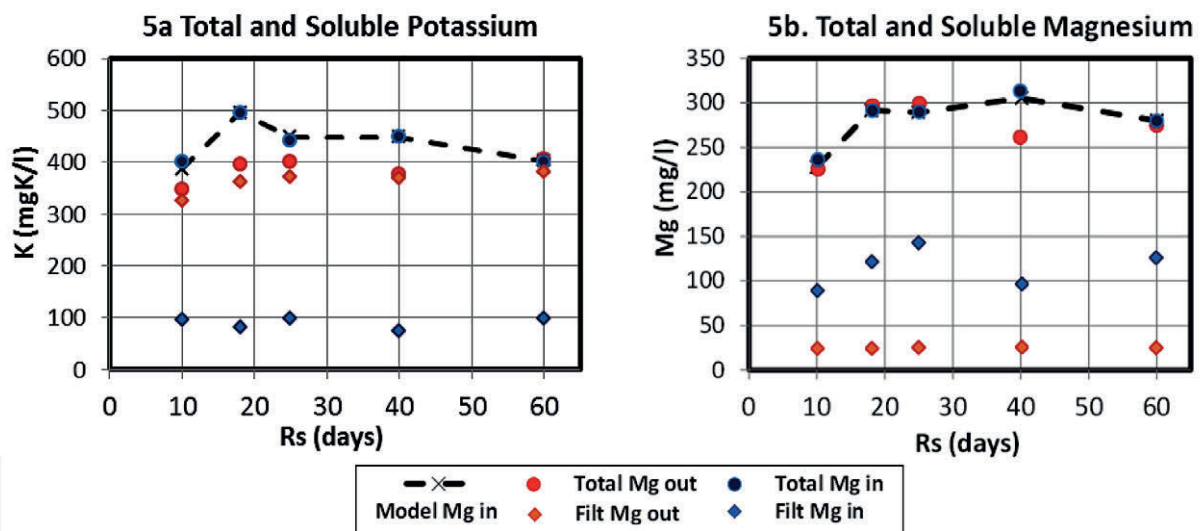


Figure 5a and 5b. Measured results for the total and soluble potassium (K; Fig. 5a) and magnesium (Mg; Fig. 5b) in the influent and effluent of the AD system treating WAS taken from the aerobic reactor of the NDEBPR AS system

tend to dissolve if the pH drops, and act as a pH buffer. They would provide a reservoir of alkalinity which is not detected by a 5-point titration (Moosbrugger et al., 1993) performed on liquid separated from particulates. This has implications for the ISS in the feed. From the experimental data, it is evident that the influent ISS varies considerably, whereas the effluent

ISS is much more consistent. In the model, the ISS is the sum of the PP and the model ISS component. To properly calibrate PP release would require experimental data that characterises the ISS towards the location of the particulate phosphorus (i.e., whether the ISS is PP or a mineral precipitate such as struvite). Hence it is recommended in the future project to

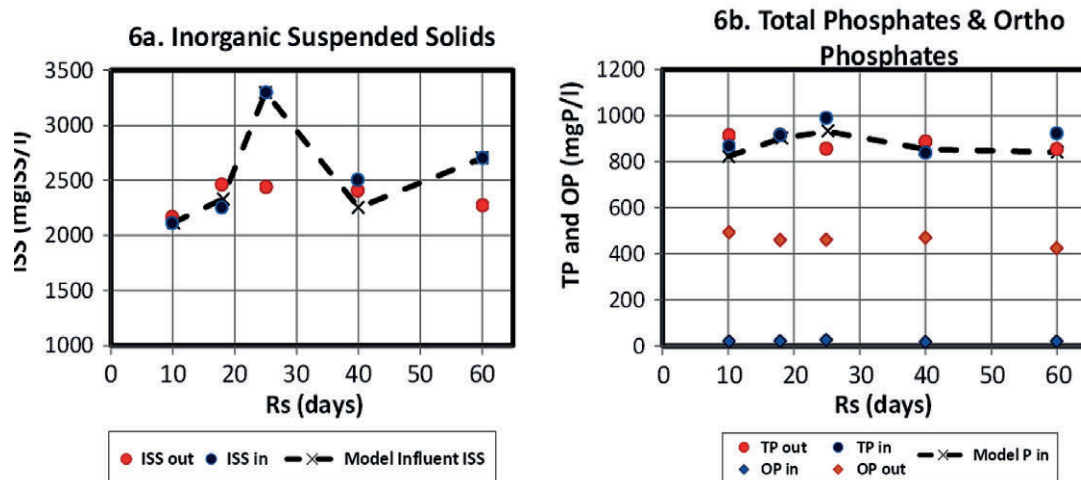


Figure 6a and 6b. Measured results for the Inorganic suspended solids (ISS; Fig. 6a) and phosphates (TP and OP; Fig. 6b) in the influent and effluent of the AD system treating WAS taken from the aerobic reactor of the NDEBPR AS system

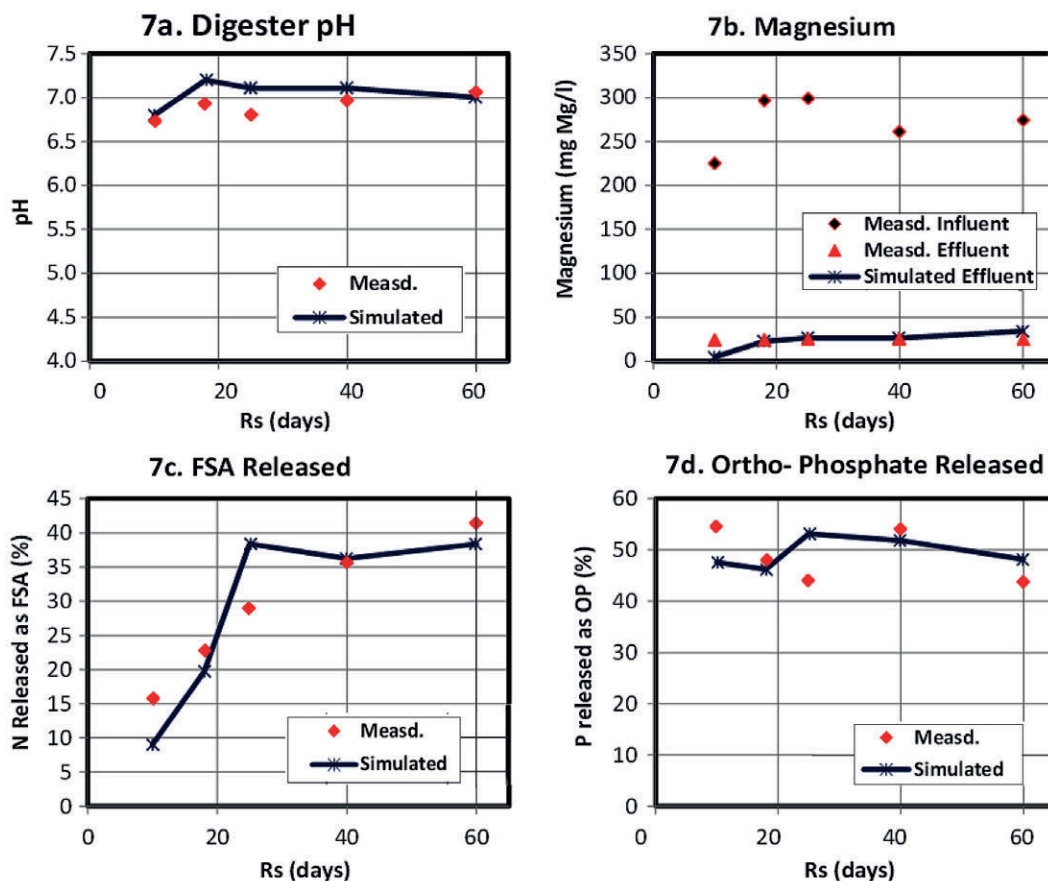


Figure 7 (a to d). Simulated results for the AD system treating WAS taken from the aerobic reactor of the NDEBPR AS system. In this case the model was adjusted, such that glycogen was not used as a reducing agent but was replaced by the sourcing of electrons from the PHB aerobically and the energy (electrons) carried internally as NADH (i.e., the reaction gets partially done in the aerobic phase). The COD results (omitted) match those of Fig. 4, but the pH prediction is higher, allowing for the mineral precipitation to occur as observed experimentally.

test for conditions (pH and temperature) that would allow for a diluted sample that achieves re-dissolution of mineral precipitates (change from solid to aqueous phase) without altering polyphosphates (stay in solid phase).

- ii. Potential for lower CO₂ solubility in the AD (e.g., reduction in the CO₂ partial pressure due to the presence of inert gas such as nitrogen; N₂) resulting in further CO₂ transfer from aqueous to the gaseous phase. The evolution of CO₂ would cause a decrease in acidity and an increase in the system pH.

The AD of WAS-containing enhanced cultures of PAO

The findings from this investigation have resulted in a few research questions, regarding the extent to which existing mathematical models virtually represent the AD of WAS from EBPR WWTP systems. To allow for more precise tracking of material components (i.e., C, H, O, N, P, metals, ISS, charge and energy) through the WWTP, in research or practical applications, necessitates having a definitive description of the biochemical mechanisms of PAO enhanced cultures. Towards this end, further investigation is required into (i) the extent to which PP is released with PHB formation (resulting in energy generation, alkalinity increase and ρCO_2 decrease) rather than PAO death (results in total alkalinity decrease and ρCO_2 increase), and (ii) the endogenous mass fraction of PAOs. The parent activated sludge experimental set-up is similar to that of Wentzel (1988), i.e., a continuous anaerobic-aerobic system fed volatile fatty acids, sufficient orthophosphate (to avoid P limitation) with the required quantity of mineral supplements (including Mg, K and Ca for synthesis of PP) and ammonia (just enough for biomass growth, hence avoiding anoxic conditions that would result from nitrification). Further, the waste sludge from this EBPR system will be fed to anaerobic digestion units and batch tests to generate data towards calibrating the maximum specific rate of anaerobic PHB uptake. These AD batch tests on enhanced cultures of PAO biomass may also be useful to investigate further the process of PAO death in AD for validating the kinetics of PAO hydrolysis and the endogenous residue fraction (the AD effluent is essentially PAO endogenous residue). In the AS system, the PAOs' endogenous respiration rate is known to be at about 0.04/d, with the endogenous residue being about 0.25 of the PAO biomass (Wentzel et al., 1990). In the AD system, it is likely that the PAOs die at a faster rate, since the AD biomass is acting on them as a substrate source. However, rather than their death rate, of importance is the rate at which their biodegradable particulate organics hydrolyse. Ikumi et al. (2014) report the biodegradability and kinetics of hydrolysis of WAS from NDEBPR AS system, as similar to that of primary sludge, but this was obtained from the AD of WAS containing mixed cultures of OHO and PAO biomass.

Within this same experimental set-up, the PHB formation rate can also be calibrated by testing the quantity of PHB formed (with PP release) at changing anaerobic reactor retention times. It is expected that if this PHB formation rate can also be used in the AD (since PAOs are again exposed to anaerobic conditions with the presence of acetate) to provide reasonable outputs for pH and ρCO_2 , this would validate the concept of PP contributing to some 'energy' carry over from the AS to AD system.

CONCLUSIONS

The behaviour of PAOs in anaerobic digestion was investigated but inadequacy in the experimental data prevented making

definitive conclusions. However, the results do indicate that PAO behaviour may influence AD pH in ways that OHO do not.

- i. In AS systems the stoichiometry of PP-accumulating organisms includes the processes of aerobic P uptake to form PP and anaerobic P release from this PP. The energy for aerobic PP formation is sourced from poly3-hydroxybutyrate (PHB) breakdown, with oxygen used as the terminal electron acceptor. Hence this oxygen used represents electrons transferred from the PHB and allows for the COD balance to be made over the activated sludge (AS) system. However, the 'energy' acquired in this PHB breakdown and PP uptake process is available in the PP when it is reused in the anaerobic reactor of the EBPR system. The energy within this PP allows for the anaerobic formation and storage of PHB using acetate in AS systems.
- ii. When the WAS from parent EBPR AS systems is fed to an anaerobic digestion (AD) system, it is expected that a similar process of PHB formation with PP release occurs (due to the combination of anaerobic conditions and presence of acetate), as for the AS system. However, this PHB eventually gets broken down to methane and carbon dioxide, since there is no alternating aerobic reactor for its utilisation in PAO growth.

Acronyms and abbreviations

AAD	Anoxic-aerobic digestion
AD	Anaerobic digestion
ADP	Adenosine diphosphate
AS	Activated sludge
ATP	Adenosine triphosphate
BNR	Biological nutrient removal
COD	Chemical oxygen demand
EBPR	Excess biological phosphorus removal
FO _c	Flux of oxygen used aerobically
HAc	Undissociated acetate (CH ₃ COOH)
ISS	Inorganic settleable solids
KH ₂ PO ₄	Di-potassium hydrogen phosphate
MMP	Multiple mineral precipitations
NAD ⁺	Nicotinamide-adenine dinucleotide
ND	Nitrification-denitrification
OHO	Ordinary heterotrophic organism
OP	Orthophosphates
PAOs	Phosphorus accumulating organisms
ρCO_2	Carbon dioxide partial pressure
PHB	Poly3-hydroxybutyrate (C ₄ H ₆ O ₂)
PP	Polyphosphate (Mg ₂ K ₃ Ca ₆ PO ₃)
SRT	Solids retention time (or sludge age)
TCA	Tricarboxylic acid
TKN	Total Kjeldahl nitrogen
Y _{biomass}	Yield coefficient for PAO converting PHB to new biomass (i.e. For growth) aerobically.
Y _{glycogen}	Yield coefficient for PAO converting PHB to glycogen aerobically.
Y _{gl}	Mols of glycogen used per mol of PHB formed anaerobically.
Y _{PHA}	Yield coefficient for PAO converting PHB to energy used in aerobic PP formation and storage
Y _{PP}	Mols of phosphorus released from polyphosphate for the formation of one mol of PHB
VFA	Volatile fatty acid
WAS	Waste activated sludge
WWTP	Wastewater treatment plant

- iii. When modelling the AD system in this way (as defined in (ii), we notice that better predictions are obtained (with acetate uptake and PHB formation included) than when modelling the AD PP release to occur with PAO death and hydrolysis. However, to validate this model, and explore the energy transfer between aerobic reactor of AS system and AD systems, requires further investigation on the AD of enhanced cultures of PAOs to amplify the PAO behaviour in the AD.

ACKNOWLEDGEMENTS

This research was supported by the Water Research Commission, the National Research Foundation and the University of Cape Town, and is published with their permission.

REFERENCES

- BATSTONE DJ, KELLER J, ANGELIDAKI I, KALYUZHNYI SV, PAVLOSTATHIS SG, ROZZI A, SANDERS WTM, SIEGRIST H and VAVILIN VA (2002) Anaerobic digestion model No 1 (ADM1), Scientific and Technical Report No 9. International Water Association (IWA), London, UK. <https://doi.org/10.2166/wst.2002.0292>
- BATSTONE DJ, AMERLINCK Y, EKAMA G, GOEL R, GRAU P, JOHNSON B, KAYA I, STEYER J-P, TAIT S, TAKÁCS I, VANROLLEGHEM PA, BROUCKAERT CJ and VOLCKE E (2012) Towards a generalized physicochemical framework. *Water Sci. Technol.* **66** (6) 1147–1161. <https://doi.org/10.2166/wst.2012.300>
- BROUCKAERT CJ, IKUMI DS and EKAMA GA (2010) A 3-phase anaerobic digestion model. *Proc. 12th IWA Anaerobic Digestion Conference (AD12)*, 1–4 November 2010, Guadalajara, Mexico.
- BRUN R, KÜHNI M, SIEGRIST H, GUJER W and REICHERT P (2002) Practical identifiability of ASM2d parameters—systematic selection and tuning of parameter subsets. *Water Res.* **36** 4113–4127. [https://doi.org/10.1016/S0043-1354\(02\)00104-5](https://doi.org/10.1016/S0043-1354(02)00104-5)
- COLE JA and HUGHES DE (1965) The metabolism of polyphosphates in chlorobiumthiosulfatophilum. *J. Gen. Microbiol.* **38** (1) 65–72.
- COMEAU Y, HALL KJ, HANCOCK REW and OLDFHAM WK (1985) Biochemical model for enhanced biological phosphorus removal. *Proc. UBC Conference on New Directions and Research in Water Treatment and Residuals Management*, June 1985, Vancouver, Canada. [https://doi.org/10.1016/0043-1354\(86\)90115-6](https://doi.org/10.1016/0043-1354(86)90115-6)
- DAIGGER TD, HENRY CL and LESLIE CP (1999) *Biological Wastewater Treatment*. (2nd edn.) Marcel Dekker Inc., New York, USA.
- EKAMA GA and WENTZEL MC (2004) A predictive model for the reactor inorganic suspended solids concentration in activated sludge systems. *Water Res.* **38** (19) 4093–4106. <https://doi.org/10.1016/j.watres.2004.08.005>
- EKAMA GA, WENTZEL MC and SÖTEMANN SW (2006) Mass balanced-based plant-wide wastewater treatment plant models – Part 2: Tracking the influent inorganic suspended solids. *Water SA* **32** (3) 277–285. <https://doi.org/10.4314/wsa.v32i3.5272>
- EKAMA GA (2009) Using bio-process stoichiometry to build a steady state plant-wide wastewater treatment plant model. *Water Res.* **43** (8) 2101–2120. <https://doi.org/10.1016/j.watres.2009.01.036>
- FLORES-ALSINA X, SOLON K, KAZADI MBAMBA C, TAIT S, JEPSSON U, GERNAEY KV and BATSTONE DJ (2016). Modelling phosphorus, sulphur and iron interactions during the dynamic simulation of anaerobic digestion processes. *Water Res.* **95** 370–382. <https://doi.org/10.1016/j.watres.2016.03.012>
- HARDING TH, IKUMI DS and EKAMA GA (2010) A steady state stoichiometric model describing the anaerobic digestion of biological excess phosphorus removal waste activated sludge. Research Report W132 (MSc thesis), Department of Civil Engineering, University of Cape Town, South Africa.
- HAUDUC H, RIEGER L, TAKÁCS I, HÉDUIT A, VANROLLEGHEM PA and GILLOT S (2010) A systematic approach for model verification: Application on seven published activated sludge models. *Water Sci. Technol.* **61** (4) 825–839. <https://doi.org/10.2166/wst.2010.898>
- HENZE M, GUJER W, MINO T, MATSUO T, WENTZEL MC and MARAIS GVR (1995) Activated sludge model No.2 (ASM2). IWA Scientific and Technical Report No. 3. IWA Publishing, London. <https://doi.org/10.2166/wst.1995.0061>
- HENZE M, VAN LOOSDRECHT MCM, EKAMA GA and BRDJANOVIC D (2008) *Biological Wastewater Treatment: Principles, Design and Modelling*. IWA Publishing, London. 620 pp. <https://doi.org/10.2166/9781780408613>
- IKUMI DS (2011) The Development of a Three-phase Plant-Wide Mathematical Model for Sewage Treatment. PhD thesis, Water Research Group (WRG). Department of Civil Engineering, University of Cape Town, South Africa.
- IKUMI D, VANROLLEGHEM PA, BROUCKAERT V, NEUMANN M and EKAMA G (2014) Towards calibration of phosphorus (P) removal plant-wide models. *Proc. 4th IWA/WEF Wastewater Treatment Modelling Seminar (WWTmod2014)*, 30 March–2 April 2014, Spa, Belgium. 197–206.
- IKUMI DS, HARDING TH, VOGTS M, LAKAY MT, MAFUNGWA HZ, BROUCKAERT CJ and EKAMA GA (2015) Mass balances modelling over wastewater treatment plants III. WRC Report No. 1822/1/14. Water Research Commission, Pretoria. <https://doi.org/10.1016/j.watres.2014.02.008>
- IKUMI DS, HARDING TH and EKAMA GA (2014) Plant-wide wastewater treatment modelling (1) – Biodegradability of wastewater and activated sludge organics in anaerobic digestion. *Water Res.* **56** (1) 267–279.
- JARDIN N and PÖPEL HJ (1994) Phosphate release of sludge from enhanced biological P-removal during digestion. *Water Sci. Technol.* **30** (6) 281–292. <https://doi.org/10.2166/wst.1994.0279>
- JEPSSON U, PONS M-N, NOPENS I, ALEX J, COPP JB, GERNAEY KV, ROSEN C, STEYER J-P and VANROLLEGHEM PA (2007) Benchmark simulation model No2: General protocol and exploratory case studies. *Water Sci. Technol.* **56** (8) 67–78. <https://doi.org/10.2166/wst.2007.604>
- KAZADI MBAMABA C, FLORES-ALSINA X, BATSTONE DJ and TAIT S (2015) A systematic study of multiple minerals precipitation modelling in wastewater treatment. *Water Res.* **85** 359–370. <https://doi.org/10.1016/j.watres.2015.08.041>
- LIZARRALDE I, FERNÁNDEZ-ARÉVALO T, BROUCKAERT CJ, VANROLLEGHEM PA, IKUMI DS, EKAMA GA, AYESA E and GRAU P (2015) A new general methodology for incorporating physicochemical transformations into multi-phase wastewater treatment process models. *Water Res.* **74** (5) 239–256. <https://doi.org/10.1016/j.watres.2015.01.031>
- LOEWENTHAL RE, KORNMULLER URC and VAN HEERDEN EP (1994) Modelling struvite precipitation in anaerobic treatment systems. *Water Sci. Technol.* **30** (12) 107–116. <https://doi.org/10.2166/wst.1994.0592>
- MEBRAHTU MK, WENTZEL MC and EKAMA GA (2007) Aerobic digestion of waste activated sludge from biological N and P removal systems. Research Report No. W126, Department of Civil Engineering, University of Cape Town, South Africa.
- MOOSBRUGGER RE, WENTZEL MC, EKAMA GA and MARAIS GVR (1993b) A five - pH point titration method for determining the carbonate and SCFA weak acid/bases in anaerobic systems. *Water Sci. Technol.* **28** (2) 237–246. <https://doi.org/10.2166/wst.1993.0112>
- MCCARTY PL (1975) Stoichiometry of biological reactions. *Prog. Water Technol.* **7** (1) 157–172.
- MINO T, SATOH H and MATSUO T (1994) Metabolism of different bacterial populations in enhanced biological phosphate removal processes. *Water Sci. Technol.* **29** (7) 67–70. <https://doi.org/10.2166/wst.1994.0309>
- MUSVOTO EV, WENTZEL MC and EKAMA GA (2000) Integrated chemical- physical processes modelling I. Development of a kinetic based model for weak acid/base systems. *Water Res.* **34** (6) 1857–1867. [https://doi.org/10.1016/S0043-1354\(99\)00334-6](https://doi.org/10.1016/S0043-1354(99)00334-6)
- NEUMANN MB (2012) Comparison of sensitivity analysis methods for pollutant degradation modelling: A case study from drinking water treatment. *Sci. Total Environ.* **433** 530–537. <https://doi.org/10.1016/j.scitotenv.2012.06.026>

- PEREIRA H, LEMOS PC, REIS MAM, CRESPO JPSG, CARRONDO MJT and SANTOS H (1996) Model for carbon metabolism in biological phosphorus removal processes based on in vivo C-NMR labeling experiments. *Water Res.* **30** (9) 2128–2138. [https://doi.org/10.1016/0043-1354\(96\)00035-8](https://doi.org/10.1016/0043-1354(96)00035-8)
- SMOLDERS GJF (1995) A metabolic model of the biological phosphorus removal (stoichiometry, kinetics and dynamic behavior). PhD thesis, Department of Biochemical Engineering, Tech. University Delft, The Netherlands.
- SMOLDERS GJF, VAN DER MEIJ J, VAN LOOSDRECHT MCM and HEIJNEN JJ (1995) A structured metabolic model for the anaerobic and aerobic stoichiometry and kinetics of the biological phosphorus removal process. *Biotechnol. Bioeng.* **47** 277–287. <https://doi.org/10.1002/bit.260470302>
- OLON K, FLORES-ALSINA X, KAZADI MBAMBA C, IKUMI D, VOLCKE EIP, VANEECKHAUTE C, EKAMA GA, VANROLLEGHEM PA, BATSTONE DJ, GERNAEY KV and JEPSSON U (2017) Plant-wide modelling of phosphorus transformations in wastewater treatment systems: Impacts of control and operational strategies. *Water Res.* **113** 97–110. <https://doi.org/10.1016/j.watres.2017.02.007>
- SÖTEMANN SW, RISTOW NE, WENTZEL MC and EKAMA GA (2005) A steady-state model for anaerobic digestion of sewage sludges. *Water SA* **31** (4) 511–527. <https://doi.org/10.4314/wsa.v31i4.5143>
- SÖTEMANN SW, WENTZEL MC and EKAMA GA (2006). Mass balance based plant-wide wastewater treatment plant models – Part 4: Aerobic digestion of primary and waste activated sludges. *Water SA* **32** (3) 297–306. <https://doi.org/10.4314/wsa.v32i3.5274>
- TAIT S, OLON K, VOLCKE EIP and BATSTONE DJ (2012) A unified approach to modelling wastewater chemistry: model corrections. *Proc. 3rd Wastewater Treatment Modelling Seminar (WWTmod2012)*, 26–28 February 2012, Mont-Sainte-Anne, Quebec, Canada.
- VANHOOREN H, MEIRLAEN J, AMERLINCK Y, CLAEYS F, VANGHELUWE H and VANROLLEGHEM PA (2003) WEST: modelling biological wastewater treatment. *J. Hydroinf.* **5** 27–50. <https://doi.org/10.2166/hydro.2003.0003>
- VAN RENSBURG P, MUSVOTO EV, WENTZEL MC and EKAMA GA (2003) Modelling multiple mineral precipitation in anaerobic digester liquor. *Water Res.* **37** (13) 3087–3097. [https://doi.org/10.1016/s0043-1354\(03\)00173-8](https://doi.org/10.1016/s0043-1354(03)00173-8)
- VANROLLEGHEM PA, INSEL G, PETERSEN B, SIN G, DE PAUW D, NOPENS I, WEIJERS S and GERNAEY K (2003) A comprehensive model calibration procedure for activated sludge models. In: *Proceedings WEF 26th Annual Technical Exhibition and Conference (WEFTEC, 2003)*, Los Angeles, CA, USA, 11–15 October 2003. <https://doi.org/10.2175/193864703784639615>
- VAN VEEN HW, ABEE T, KORTSTEE GJJ, PEREIRA H, KONINGS WN and ZEHNDER AJB (1994) Generation of a proton motive force by the excretion of metal phosphate in the polyphosphate-accumulating *Acinetobacter johnsonii* 210A. *J. Biol. Chem.* **269** 29509–29514. <https://doi.org/10.1128/jb.175.1.200-206.1993>
- VOGTS M, IKUMI DS and EKAMA GA (2014) The removal of N and P in aerobic and anoxic-aerobic digestion of waste activated sludge from biological nutrient removal systems. *Proc. WISA Conference 2014*, 25–29 May 2014, Mbombela. <https://doi.org/10.4314/wsa.v41i2.05>
- WANG R, YONGMEI L, CHEN W, ZOU J and CHEN Y (2016) Phosphate release involving PAOs activity during anaerobic fermentation of EBPR sludge and the extension of ADM1. *Chem. Eng. J.* **297** (1) 436–447. <https://doi.org/10.1016/j.cej.2015.10.110>
- WENTZEL MC (1988) Biological excess phosphorus removal in activated sludge systems. PhD thesis, Water Research Group, Department of Civil Engineering, University of Cape Town, South Africa.
- WENTZEL MC, EKAMA GA, DOLD PL and MARAIS GVR (1990) Biological excess phosphorus removal – steady state process design. *Water SA* **16** (1) 29–48.

Modelling mesophilic-thermophilic temperature transitions experienced by an aerobic membrane bioreactor treating furfural plant effluent

LG Kay¹, CJ Brouckaert¹ and RC Sindall¹

¹Pollution Research Group, Department of Chemical Engineering, University of KwaZulu-Natal, King George V Avenue, Durban, 4041, South Africa

ABSTRACT

A mathematical model was developed of an aerobic membrane bioreactor (MBR) treating effluent from a by-products facility at a sugar mill producing furfural, based on measurements of microbial kinetics and stoichiometry at different temperatures. The model was calibrated and validated against plant data using volumetric flow into the MBR and volumetric sludge wasting from the MBR as inputs. The model is able to predict steady-state and unsteady-state operation of the MBR under both mesophilic and thermophilic conditions, and the transitions between the two regimes. Comparison of model simulations and plant data suggests that thermophilic operation is advantageous, but it is less stable than mesophilic operation and frequent feed disruptions can have detrimental effects on MBR operation.

Keywords: wastewater treatment modelling, furfural process effluent, thermophilic, physico-chemical framework

INTRODUCTION

The Sezela Mill Complex, operated by the Illovo Group, consists of a sugar mill with an attached downstream products facility. The downstream site includes a furfural production plant and a range of smaller plants that produce derivatives of furfural. The furfural plant generates an effluent as a by-product of the furfural production process, which is acidic in nature and has a high chemical oxygen demand (COD). The COD consists primarily of acetic acid, with minor amounts of formic acid and intermittent furfural contamination (Judd, 2011). The discharge of the effluent has a negative effect on the mill's water balance. Possible treatment of the effluent therefore provides an opportunity for water recovery, attractive for financial and environmental reasons. This prompted the construction of a pilot-scale aerobic membrane bioreactor (MBR).

The MBR has a hydraulic design capacity of 1 200 m³·d⁻¹ but in practice treats a feed flow rate of no more than 1 000 m³·d⁻¹. The MBR is a 4 600 m³ (29 m diameter, 7 m depth) open cylindrical tank and air is supplied to the mixed liquor through diffusers distributed along the bottom of the tank. It is fitted with a bank of 12 EK400 Kubota flat sheet modules, with a total membrane area of 2 840 m², submerged within the tank. Two 224 kW blowers, rated at 7 060 Nm³·h⁻¹ at 740 mbar, supply air via the fine bubble diffusers along the floor of the tank. A third 61.5 kW blower, rated at 2 880 Nm³·h⁻¹ at 500 mbar, also supplies air as coarse bubbles to scour and clean the membranes. MBR technology was selected due to the suspected presence of an unknown trace toxin in the process effluent that is thought to inhibit conventional aerobic or anaerobic treatment (Judd, 2011). The high (12 to 14 g/L) mixed liquor suspended solids (MLSS) achieved by the MBR is thought to overcome this limitation.

When a steady high feed rate can be sustained, the temperature rises to around 50°C (thermophilic operation), and at a steady low feed rate it operates at around 40°C (mesophilic operation). During mesophilic operation the hydraulic retention time (HRT) is approximately 13 days, and

the sludge retention time (SRT) approximately 105 days. For thermophilic operation, the HRT and SRT are about 6 days and 55 days, respectively.

Feed fluctuations to the MBR occur frequently, due to both external and internal operational factors. During these fluctuations, the temperature of the MBR shifts, which often results in a transition between mesophilic and thermophilic temperature regimes. The transition to lower temperature is marked by a dramatically reduced biomass activity, which leads to operational instability.

Thermophilic operation has the advantages of specific reaction rates several times higher than those for mesophilic operation, and lower sludge production. However, more aeration is required, there is an increased tendency for foaming, and the sludge may have poor settling characteristics (LaPara and Alleman, 1999).

It was proposed that an integrated model capable of predicting temperature, pH and biomass activity (via oxygen utilization rate) would be a useful tool to explore design options and to devise operational strategies that best mitigate feed fluctuations and keep the process as stable as possible.

There has recently been a coordinated effort to establish a comprehensive modelling framework for bio-processes, extending the representation of the biologically mediated reactions to include other physico-chemical phenomena which interact with them (e.g. Batstone et al., 2012; Lizarralde et al., 2015; Solon et al., 2017). Although the energy balance is a logical part of such a framework, it has not received much attention up to now, as few bio-processes involve sufficiently large energy transfers to cause significant interactions with the material transformations taking place. The Illovo MBR model provided an opportunity to demonstrate the incorporation of the energy balance into the framework in a particularly uncomplicated example. It has only one rate-limited biological reaction, four ionic equilibrium systems (acetate, carbonate, ammonia and phosphate), two phase transfer processes (evolution of carbon dioxide and evaporation of water), and one energy balance. Although the energy balance contains a number of terms, those representing transfers to or from the environment turned out to be minor compared to those originating within the process itself.

*Corresponding author, email: brouckae@ukzn.ac.za

Received 1 September 2017; accepted in revised form 3 July 2019

This study describes the development of a mathematical model of the mass and energy balances over the MBR, based on measurements of the microbial kinetics and stoichiometry at different temperatures. The model was used to simulate the dynamic operation of the MBR under mesophilic and thermophilic modes of operation and the transition between the two temperature regimes.

METHODS

The investigation involved a combination of experimental work and model development.

Experimental methods

Influent to the MBR was sampled daily and used to produce weekly composite samples. The composite samples were tested for COD using the standard closed reflux, colorimetric method (5220 D) and for total acidity using a titrimetric method using 0.1 N sodium hydroxide (2310 B) (Bridgewater et al., 2012).

The mixed liquor was sampled near the MBR surface when the blowers were running to ensure adequate homogeneity, and transported immediately to the respirometer to ensure negligible thermal shock. The MBR operating temperature was between 40 and 50°C; the temperature of the sample dropped by no more than 5°C during transportation to the respirometer. The temperature was then increased by the respirometer to the original sampling temperature.

The MLSS concentration of the MBR was obtained daily, following standard methods for total suspended solids (TSS) determination (2540 D) (Bridgewater et al., 2012).

A BM-EVO respirometer (Surcis, 2019) was used to carry out oxygen uptake rate (OUR) tests. This is essentially a 1 L, 2 compartment, mixed reaction vessel, equipped with aeration in one compartment, a circulating pump and a dissolved oxygen (DO) probe in the un-aerated compartment.

The endogenous respiration rate of the MBR mixed liquor was measured by a cyclic OUR test, in which intermittent aeration is used to drive the DO concentration between set limits, and the rate of decline of the oxygen concentration is measured while the aeration is off.

Exogenous respiration rate was measured using a dynamic oxygen uptake response test. After continuous aeration until conditions of endogenous respiration were achieved (about 24 h), the sample was circulated between the aerated and non-aerated compartments, with the DO electrode located in the non-aerated compartment. The drop in DO concentration when substrate is added is a measure of the increased rate of respiration caused by the uptake of the substrate. The relationship between the drop in DO and the reaction rate was established by calibrating the apparatus with a substance with a known chemical oxygen demand (sodium sulphite), and one with a known biological oxygen demand (sodium acetate).

Model development

The model of the Illovo MBR was an assembly of features taken from literature models to match the particular aspects of the system. The biological reactions were formulated following the IWA ASM1 (International Water Association Activated Sludge Model No 1) (Henze et al., 1987), simplified as acetic acid was the only substrate under consideration. ASM1 does not consider energy balances, so the energy balance model of Sedory and Stenstrom (1995) was selected due to its ability to give a complete

breakdown of the heat exchange mechanisms occurring, and its extensive use by various authors (LaPara and Alleman, 1999; Gillot and Vanrolleghem, 2003; Makinia et al., 2005).

The development of the model was carried out over the following series of steps:

- A mass balance over the MBR was formulated with a suitable description of the pertinent kinetic processes, and an ionic speciation subroutine was added for prediction of pH
- An energy balance over the MBR was formulated to predict temperature
- A temperature-dependent description of the kinetic parameters in the mass balance was obtained from laboratory tests at mesophilic (40°C) and thermophilic (50°C) temperatures
- The combined mass and energy balance were calibrated using parameters obtained from the experimental work and the literature
- The dynamic model was validated against an independent set of plant data
- A sensitivity analysis was carried out to determine the effects of various parameters in the energy balance on temperature
- The model was used to simulate both mesophilic and thermophilic operation of the MBR and the results used to assess how process operation could be improved

Model components and reactions

The model was formulated in two stages. Initially it only considered biological reactions, using a formulation similar to ASM1. Later, when it was realised that pH was important, because the plant operators used it to regulate the feed to the reactor, equilibrium ionic reactions were added to the model for pH prediction, which required some additional components, and some additional detail in the representation of the existing components.

The biological reactions included only growth and decay of heterotrophic microorganisms as the feed contained only soluble, readily biodegradable organic substrate (assumed to be acetic acid). Nitrification was not included in the model, as just sufficient nitrogen and phosphorus were dosed in to satisfy the biomass nutrient requirements. Table 1 lists the components involved in the biological reactions; the stoichiometry and kinetics of the biological reactions are represented as a Gujer matrix in Table 2; the ionic model components are listed in Table 3; the transformed Gujer matrix including ionic components appears in Table 4.

Including pH prediction requires a transformation that is characteristic of the physico-chemical framework: adding relevant ionic components and assigning atomic content to the biological reaction components, so that the interaction between the biological and ionic reactions can be represented. Thus S_s was assumed to be acetic acid $C_2H_4O_2$, and X_H and X_p were assigned the same elemental formula $C_5H_7O_2N$.

The stoichiometric coefficients in Table 4 are expressed in molal units rather than COD units as in Table 2, and reflect balances over the reactions on the elements C, H, O, N and electrons. The COD of each component is inherent

Table 1. Biological model components

S_s	Soluble readily biodegradable substrate
X_H	Active heterotrophic biomass (assumed to be $C_5H_7O_2N$)
X_p	Un-biodegradable particulate material resulting from cell death
S_o	Dissolved oxygen

Table 2. Gujer matrix for the biological reactions (in COD units)

<i>i</i>	Components →	1	2	3	4	Rate
<i>j</i>	Processes ↓	S_s	X_H	X_p	S_o	expression
1	Aerobic growth of biomass	$-1/Y$	1		$(1-Y)/Y$	$\frac{\mu_m S_s X_H}{K_s + S_s}$
2	Biomass decay		-1	f_p	$(1-f_p)$	$k_d X_H k_d X_H$
		Biodegradable substrate	Active biomass	Inert matter	Dissolved oxygen	

Y : yield coefficient in biomass growth; μ_m : maximum specific growth rate; K_s : half saturation coefficient; f_p : yield of inert residue; k_d : specific death rate constant

in its elemental formula, and the unit conversions are 64 g COD·mol⁻¹ for C₂H₃O₂⁻ and 160 g COD·mol⁻¹ for C₅H₇O₂N. The parameter Y is not affected by the change in units. The rate expressions continue to use COD units.

Mass balance

The mass of fluid within the MBR (i.e. liquid level) fluctuates depending on the feed rate into the MBR as well as the sludge and permeate withdrawal rates; these are all independent of one another. To simplify the mass balance, the mass content of the MBR was assumed constant, as plant data shows only small fluctuations in liquid level during operation, ±0.1 m in 6.5 m.

The overall mass balance was therefore represented as:

$$m_o = m_e + m_{sw} + m_p \tag{1}$$

Where:

m_o is the mass flow rate of the furfural plant effluent fed into the MBR (kg·s⁻¹)

m_e is the mass flow rate of evaporation from the MBR (kg·s⁻¹)

m_{sw} is the mass flow rate of sludge wasting from the MBR (kg·s⁻¹)

m_p is the mass flow rate of the permeate from the MBR (kg·s⁻¹)

Assuming a uniform density throughout the reactor, and of the feed, evaporated water, sludge wasting and permeate streams, the mass balance is written in volumetric terms as follows:

$$q_o = q_e + q_{sw} + q_p \tag{2}$$

Where:

q_o is the volumetric flow rate of the furfural plant effluent fed into the MBR (m³·s⁻¹)

q_e is the volumetric flow rate of evaporation from the MBR (m³·s⁻¹)

q_{sw} is the volumetric flow rate of sludge wasting from the MBR (m³·s⁻¹)

q_p is the volumetric flow rate of the permeate from the MBR (m³·s⁻¹)

This is illustrated in Fig. 1.

Table 3. Ionic model components

H ⁺	Hydrogen ion
C ₂ H ₃ O ₂ ⁻	Acetate ion (assumed to be ionised S_s)
NH ₄ ⁺	Ammonium ion
CO ₃ ⁼	Carbonate ion
PO ₄ ⁻³	Phosphate ion

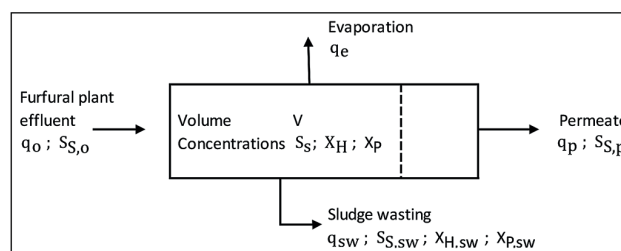


Figure 1. MBR mass balance (modified from Gent, 2012)

Component mass balances

Complete mixing of the MBR contents was assumed for the mass balance model; therefore the soluble concentrations of each of the components in the outlet streams (permeate and sludge wasting) were taken as equal to the concentrations of each component within the MBR. It was assumed that no solids pass through the membranes; consequently, the solids concentration of the waste sludge stream is the same as that in the MBR.

Readily biodegradable substrate, S_s

S_s enters the MBR with the feed, exits through the sludge wasting and permeate streams, and is consumed by reaction. It was assumed that soluble S_s can pass through the membrane. The furfural plant effluent is primarily acetic acid; S_s is assumed to consist entirely of acetic acid. The mass balance for the substrate is:

$$\frac{d(VS_s)}{dt} = q_o S_{S,o} - q_{sw} S_{S,sw} - q_p S_{S,p} + r_{S_s} V \tag{3}$$

Where:

$S_{S,o}$ is the concentration of the substrate in the furfural plant effluent stream (kg·m⁻³)

Table 4. Gujer matrix, transformed according to the physico-chemical modelling framework

<i>i</i>	1	2	3	4	5	6	7	8	Rate
<i>j</i>	S_s	X_H	X_p	S_o	H ⁺	CO ₃ ⁼	NH ₄ ⁺	H ₂ O	expression
1	$-2.5/Y$	1		$\frac{5(1-Y)}{Y}$	$\frac{(7.5-9Y)}{Y}$	$\frac{5(1-Y)}{Y}$	-1	3	$\frac{\mu_m S_s X_H}{K_s + S_s}$
2		-1	f_p	$-5(1-f_p)$	$9(1-f_p)$	$5(1-f_p)$	$(1-f_p)$	$3(1-f_p)$	$K_d X_H$
	C ₂ H ₃ O ₂ ⁻	C ₅ H ₇ O ₂ N	C ₅ H ₇ O ₂ N	O ₂	H ⁺	CO ₃ ⁼	NH ₄ ⁺	H ₂ O	

$S_{s,sw}$ is the concentration of the substrate in the sludge wasting stream (equal to residual substrate concentration S_s) ($\text{kg}\cdot\text{m}^{-3}$)
 $S_{s,p}$ is the concentration of substrate in the permeate (equal to residual substrate concentration S_s) ($\text{kg}\cdot\text{m}^{-3}$)
 r_{s_s} is the rate of consumption of the substrate ($\text{kg}\cdot\text{m}^{-3}\cdot\text{s}^{-1}$)
 V is the volume of the mixed liquor within the MBR (m^3)

Active heterotrophic biomass, X_H

The model assumes no biomass enters the MBR with the feed; it is only generated from growth on S_s . Sludge dosing from the neighbouring conventional activated sludge (CAS) plant was not considered, which occasionally occurs in practice to boost the microbial population. The biomass was modelled with complete retention by the membranes; it was assumed that particulate material can only be removed through sludge wasting. The mass balance for active biomass is:

$$\frac{d(VX_H)}{dt} = -q_{sw}X_{H,sw} + r_{X_H}V \quad (4)$$

Where:

$X_{H,sw}$ is the concentration of the biomass in the sludge wasting stream (equal to residual biomass concentration X_H) ($\text{kg}\cdot\text{m}^{-3}$)
 r_{X_H} is the biomass growth rate ($\text{kg}\cdot\text{m}^{-3}\cdot\text{s}^{-1}$)

Inert organic matter from decay, X_p

X_p is particulate material generated during the decay of biomass, and is only removed through sludge wasting. The mass balance for X_p within the MBR is:

$$\frac{d(VX_p)}{dt} = -q_{sw}X_{p,sw} + r_{X_p}V \quad (5)$$

Where:

$X_{p,sw}$ is the concentration of inert organic matter from decay (equal to residual inert organic matter concentration X_p) ($\text{kg}\cdot\text{m}^{-3}$)
 r_{X_p} is the rate of inert organic matter from decay formation ($\text{kg}\cdot\text{m}^{-3}\cdot\text{s}^{-1}$)

Energy balance

The energy balance model of Sedory and Stenstrom (1995) was selected due to its detailed set of heat exchange mechanisms, and its extensive use by several authors (LaPara and Alleman, 1999; Gillot and Vanrolleghem, 2003; Makinia et al., 2005). The assumption of complete mixing implies a uniform temperature in the MBR, equal to the outlet stream temperature.

The overall energy balance is represented by:

$$V\rho C_p \frac{dT}{dt} = \text{Input}(H_{liq}) - \text{Output}(H_{liq}) + Q \quad (6)$$

Where Q is the sum of the various heat transfer terms illustrated in Fig. 2:

$$Q = Q_{SR} + Q_{AR} + Q_C + Q_{EV} + Q_A + Q_{TW,l} + Q_{RX} + Q_P \quad (7)$$

Where:

Q_{SR} is the heat gain from solar radiation (W)
 Q_{AR} is the heat loss from atmospheric radiation (W)
 Q_C is the heat loss due to surface convection (W)
 Q_{EV} is the heat loss due to surface evaporation (W)
 Q_A is the heat loss due to aeration (W)
 Q_{TW} is the heat loss due to convection from the tank sides and floor (W)
 Q_{RX} is the heat gain from the exothermic reaction (W)
 Q_P is the heat gain from the compressors (W)
 $\text{Input}(H_{liq})$ and $\text{Output}(H_{liq})$ are the enthalpy input and output terms, respectively (W).

Enthalpy flows

The enthalpy terms consider the heat provided or lost by the liquid streams entering and leaving the MBR. They do not include enthalpy lost due to evaporation of water from the MBR, which is accounted for separately through the surface evaporation and aeration heat transfer terms.

The enthalpy terms are defined as follows:

$$\text{Input}(H_{liq}) = q_o \rho_l C_{p,l} T_i \quad (8)$$

$$\text{Output}(H_{liq}) = (q_{sw} + q_p) \rho_l C_{p,l} T \quad (9)$$

Where:

ρ_l is the density of the MBR inlet and outlet streams, assumed equal to water ($\text{kg}\cdot\text{m}^{-3}$)
 T_i is the influent temperature ($^{\circ}\text{C}$)
 T is the MBR temperature ($^{\circ}\text{C}$)
 $C_{p,l}$ is the liquid heat capacity, assumed to be constant at $4170 \text{ J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$

Solar radiation

Radiation from the sun is an important factor due to the open surface of the MBR. A correlation was developed by Raphael (1962) to predict the contribution from solar radiation to the energy balance:

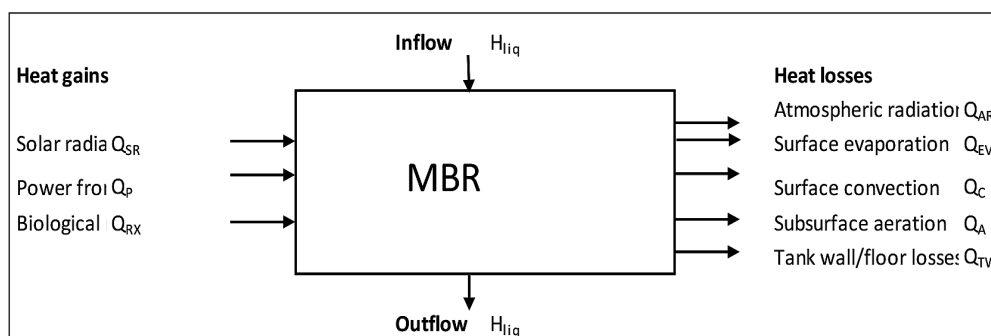


Figure 2. MBR heat exchange components (after Talati and Stenstrom, 1990)

$$Q_{SR} = H_{SR,0} (1 - 0.0071 C_c^2) A_s \quad (10)$$

Where:

C_c is the cloud cover (tenths)

A_s is the surface area of the reactor contents in direct contact with the environment (m^2)

$H_{SR,0}$ is the average daily absorbed solar radiation for clear sky conditions ($W \cdot m^{-2}$)

This is dependent on meteorological conditions, site latitude, and the time of year. The average daily absorbed solar radiation may be calculated from a simplified form presented by Talati and Stenstrom (1990):

$$H_{SR,0} = a - b \cdot \sin\left(\frac{2\pi(d + 183)}{366} + c\right) \quad (11)$$

Where:

d is the day of the year (out of 366)

The values for a , b and c are obtained from the following correlations:

$$\begin{aligned} a &= (4.843 \times 10^{-5})(95.1892 - 0.3591l - 8.4537 \times 10^{-3}l^2) \\ b &= (4.843 \times 10^{-5})(-6.2484 + 1.6645l - 1.1648 \times 10^{-2}l^2) \\ c &= 1.4451 + 1.434 \times 10^{-2}l - 1.745 \times 10^{-4}l^2 \end{aligned}$$

Where:

l is the latitude of the reactor ($^\circ$)

This correlation is valid between 26° and 46° latitude and must be adjusted by adding 183 days for use in the Southern Hemisphere.

Atmospheric radiation

The heat exchange that results from atmospheric radiation is based on Stefan Boltzmann's radiation law. This is expressed as the difference between the incoming and back radiation as follows:

$$Q_{AR} = [\epsilon\sigma(T + 273.15)^4 A_s] - [(1 - \lambda)\beta\sigma(T_a + 273.15)^4 A_s] \quad (12)$$

Where:

ϵ is the water-surface emissivity

σ is the Stefan-Boltzmann constant ($W \cdot m^{-2} \cdot K^{-4}$)

λ is the water-surface reflectivity

β is the atmospheric radiation factor

T_a is the ambient temperature ($^\circ C$)

The atmospheric radiation factor β ranges between 0.75 and 0.85 for most conditions. Previous research has found that 0.97 and 0.03 are good estimates for the emissivity and reflectivity of water (ϵ and λ), respectively (Talati and Stenstrom, 1990).

Surface convection

The temperature difference between the air and the water surface provides the driving force for heat loss by surface convection. The following was obtained from Novotny and Krenkel (1973):

$$Q_C = \rho_g C_{p,g} h_v A_s (T - T_a) \quad (13)$$

Where:

ρ_g is the density of air ($kg \cdot m^{-3}$)

$C_{p,g}$ is the specific heat of air at constant pressure ($J \cdot kg^{-1} \cdot K^{-1}$)

h_v is the convective transfer coefficient ($m \cdot s^{-1}$)

The rate of convective heat loss is affected by the vapour transfer coefficient, which is dependent on the wind velocity. The following equation was developed by Novotny and Krenkel (1973):

$$h_v = (4.537 \times 10^{-3}) A_s^{-0.05} WS \quad (14)$$

Where WS is the wind speed ($m \cdot s^{-1}$)

Surface evaporation

The calculation of the heat loss due to evaporation from the MBR mixed liquor surface developed by Novotny and Krenkel (1973) is dependent on wind velocity, relative humidity, and temperature:

$$Q_{EV} = [55.448 \left(1 - \frac{r_h}{100}\right) + 3.322(T - T_a)] e^{0.06047a} \cdot WS \cdot A_s^{0.95} \quad (15)$$

Where:

r_h is the relative humidity of ambient air (%).

Aeration

Evaporation of water occurs in the course of contact between air from aerators and water in the tank. Air bubbles are assumed to enter the MBR at ambient temperature and humidity, and leave the system at the MBR operating temperature, saturated with water vapour (Novotny and Krenkel, 1973). The amount of water transferred depends on the air flowrate, tank temperature, ambient air temperature, and relative humidity.

The evaporative heat losses are dependent on the difference in the vapour pressure between the water and air. The equation was developed by Novotny and Krenkel (1973) and modified to this final form by Talati (1988):

$$Q_{AL} = \frac{MM_w q_g \Delta H_{vap}}{100R} \left\{ \frac{v_w [r_h + h_f(100 - r_h)]}{(T + 273.15)} - \frac{v_a r_h}{(T_a + 273.15)} \right\} \quad (16)$$

Where:

Q_{AL} is the evaporative heat loss due to aeration (W)

R is the universal gas constant ($8.314 J \cdot mol^{-1} \cdot K^{-1}$)

v_a is the vapour pressure of water at air temperature (Pa)

v_w is the vapour pressure of water at reactor temperature (Pa)

h_f is the exit air humidity factor – assumed to be 1, as air is assumed to be saturated at exit

MM_w is the molar mass of water ($kg \cdot mol^{-1}$)

ΔH_{vap} is the latent heat of evaporation ($J \cdot kg^{-1}$)

Tank wall and floor conduction/convection

Heat losses from the aeration tank walls and floor depend upon the material of construction, the heat transfer area, and its thickness. Heat transfer coefficients for the tank material to air and the tank material to earth are different. Therefore, this model includes two terms: one for the MBR tank wall area exposed to air and one for the MBR tank area exposed to the ground.

The governing equation is as follows:

$$Q_{TW} = U_{a/g} A_t (T - T_{a/g}) \quad (17)$$

Where:

$U_{a/g}$ is the overall heat transfer coefficient for conduction from liquid phase through the reactor walls to air/ground ($W \cdot m^{-2} \cdot K^{-1}$)
 A_i is the area of the reactor that surrounds the liquid phase (m^2)
 $T_{a/g}$ is the temperature of the air/ground ($^{\circ}C$)

The overall heat transfer coefficient is given by:

$$U = \frac{1}{\frac{1}{K_i} + \frac{x_1}{k_1} + \frac{x_2}{k_2} \dots \dots + \frac{1}{K_0}} \quad (18)$$

Where:

x_i is the thickness of materials (m)
 k_i is the thermal conductivity of materials ($W \cdot m^{-1} \cdot K^{-1}$)
 K_i is the surface conductance at the air-surface area inside the tank ($W \cdot m^{-2} \cdot K^{-1}$)
 K_0 is the surface conductance at the air-surface area outside the tank ($W \cdot m^{-2} \cdot K^{-1}$)

The factor $1/K_i$ becomes zero if liquid is in contact with the surface of the wall. If the outside wall is in contact with air, an approximate value of K_0 is taken as $33.90 W \cdot m^{-2} \cdot K^{-1}$. If the wall is surrounded by an earth embankment greater than 3 m thick, K_0 becomes $0.285 W \cdot m^{-2} \cdot K^{-1}$ (Sedory and Stenstrom, 1995).

Heat of reaction

The heat generated from reaction is calculated for the two principle reactions, the growth of biomass and the subsequent oxidation of decaying biomass:

$$Q_{RX} = \left[\frac{\mu_m S_S X_H}{K_S + S_S} \Delta H_{rxn, X_H} + (1 - f_p) k_d X_H \Delta H_{rxn, d} \right] V \quad (19)$$

Where:

$\Delta H_{rxn, X_H}$ is the heat of reaction for the growth reaction ($J \cdot mol X^{-1}$)
 $\Delta H_{rxn, d}$ is heat of reaction for the endogenous respiration reaction ($J \cdot mol X^{-1}$)
 V is the reactor volume (m^3)

The reaction rate expressions in Eq. 19 were previously defined in Table 2, and were assumed to be independent of temperature, and calculated at standard conditions. The heat of reaction for growth (ΔH_{rxn}) is dependent on the biological yield (Y).

Mechanical power

In diffused aeration systems, the air stream is heated by compression. The heat input to the system is dependent upon the efficiency of the compressor. A fraction of the temperature increase during compression is lost as the bubbles expand when they rise through the medium.

$$Q_p = B(1 - \eta/100) \quad (20)$$

Where:

B is the power of the aerator/compressor (W)
 η is the efficiency of the aerator/compressor (%)

Implementation in MATLAB

The mass balance, energy balance, and speciation routine for the prediction of pH were simulated using MATLAB R2010a.

The mass and energy balances form a set of linked ordinary

differential equations (ODE's). Their solution was found by numerical integration using the MATLAB function 'ode23t'.

RESULTS

The results from the laboratory investigation were values and temperature dependence of the reaction kinetic parameters. These were then combined with several sets of historical plant data to calibrate and validate the whole model.

Reaction kinetic parameters

Specific death rate constant, k_d

To determine k_d of the activated sludge, cyclic OUR tests were performed using the respirometer. The mixed liquor sample was aerated for 24 h prior to the test to ensure any external substrate present in the sample was consumed, and endogenous respiration reached. A plot of the natural logarithm of the OUR during endogenous respiration, as a function of time, describes the exponential decay of biomass as a straight line with slope k_d . For an exponential decay, the relationship between OUR and active biomass concentration does not need to be known, since it does not affect the slope of the logarithmic plot. Three samples were tested at $40^{\circ}C$ (mesophilic) and two samples were tested at $50^{\circ}C$ (thermophilic). The results of these tests and the average values for each temperature of study are shown in Table 5.

Table 5 shows a two-fold increase in k_d between mesophilic and thermophilic operation. This is within the range expected based on other related studies (Vogelaar et al., 2003; Abeynayaka, 2009).

Aerobic yield of heterotrophic biomass, Y

Y was estimated using the dynamic response test performed on the respirometer. Six tests were performed at $40^{\circ}C$, and four tests were carried out at $50^{\circ}C$.

The integral of the respirogram gives the oxygen consumed during the test.

Y was determined from the ratio of the oxygen consumed to the COD of the added substrate. The calculated Y values are shown in Table 6 and it can be seen that it is significantly lower at the higher temperature.

Heterotrophic maximum growth rate μ_m and half saturation coefficient K_s

μ_m and K_s were determined by fitting a mass balance model

Table 5. Summary of specific death rate for mesophilic and thermophilic temperatures

Temperature ($^{\circ}C$)	k_d (h^{-1})
40	0.0123 ± 0.0053
50	0.0249 ± 0.0069

Table 6. Summary of yield results for mesophilic and thermophilic temperatures

Temperature ($^{\circ}C$)	Y (units)
40	0.620 ± 0.031
50	0.512 ± 0.013

of the respirometer to the OUR data. The experimentally determined values for k_d and Y were fixed, and μ_m and K_s were found by regression.

The active biomass concentration was estimated as a fraction of the MLSS concentration. A MLSS to MLVSS (mixed liquor volatile suspended solids) ratio of 0.75, and an active biomass concentration (X_H) to MLVSS ratio of 0.35 was used (Casey, 2006; Ubisi et al., 1997).

Four respirograms were used for the mesophilic kinetic regression and three respirograms were used for the thermophilic kinetic regression. The average parameter values for each of the temperatures were found and the standard deviation calculated. These results are summarised in Table 7.

Temperature dependence

The effect of temperature on the reaction rate of biological processes was expressed empirically as $r(T) = r(T_{ref}) \cdot \theta^{(T-T_{ref})}$ (Tchobanoglous et al., 2003). Estimates of the temperature dependence of these parameter values were derived from values at 40 and 50°C. The values of θ for the various kinetic parameters are shown in Table 8. These experimentally obtained parameters were used as initial estimates in the calibration of the combined MBR model.

pH model

In the physico-chemical modelling approach, the concentrations of ionic components are inputs to an ionic speciation subroutine at each iteration step (Brouckaert et al., 2010). The variables of interest that the speciation routine calculates are the pH of the solution, the activity coefficients (γ), the concentrations of the species (c), and the dissolved CO_2 concentration. The driving force for CO_2 transfer depends on the dissolved CO_2 concentration and the partial pressure of CO_2 in air bubbles in contact with the liquid. The rate of CO_2 evolution to the bubbles controls the accumulation of dissolved CO_2 , and therefore strongly influences the system pH.

The molar concentration of hydrogen ions entering the MBR was assumed to be the same as the molar concentration of acetic acid, as determined from the measured COD. Urea and phosphoric acid are dosed daily into the MBR to maintain a healthy microbial community. A nutrient dosing ratio of nitrogen and phosphorus was assumed from literature, as COD:N:P of 100:11:2 on a mass basis (1:0.25:0.021 on a molar basis) (Milenko and Vrtovsek, 2004). The pH of the feed stream was matched to plant data by adjusting the carbonate concentration.

The effect of the pH on the biomass activity in the Illovo MBR had been investigated by Kennedy and Young (2006). Their data were incorporated into the model to simulate the effect of lowered pH on the MBR microbial system. The biomass activities were normalised as percentages, and piecewise linear interpolation between the values was used, as shown in Fig. 3. The biomass activity factor was applied to the growth rate parameter μ_m .

Model calibration

Calibration of the model against data from the full-scale plant was necessary in order to test the applicability of the kinetic parameters obtained from the laboratory tests, and to establish parameters related to the plant operation. Unfortunately, the reactor was not operating normally during the laboratory and modelling investigation. A fault in a distillation unit in the furfural process resulted in a reduced effluent supply, with the

Table 7. Summary of maximum growth rate and half saturation coefficient results for mesophilic and thermophilic temperatures

Temperature (°C)	μ_m (h ⁻¹)	K_s (g·L ⁻¹)
40	0.0209 ± 0.0032	0.895 ± 0.187
50	0.0407 ± 0.0072	1.00 ± 0.47

Table 8. Temperature dependence of mass balance parameters, with $T_{ref} = 40^\circ\text{C}$

μ_m	K_s	Y	k_d
1.07	1.01	0.981	1.07

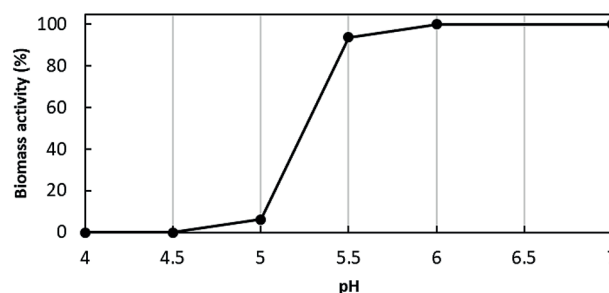


Figure 3. Biomass activity as a function of pH

result that thermophilic operation was not achieved during that year (2013), and historical data from 2012 had to be used for the calibration. This had the disadvantages that the biomass used in the laboratory tests might not have been fully representative of the biomass present at the previous time, and that it was not possible to check conditions that had not been recorded, such as wind speed or cloud cover. Furthermore, certain online measurements, notably the temperature of the feed to the reactor, had been overwritten, due to the limited storage capacity of the SCADA control system. In the calibration regression these unknown variables were treated as constant parameters, constrained to fall within their known ranges.

The calibration was carried out in several stages, using different selected sets of plant data from 2012 and 2013. Firstly, 5 periods of relatively steady operation, in which the feed rate, sludge wasting rate and operating temperature remained fairly constant, were selected for steady state calibrations at temperatures between 40°C and 50°C. Then a period was selected during which varying feed rates resulting in a fluctuating temperature for dynamic calibration.

Recorded furfural plant effluent feed rates (q_0), and sludge wasting rates (q_{sw}) were used as model inputs for the calibrations. All other input parameters were assumed constant (feed temperature, feed COD concentration, weather conditions, etc.; the final values for these parameters are listed in the Appendix). The regressed parameters were adjusted to fit the measured reactor temperature and sludge concentration, the effluent COD, and, in the final calibration stage, the reactor pH.

The calibration procedure was able to match the measured temperatures, sludge concentrations and effluent CODs satisfactorily, with parameter values that fell within their expected ranges. For example, the calibrated fit of the reactor temperature is shown in Fig. 4. The reasonable correspondence between simulation and measurements is an indication that the heat transfer terms that were represented by fixed parameter values, although these must have varied considerably over the time period, had relatively little influence on the energy balance.

However, the simulated pH was much less satisfactory, exhibiting much larger fluctuations than observed on the plant (Fig. 5, before re-calibration). The reason was found to be that the active biomass population in the model responded too slowly to changes in the loading rate of acetic acid, to which the reactor pH is very sensitive; i.e., the maximum growth rate μ_m had been set too low. This had been based on the laboratory data, together with an assumed X_H to MLVSS ratio of 0.35 (see Table 7), which was clearly not appropriate for this particular system. It is a recurring problem of such biological reaction models that the growth rates are based on the concentration of live biomass: a modelling construct that is not directly measurable. When a reaction rate is represented as $\frac{\mu_m S_i X_H}{K_s + S_i}$, μ_m and X_H cannot both be inferred from just a measured reaction rate: all that can be inferred is the product ($\mu_m X_H$).

The dynamic calibration was re-run, including μ_m as a regression parameter. The pH simulation in Fig. 5 shows the substantial improvement between the initial and final calibrations with the root mean squared difference (RMSD) decreasing from 1.14 to 0.47. The decreased fluctuations are a direct result of the increased μ_m value. This higher value of μ_m corresponds to lower, but more rapidly varying, concentrations of active biomass (on average 2.5 g/L compared to 3.7 g/L before recalibration). The re-calibration had only marginal effects on the other variables; for example, the RMSD for the temperature decreased very slightly from the 1.66°C of Fig. 4 to 1.61°C.

Even where it is not strictly correct for the biomass being tested, an assumed value for the active biomass concentration should usually be good enough for most purposes, and indeed proved to be good enough for the steady-state calibration of the model. It was even good enough for most aspects of the dynamic calibration; its limitation was only revealed by the poor prediction of the dynamic pH behaviour. So, the dynamic re-calibration raised μ_m to increase the responsiveness of the reaction rate, while reducing the average X_H to maintain the same average reaction rate.

Validation

The final validation of the model was carried out using plant data from 2010, a different period from that of the calibration data. Only q_0 and q_{sw} were used as variable model inputs, all other parameters were considered to be constant, and fixed at the values determined by the calibration (see Appendix).

Mesophilic-thermophilic temperature transition

The overall objective of the model was the ability to simulate the transition from mesophilic to thermophilic temperature. To test this, a period of data was found from 2010 that showed a clear transition from under 40°C to above 50°C, with no process upsets and minimal external sludge dosing during the transition. Only q_0 and q_{sw} obtained from the plant data were used as model input variables.

The model was able to describe the temperature transition between mesophilic and thermophilic regimes as shown in Fig. 6, with RMSD of 1.7°C, following the trends of the data during the transition; thus meeting the primary objective of the model.

Sensitivity analysis

A sensitivity analysis was performed on the final calibrated model to determine the parameters that had the most significant effect on the MBR temperature.

There was a small effect on the MBR temperature for the observed feed temperature range of 36 to 40°C. The effect was more significant for thermophilic operation than for mesophilic, the lowest feed temperature bringing the MBR temperature down to 45°C. The operation of blowers did not have a significant effect on the MBR temperature, with temperature changes of no more than 1°C.

A wind speed of 2 m·s⁻¹ increases the MBR temperature by around 5°C for mesophilic and thermophilic operation, while a wind speed of 8 m·s⁻¹ reduces the temperature by around 4°C for both thermophilic and mesophilic operation. The MBR operating temperature is sensitive to changes in the ambient temperature, between 10 and 40°C. Mesophilic operation is more sensitive, with temperatures varying by ± 9°C while thermophilic temperatures vary by ± 6°C.

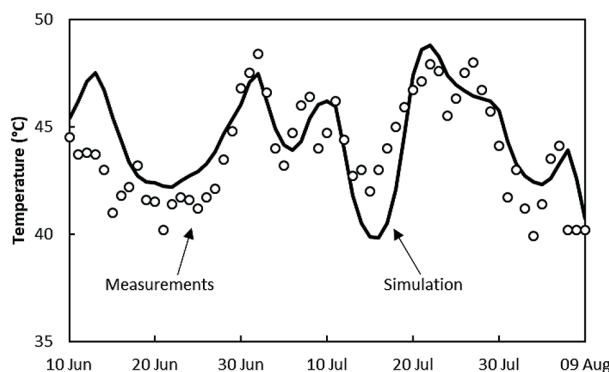


Figure 4. Comparison of measured and simulated temperatures after dynamic calibration

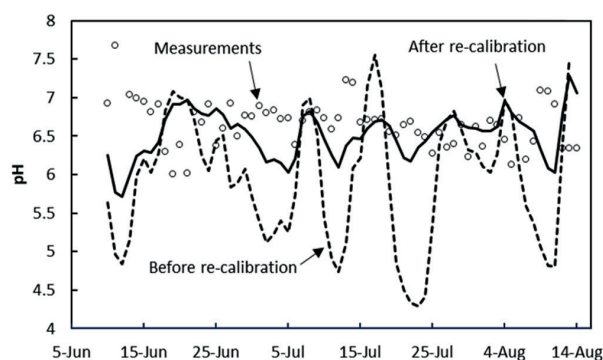


Figure 5. pH comparison for 2012 period before and after dynamic re-calibration

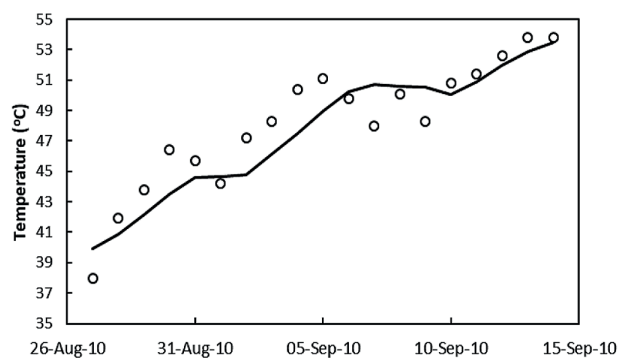


Figure 6. Mesophilic to thermophilic temperature transition during 2010

The assumption of an average wind speed and the ambient temperature throughout the energy balance calibration will therefore cause some error; these parameters vary considerably in reality. Rainfall by itself does not have a significant effect on the steady-state MBR temperature; however, it is the conditions that tend to come with rainfall, i.e., high wind speeds and low ambient temperatures, which affect the MBR temperature.

The aeration rate and the relative humidity have a significant effect on the evaporation rate from the MBR. As expected, the evaporation rate increases with an increase in aeration, and increases with a decrease in relative humidity. The evaporation rate is always greater at a higher temperature. However, a greater amount of water is evaporated per volume of feed for mesophilic temperatures, approximately double that for thermophilic operation. This is due to the longer hydraulic residence time (HRT) experienced for mesophilic operation.

DISCUSSION AND APPLICATION OF THE MODEL

The model has been shown to be capable of steady-state and dynamic prediction of temperature and pH within the required range for the design and operational model predictions.

The temperature can be accurately predicted to within 2°C during a dynamic simulation over an extended period of time (over 60 days), with only the feed rate (q_0) and the sludge wasting rate (q_{sw}) required as inputs. The model is also capable of pH predictions to within 0.5 pH units.

Both temperature regimes are adequately represented by a single biomass population with temperature-dependent kinetic parameters. This does not necessarily mean that microorganism population does not change, merely that it was not necessary for the model to represent a change to match the available experimental data.

The heat generated by reaction is the largest energy term for both modes of operation, constituting 25% and 36% for mesophilic and thermophilic regimes, respectively. There is a significant difference between the heats of reaction for the regimes, as at thermophilic operation higher microorganism death rates occur, leading to higher heat generation from the exothermic biomass decay.

After final calibration of the model the influent flow rate at mesophilic (40°C) and thermophilic (50°C) operation was determined to be 12.9 m³·h⁻¹ and 29.3 m³·h⁻¹ each, respectively. Thus for the same furfural plant effluent feed rate to the MBR, mesophilic operation would require a reactor working volume that is 2.2 times larger than thermophilic operation.

Operational predictions

The model was used to simulate a number of process upset scenarios that the MBR would typically encounter during a season of operation.

Thermophilic-mesophilic/mesophilic-thermophilic process transitions

The feed rate to the MBR was altered following steady-state operation to cause the operating regime to shift from mesophilic to thermophilic, and from thermophilic to mesophilic.

The model was run for 100 simulated hours at the thermophilic flow rate of 29.3 m³·h⁻¹ to ensure steady conditions, then the feed rate was reduced to 12.9 m³·h⁻¹ or vice versa, corresponding to HRTs of 5.8 and 13.1 days, respectively. The sludge wasting flow rate was simultaneously switched

between the corresponding values of 3.1 and 1.6 m³·h⁻¹, with corresponding SRTs of 105.7 and 54.6 days, respectively.

Feed flow decrease

The temperature, pH and biomass responses caused by this transition are shown in Figs 7 and 8. As the feed rate is cut the pH of the MBR initially rises rapidly as a result of the rapid depletion of acetic acid at the lower feed rate. The lower X_H leads to lower X_H growth rates and to an initial decrease in X_H . However, as the temperature in the MBR decreases, the X_H death rate decreases, which leads to an increase in X_H above that observed for thermophilic operation.

A trial and error search was performed on the model to determine the heating required to maintain the temperature at 50°C at the lower feed rate. The power required to maintain thermophilic operation for a mesophilic feed rate was found to be 790 kW.

Feed flow increase

The model predicts that an abrupt increase in feed rate has an adverse effect, as the pH rapidly drops below 5, inhibiting the biological reaction as shown in Fig. 3. Figure 9 shows the modelled pH and temperature responses. The process is unable to recover from this condition without intervention.

To avoid the sharp pH drop, the abrupt increases in feed and sludge wasting rates were replaced by a gradual transition. The simulated transition from mesophilic to thermophilic operation in this way takes about 400 h (17 days) as seen in Fig. 10.

X_H initially increases due to an increased S_S concentration, leading to increased X_H growth rates. An increase in temperature leads to an increased death rate, which decreases X_H , as well as an increase in μ_m which brings S_S back to a steady-state value.

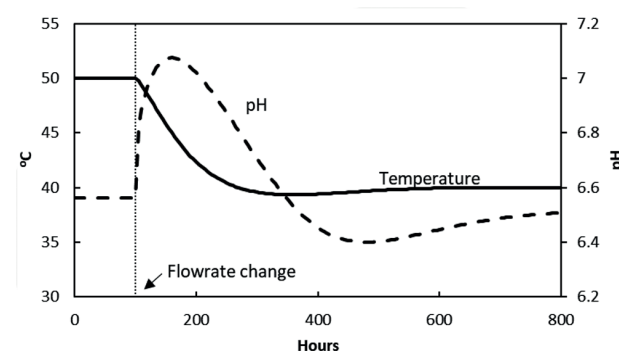


Figure 7. Simulated temperature and pH responses to a decrease in feed flowrate

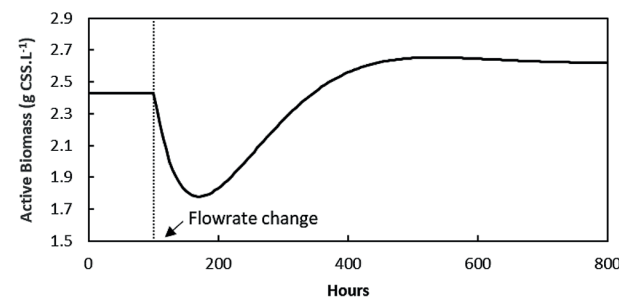


Figure 8. Simulated active biomass response to a decrease in feed flowrate

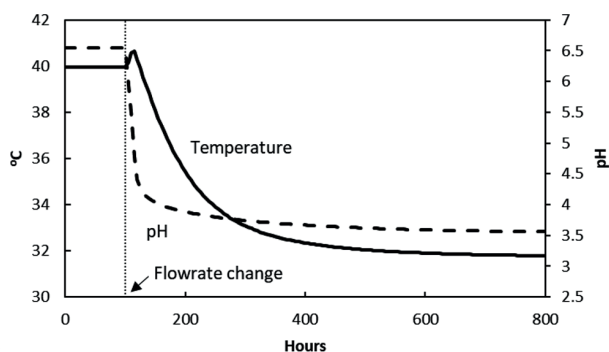


Figure 9. Simulated temperature and pH responses to an abrupt increase in feed flowrate

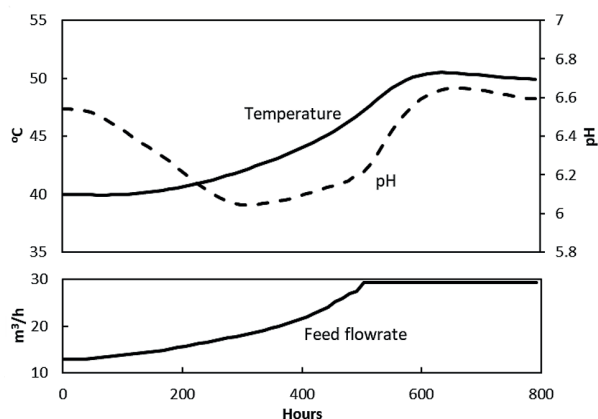


Figure 10. Simulated temperature and pH responses to a controlled increase in flowrate

To determine the amount of cooling required to maintain mesophilic operation for a thermophilic feed rate, cooling was applied. It was found by trial that 920 kW is required to maintain a mesophilic temperature for a thermophilic feed rate.

It is time consuming to increase the operating temperature by controlling the feed only, due to its low pH. Temperature drops can occur considerably faster, as there is no pH limitation. It is critical to keep the MBR pH above 6 to maintain biomass activity. For mesophilic operation the feed rate is tightly constrained; any sudden increase results in a detrimental drop in pH.

MBR instability (feed – no feed – feed)

The transitions described in the previous sections, and illustrated in Figs 7 to 9, provide a pattern for understanding various process disturbances. If feed to the MBR is cut for a significant amount of time, the lack of substrate leads to a depletion of biomass. When the feed is subsequently increased, there is insufficient biomass to assimilate the additional acetic acid load: the pH drops, inhibiting the biological reaction and causing failure of the process if there is no intervention.

When operating in the mesophilic regime (around 40°C) the maximum time for a feed interruption without causing instability was found by trial to be 98 h (4.1 days).

If, after such an incident, the feed flow is increased under pH control (as illustrated in Fig. 10), the temperature gradually climbs back to 40°C, taking over 300 h (12.5 days) to fully recover.

When operating in the thermophilic regime (around 50°C),

the situation is similar, except that the maximum duration of a feed interruption is only 30 h before instability will ensue. During the recovery, the temperature drops to below 42°C in the first 44 h, then starts to increase again. The total recovery time back to a thermophilic temperature of 50°C is again approximately 300 h (12.5 days).

In order to obtain a clear answer as to which temperature regime to design for, a full technical and economic analysis in the context of the entire process design is required, which takes into account the nature and frequency of process upsets. Under steady operating conditions, thermophilic operation can handle more than double the effluent load for a given reactor size. If a method were found to reduce instability under thermophilic operation, this would be the preferred solution. The primary cause of the observed instability is frequent feed fluctuations, and the time taken to recover normal operation.

One of the possible solutions to reduce feed fluctuations to the MBR may be a buffer tank. This would increase the capital costs involved with thermophilic operation, but would likely be less than the capital costs involved with constructing the larger MBR required for mesophilic operation.

CONCLUSIONS

The dynamic mass and energy balance model is successful in predicting steady-state and unsteady-state operation of the MBR under both mesophilic and thermophilic conditions. The temperature can be predicted within 2°C, and the pH within 0.5 units over a period greater than 60 days, using only the daily furfural feeding and sludge wasting rates as model inputs. A single biomass population can be used to model the MBR by including temperature dependencies in the kinetic and stoichiometric parameters.

There are advantages and disadvantages associated with both mesophilic and thermophilic operation of the MBR. A detailed economic analysis would be required to determine which regime is optimal. Both the model simulations and plant data suggest that thermophilic operation could be advantageous, but it is less stable than mesophilic operation, and frequent feed disruptions will have serious adverse effects.

ACKNOWLEDGEMENTS

The authors express their thanks to Illovo Sugar for sponsoring this study, especially Charles Krüger of Illovo Sugar who initiated it, and the technical staff for their practical support.

REFERENCES

- ABEYNAYAKA A (2009) Thermophilic aerobic membrane bioreactor for industrial wastewater treatment. Master of Engineering, Asian Institute of Technology.
- BATSTONE DJ, AMERLINCK Y, EKAMA GA, GOEL R, GRAU P, JOHNSON B, KAYA I, STEYER JP, TAIT S, TAKACS I, VANROLLEGHEM PA, BROUCKAERT CJ and VOLKE E (2012) Towards a generalized physicochemical framework. *Water Sci. Technol.* **66** (6) 1147–1161. <https://doi.org/10.2166/wst.2012.300>
- BRIDGEWATER L, RICE EW, BAIRD RB, EATON AD and CLESCERI LS (eds) (2012) *Standard Methods for the Examination of Water and Wastewater*. APHA-AWWA-WEF, Washington D.C.
- BROUCKAERT CJ, IKUMI D and EKAMA GA (2010) Modelling of anaerobic digestion for incorporation into a plant-wide wastewater treatment model. *WISA Biennial Conference*, 18–22 April 2010, Durban, South Africa.
- CASEY TJ (2006) *Unit Treatment Processes in Water and Wastewater Engineering*. Aquavarra Research Limited, Blackrock, Dublin.

- GENT RD (2012) Experimental effluent treatment at Sezela. *Proceedings of the South African Sugar Technologists' Association* **85** 267–277.
- GILLOT S and VANROLLEGHEM PA (2003) Equilibrium temperature in aerated basins—comparison of two prediction models. *Water Res.* **37** 3742–3748. [https://doi.org/10.1016/S0043-1354\(03\)00263-X](https://doi.org/10.1016/S0043-1354(03)00263-X)
- HENZE M, GRADY CPL (Jnr), GUJER W, MARAIS GVR and MATSUO T (1987) Activated sludge model No 1. *IWA Scientific and Technical Report No 1*. International Water Association, London. ISSN 1010-707X.
- JUDD S (2001) *The MBR Book: Principles and Applications of Membrane Bioreactors for Water and Wastewater Treatment*. Elsevier Ltd, Oxford, UK.
- KENNEDY S and YOUNG T (2006) *Membrane bioreactors for water re-use in Southern Africa*. URL: [http://www.kubotafreeair.com/files/S_%20Kennedy%20&%20T_%20Young%20\(2007\)_%20Membrane%20Bioreactors%20for%20Water%20ReUse%20in%20Southern%20Africa.pdf](http://www.kubotafreeair.com/files/S_%20Kennedy%20&%20T_%20Young%20(2007)_%20Membrane%20Bioreactors%20for%20Water%20ReUse%20in%20Southern%20Africa.pdf) (Accessed 22 March 2013).
- LAPARA TM and ALLEMAN JE (1999) Thermophilic aerobic biological wastewater treatment. *Water Res.* **33** 895–908. [https://doi.org/10.1016/S0043-1354\(98\)00282-6](https://doi.org/10.1016/S0043-1354(98)00282-6)
- LIZARRALDE I, FERNANDEZ-AREVALO T, BROUCKAERT CJ, VANROLLEGHEM PA, IKUMI DS, EKAMA GA, AYESA E and GRAU P (2015) A new general methodology for incorporating physico-chemical transformations into multiphase wastewater treatment process models. *Water Res.* **74** 239–256. <https://doi.org/10.1016/j.watres.2015.01.031>
- MAKINIA J, WELLS SA and JIMA P (2005) Temperature modeling in activated sludge systems: A case study. *Water Environ. Res.* **77** 8. <https://doi.org/10.2175/106143005X67449>
- MILENKO R and VRTOVSEK J (2004) The study of nutrient balance in sequencing batch reactor wastewater treatment. *Acta Chim. Slovenika* **51** 779–785.
- NOVOTNY V and KRENKEL PA (1973) Simplified mathematical model of temperature changes in rivers. *Water Pollut. Control Fed.* **45** 240–248.
- RAPHAEL JM (1962) Prediction of temperature in rivers and reservoirs. *J. Power Div.* **88** 157–188.
- SEDORY PE and STENSTROM MK (1995) Dynamic prediction of wastewater aeration basin temperature. *J. Environ. Eng.* **121** 609–618. [https://doi.org/10.1061/\(ASCE\)0733-9372\(1995\)121:9\(609\)](https://doi.org/10.1061/(ASCE)0733-9372(1995)121:9(609))
- SOLOMON K, FLORES-ALSINA X, KAZADI MBAMBA C, VOLKE EIP, TAIT S, BATSTONE DJ, GERNAEY KV and JEPPESSON U (2015) Effects of ionic strength and ion pairing on (plant-wide) modelling of anaerobic digestion. *Water Res.* **70** 235–245. <https://doi.org/10.1016/j.watres.2014.11.035>
- SURCIS SL (2019) Multifunctional respirometry systems. URL: http://www.surcis.com/en/products_362 [Accessed 8 January 2019]
- TALATI SN (1988) *Heat Loss in Aeration Tanks*. MSc Civil Engineering, University of California.
- TALATI SN and STENSTROM MK (1990) Aeration-basin heat loss. *J. Environ. Eng.* **116** 70–86. [https://doi.org/10.1061/\(ASCE\)0733-9372\(1990\)116:1\(70\)](https://doi.org/10.1061/(ASCE)0733-9372(1990)116:1(70))
- TCHOBANAGLOUS G, BURTON FL and STENSEL HD (2003) *Wastewater Engineering: Treatment and Reuse*. McGraw Hill Education, New York.
- UBISI MF, JOOD TW, WENTZEL MC and EKAMA GA (1997) Activated sludge mixed liquor heterotrophic active biomass. *Water SA* **23** 239–248.
- VOGELAAR JC, KLAPWIJK B, TEMINK H and VAN LIER JB (2003) Kinetic comparisons of mesophilic and thermophilic aerobic biomass. *J. Ind. Microbiol. Biotechnol.* **30** 81–88. <https://doi.org/10.1007/s10295-002-0015-z>

APPENDIX

Table A1. Model parameters

Symbol	Description	Value	Units
A_l	Area of the reactor that surrounds the liquid phase	624.6	m ²
A_s	Surface area of the reactor contents in direct contact with the environment	575.9	m ²
B	Power of the aerator/compressor	288.5	kW
C_c	Fractional cloud cover	0.45	-
f_p	Fraction of inert COD generated by cell lysis	0.15	kg·kg ⁻¹
h_f	Exit air humidity factor	1	-
K_i	Surface conductance at the air-surface inside reactor	33.91	W·m ⁻² ·K ⁻¹
K_o	Surface conductance at the air-surface outside reactor	0.285	W·m ⁻² ·K ⁻¹
K_s	Half saturation constant	See Table 4	kg·m ⁻³
k_d	Specific death rate constant	See Table 2	h ⁻¹
l	Latitude of the reactor	30	°
P	Pressure	101 325	Pa
$S_{s,0}$	Soluble substrate (assumed acetic acid) in feed stream.	17 800	mgCOD·kg ⁻¹
r_h	Relative humidity percentage	83	%
T_a	Ambient temperature	25	°C
T_e	Earth temperature	25	°C
V_r	Volume of the reactor	4 060	m ³
WS	Wind velocity	5	m·s ⁻¹
Y	Biological yield	See Table 3	-
β	Atmospheric radiation factor	0.85	-
ϵ	Water-surface emissivity	0.97	-
η	Efficiency of the aerator/compressor	0.6	-
θ	Reaction rate temperature coefficient	See Table 6	-
λ	Water-surface reflectivity	0.03	-
μ_m	Maximum specific growth rate	0.0763 at 40°C	h ⁻¹

Reconsideration and upgrading of sampling and analysis methods for avoiding measurement-related design and operation failures in wastewater treatment

V Bakos^{1*}, A Deák² and A Jobbágy¹

¹Dept. Applied Biotechnology and Food Science, Budapest University of Technology and Economics, H-1111 Budapest, Szent Gellért tér 4, Hungary

²Dept. Sewage Operations, Budapest Central Wastewater Treatment Plant, Budapest Waterworks Pte. Ltd., H-1211 Budapest, Nagy Duna sor 2, Hungary

ABSTRACT

Success of design and high operational efficiency may basically stand or fall on the quality of measured (or estimated) input data. Even small mistakes committed in the initial steps of sampling and analysis may become large once scaled up in the design process or during full-scale operation. The paper provides several experiment-based practical recommendations and easily implemented, powerful methods for appropriate sampling and analysis practice in wastewater treatment. Representative wastewater characterization is crucial for satisfactory design and cost-effective operation. The paper highlights hidden problems and challenges of sampling and analysis in activated sludge wastewater treatment which may strongly affect the quality of input data, and thus basically determine the modelling outputs. Full-scale results proved that wastewater quality may change significantly in the sampling tubing and vessels; during the sampling process even nitrification can happen. Regarding sludge settling measurements, effects of dilution, temperature, floc structure, nitrate and dissolved oxygen concentrations as well as current biochemical condition of the sludge sample have been studied and important recommendations provided. A combined comparative method including SVI and DSVI measurements has been elaborated for indication and early warning alert of undesired floc structure transformations. Influent BOD₅ concentration is a key factor for describing biodegradability and denitrification capacity of wastewater to be treated. Results of the two most commonly used BOD testing methods were compared for preclarified wastewater. An electrochemical measurement technique provided significantly lower BOD₅ concentrations compared to manometric analysis results with a difference of 23% and 15% on average for unfiltered and filtered samples, respectively. Effects of BOD-based fractionation deviations on predictable denitrification efficiency were studied at different inlet C/N ratios by simulating existing full-scale wastewater treatment plants resulting in remarkable differences in effluent nitrate concentrations. Based on the results, application of the manometric BOD measurement method proved to be preferable.

Keywords: wastewater sampling, measurement errors, data accuracy, biochemical oxygen demand, sludge volume index, activated sludge model

INTRODUCTION

The quality of the data acquired for design, optimization and operation of wastewater treatment depends strongly on the sampling and measurement methods applied. While an extremely wide range of sources of failure may be considered, appropriate analysis and sensor calibration (Rieger et al., 2004; 2006), as well as having a good concept of data acquisition and critical evaluation (Rieger et al., 2010), may help to provide adequate simulation results and accurate estimations of treatment efficiency (Hauduc et al., 2013). Although instrumentation, control and automation applications for wastewater treatment systems are continuously developing, there are still 'blank spots' in this field, principally in sewer control (Olsson et al., 2014). Advanced sampling, measurement and laboratory techniques ensure appropriate technical support to choose technologies with special regard to local conditions, instead of using generalized rules of thumb as predominant design guidelines (Tardy et al., 2012). Despite its importance, little care may sometimes be taken on the origin of initial data, background of sampling and analysis methodology; moreover, this information may not be available at all. This shortage, coupled with sampling and/or measurement failures, can lead to high

uncertainties in wastewater quality, and may be the bottleneck limiting the success of technological concepts proposed.

The current study illustrates effects that may largely influence the accuracy of measured data used for design and operation. Sludge volume index (SVI) is a key design parameter for secondary clarifier dimensioning and for appropriate operation. Numerous important models and methods have been elaborated and widely studied for the determination of activated sludge settleability (Ekama et al., 1984 and 1997; Jenkins et al., 2004; Wanner and Jobbágy, 2014). However, measurements for sludge settling may still involve possible major problems which have not been thoroughly discussed to date, and which may result in inaccurate data and failures in design and operation. Methods that might be used at different wastewater treatment plants (WWTPs) are still diverse and not transparent to enable comparison from site to site.

Another basic parameter, biochemical oxygen demand (BOD), has a key role in bioreactor and aeration dimensioning; moreover, it is an important background parameter influencing the biodegradability-based chemical oxygen demand (COD) fractionation (Ekama et al., 1986; Roeleveld and Loosdrecht, 2002; Drownowski and Makinia, 2014) considerably in ASM (activated sludge model) based simulations. Although several different standard methods are widely used in international laboratory practice, the detailed analytical background of the data is generally unknown during evaluation for simulation studies.

*Corresponding author, email: vbakos@mail.bme.hu

Received 4 April 2018; accepted in revised form 14 June 2019

This paper highlights a number of typical sampling and monitoring failures and uncertainties that can lead to inadequate wastewater characterization, and consequently to poor system design and operation. The purpose of this paper is to show the importance of representative and appropriate sampling and analysis by presenting the high operational impacts of measurement and sampling failures. Moreover, the paper aims to provide useful experiment-based recommendations towards upgrading sampling as well as SVI and BOD measurement methods for avoiding design and operation failures and supporting cost-effective plant operation.

MATERIALS AND METHODS

Analytical methods

COD (chemical oxygen demand), TSS (total suspended solids), VSS (volatile suspended solids), MLSS (mixed liquor suspended solids), ammonium, nitrate, nitrite and DO (dissolved oxygen) concentrations were measured according to standard analytical methods (APHA, 1999). For the measurement of filtered, flocculated COD the flocculation method of Mamais was used (Mamais et al., 1993). DO concentration was measured by portable WTW Multi 3420 FDO Check meter (WTW GmbH, Weilheim, Germany).

SVI and DSVI measurements

For both on-site and laboratory settling measurements activated sludge samples were taken between the 30th and 60th minutes of aeration from intermittently aerated bioreactors of 4 Hungarian activated sludge WWTPs, namely, Budapest Central WWTP (BCWWTP, max. hydraulic capacity: 350 000 m³·d⁻¹, the largest facility in Hungary), Biatorbágy WWTP (BIAWWTP, max. hydraulic capacity: 2 000 m³·d⁻¹), North Budapest WWTP (NPWWTP, ave. hydraulic load: 150 000 m³·d⁻¹) and Kecskemét WWTP (KWWTP, ave. hydraulic load: 22 000 m³·d⁻¹) in the period between January and June, 2015. BCWWTP, BIAWWTP and NPWWTP basically have a pre-denitrifying MLE (Modified Ludzack-Ettinger) bioreactor arrangement, applying intermittent aeration in the aerobic zone for reducing effluent nitrate content. KWWTP has only intermittently aerated carousel-type bioreactors without a preceding non-aerated zone. SVI and DSVI (diluted sludge volume index) were determined according to standard methods (HSI, 2006, No. MSZ EN 14702-1,2:2006; Jenkins et al., 2004) where treated effluent was used for dilution. During the on-site measurements, 10 L of additional mixed liquor sample were taken for further laboratory investigations. In the laboratory the sample was homogenized, and divided into 4 different lab-scale reactors. Two of these were run without

aeration with smooth mixing (i.e. non-pre-aerated samples), and the other two were intensively aerated for 2 h (i.e. pre-aerated samples). The experiments were carried out at 20 and 25°C. Before the settling measurements nitrate concentrations of the mixed liquors were determined. Then appropriate amounts of KNO₃ solution were added to each of the systems in order to set nitrate concentration at a lower (10 mg·L⁻¹ NO₃N) or a higher (25 mg·L⁻¹ NO₃N) value. One non-aerated and one aerated system was applied for each nitrate setting. After 10 min both SVI and DSVI measurements were carried out. Native activated sludge samples were investigated by Olympus CX41 microscope (magn. 200x, phase contrast).

BOD measurements, statistical analysis and mathematical modelling

For the comparative study on BOD₅ measuring protocols, both the electrochemical method (HSI, 2000, No. MSZ EN 1899-1,2:2000; APHA, 1999, No. 5210B., membrane electrode) and manometric technique (HSI, 2004, No. MSZ E 21420-9:2004; APHA, 1999, No. 5210D.) were applied. For manometric measurements WTW OxiTop Control System was used (WTW GmbH, Weilheim, Germany). The measurements were carried out both for unfiltered and filtered raw inlet and preclarified wastewater (5 times, applying 5 parallels) samples taken at BIAWWTP and BCWWTP, respectively. For wastewater characterization ultimate BOD was determined from raw inlet and preclarified wastewater samples and biodegradable COD was calculated according to related guidelines (STOWA, 1996; Roeleveld and Loosdrecht, 2002) and standards (APHA, 1999, No. 5210C.). Raw wastewater samples were taken from the raw inlet of BIAWWTP after the screens, and preclarified wastewater represents the effluent of the Sedipac 3D type primary clarifiers (i.e. combined grit and grease removal and primary settling) of BCWWTP. Statistical analysis of measured BOD results was performed by Dell Statistica (Dell Inc., 2015, version 12, software. dell.com). Since excess biological phosphorus removal was beyond the scope of the study, simulations were carried out using the ASM1 (Activated Sludge Model No. 1, Henze et al., 1987) based WEST software (<http://www.mikepoweredbydhi.com/>). Calculations were carried out for BCWWTP and BIAWWTP, both having pre-denitrification and intermittently aerated carousel-type bioreactors located between the non-aerated zone and the secondary clarifiers, according to the technological layout of the investigated plants presented in Figs 1 and 2.

Table 1 lists a possible characteristic spectrum of raw influent wastewater qualities of BIAWWTP as well as a characteristic preclarified wastewater quality of BCWWTP applied for mathematical simulation studies. The presented influent

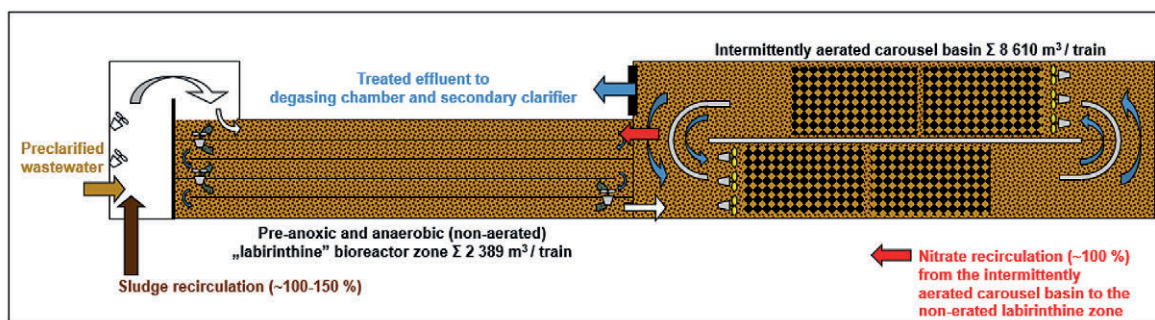


Figure 1. Technological layout of one biological train of BCWWTP; in the period of the investigations 14 out of the 18 biological trains were in operation

wastewater qualities were determined on the basis of on-site measurements. Effluent concentrations were also measured for model fitting and validation. A BOD-based COD fractionation model was developed according to the guidelines of the STOWA simulation protocol (STOWA, 1996; Roeleveld and Van Loosdrecht, 2002). The fractionation model was fitted by supplementary analysis for flocculated and filtered COD concentration in order to confirm its applicability. According to the characteristic bioreactor temperatures measured in June 2015, steady-state simulations were run at 20°C after appropriate fitting and validation of the model. The real influent quality and operational settings of BCWWTP (for

1 June 2015) and BIAWWTP (for 5 different operational days in June, 2015) were applied for simulations and the calculated effluent concentrations of N forms were compared to the measured values of the treated effluent. The measured inlet flow rate was 225 900 m³ d⁻¹ at BCWWTP and varied between 1 722 and 2 681 m³·d⁻¹ at BIAWWTP on the sampling days. In the period of the investigations, intermittently aerated carousel basins operated with cycle settings of alternating 2-h long aerated and 1-h long non-aerated periods, at MLSS concentration of approx. 2.8 kg·m⁻³ at BCWWTP, and settings of 50-min long aerated and 10-min long non-aerated periods at approx. 4.2 kg·m⁻³ MLSS at BIAWWTP.

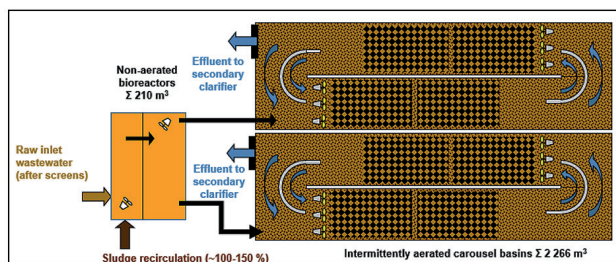


Figure 2. Technological layout of the biological step of BIAWWTP

RESULTS AND DISCUSSION

Failures and latent challenges in sampling and determination of inlet wastewater quality

As illustrated in Fig. 3 autosamplers and sampling vessels may operate as bioreactors due to non-desired attached growth even leading to nitrification while resulting in false analytical data. At BCWWTP the significant change of the inlet wastewater quality was detected after the translocation of the inlet sampling point by approx. 10 m along the inlet wastewater receiving channel. Since only the autosampler inlet tubing was transferred without

Table 1. The influent wastewater quality and effluent dissolved COD concentrations at BCWWTP and BIAWWTP applied for mathematical simulations

Parameter		BCWWTP		BIAWWTP			
		Pre-clarified	raw inlet (after screens)	19 June 2015	24 June 2015	30 June 2015	
Total COD	g·m ⁻³	513	782	737	1057	614	883
Dissolved COD	g·m ⁻³	223	292	234	365	150	305
Total BOD ₅ *	g·m ⁻³	240	434	353	626	287	499
TSS	g·m ⁻³	200	371	547	685	530	564
NH ₄ N	g·m ⁻³	45.4	78.9	95.3	69.6	42.2	66.5
TKN	g·m ⁻³	62.7	87.4	108.5	84.3	54.9	75.4
Effluent dissolved COD	g·m ⁻³	27	25	30	42	31	39
BOD ₅ /NH ₄ N	-	5.3	5.5	3.7	9.0	6.8	7.5

*BOD₅ concentrations are given according to manometric BOD measurement method

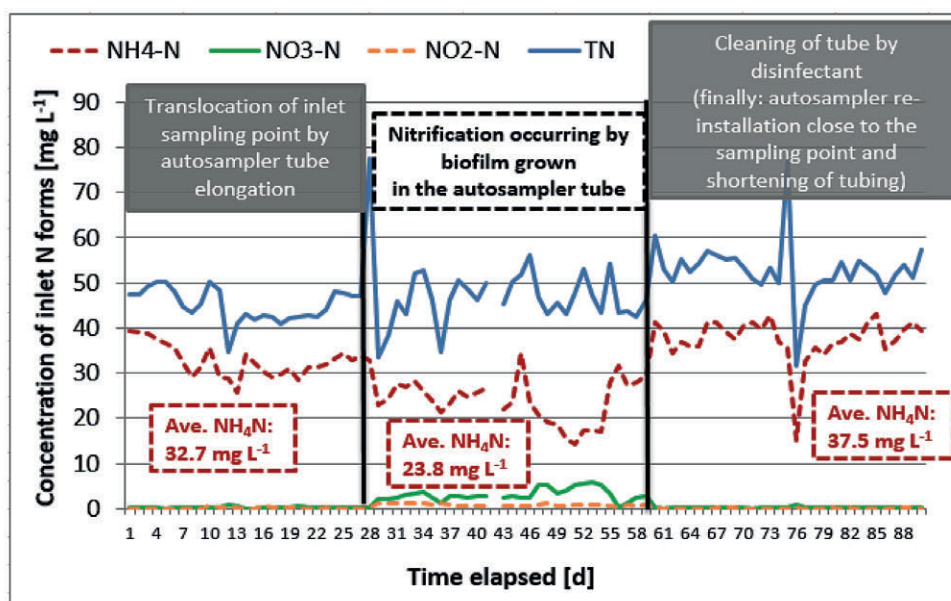


Figure 3. Spontaneous nitrification in autosampler tubing and results of the corrective actions

the translocation of the sampling device, partial sedimentation of the sample and inner biofilm formation occurred in the long and curvy hose. This led to the increase of inlet total COD and TSS concentrations measured as well as to relatively intensive nitrification which was detectable both by the remarkably decreased inlet ammonium concentration and by the increasing inlet nitrate level up to approx. 5 mg·L⁻¹ NO₃N, and additionally, nitrite also appeared in the influent. The problem was quickly detected and other possible causes were carefully investigated and excluded (e.g. by comparison of grab samples quality at both sampling points). The sampling nuisance was eliminated by changing and cleaning the inlet tube using disinfectant and the autosampler was brought directly next to the new sampling point which allowed the shortening of the sampling tube as well.

Recommendations and methods for appropriate autosampler operation and maintenance practice as well as for detection of measurement failures

Cleaning of the inner spare parts and hose of autosamplers is recommended to be a regular preventive maintenance process. Operators are advised to position the autosampler close to the sampling point for minimizing the hose length and to fix the sampling tube properly without saggings and loops where suspended solids can settle down and danger of biofilm formation can be increased. For quick detection of deficient data and measurement failures, regular follow-up of historical data and application of cross-checking calculations (e.g. calculating BOD₅/COD ratio, influent N forms, etc.) are highly recommended. In case of the introduction of new sampling and/or measurement techniques or relevant modifications in the former practice, the careful tracking of the new results, their comparison with earlier data or even more their simultaneous check and comparison with

the results of the formerly applied (reference) method of long standing are strongly recommended preventive actions.

SVI and DSVI measurements and effects of several biochemical influencing factors

On-site measurement of SVI and DSVI at four full-scale WWTPs

In the common practice for testing sludge settleability both SVI and DSVI methods have been widely applied. When the settled sludge volume exceeds 200–250 cm³·L⁻¹, depending on the protocol (e.g. Ekama et al., 1997; ATV, 2000; Jenkins et al., 2004; Wanner and Jobbágy, 2014), dilution may be recommended. However, typically from spring to fall at the relatively low applicable MLSS concentrations this technique may result in excessively diluted MLSS content close to (or lower than) 1 kg·m⁻³ in the test-cylinder, which may just be marginally appropriate for sludge flocculation. Figure 4 illustrates selected characteristic SVI and DSVI measurement results of a 6-month long investigation period (January–June, 2015). Samples were taken several times at the four investigated WWTPs operated at diverse MLSS concentrations and the results of samples with different filamentous abundance are presented.

During investigation of the factors possibly influencing SVI and DSVI it has been found that increasing filament abundance not only increases these values, but the difference between them as well. Moreover, MLSS had no remarkable effect on this in the investigated biomass concentration range of 2.8–6.0 kg·m⁻³ MLSS. As illustrated in Fig. 4, the gap between SVI and DSVI proved to be higher with higher filament abundance. This observation is in accordance with the experience of Lee et al. (1983) regarding the relationship found between TEFL (total extended length of filaments) values and sludge volume indices.

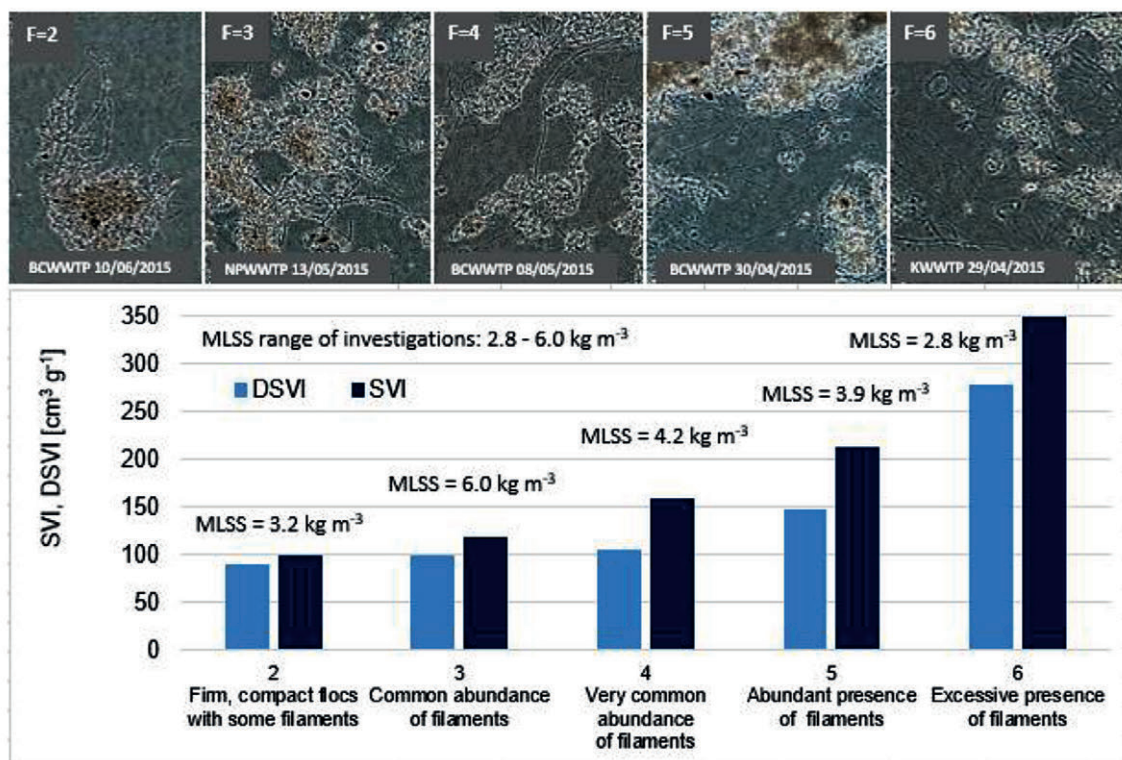


Figure 4. Correlation of filament abundance with the difference between SVI and DSVI values measured. Filament abundance F was determined by the scoring method of Jenkins: 0–none; 1–few; 2–some; 3–common; 4–very common; 5–abundant; 6–excessive (Jenkins et al., 2004).

Combined comparative method for following sludge settleability

At BCWWTP the settleability of activated sludge characteristically deteriorates in the winter period (see Fig. 5a and b), in accordance with the common experience of full-scale plant operators. Both decreasing bioreactor temperature and decreasing F/M ratio (due to higher MLSS concentration maintained in winter for increasing sludge age) may be cited to explain increasing SVI values. The typical MLSS concentration varies between 2.0 and 2.5 kg·m⁻³ for the summer and is in the range of 3.0–4.0 kg·m⁻³ for the winter season. In the most affected periods of floc structure transformations (i.e. at remarkable increase or decrease of filament abundance) both SVI and DSVI measurement methods were applied simultaneously. When the floc structure was close to ideal (i.e. firm flocs with the presence of some filaments), the SVI and DSVI values were practically equal in the investigated MLSS range. However, the higher the filamentous abundance was, the larger the difference that could be detected between the SVI and DSVI data measured.

Full-scale experiences at BCWWTP showed that SVI may be more sensitive for detecting changes in filament abundance, whereas in winter, when relatively high MLSS concentration can be coupled with increased filament abundance, DSVI may better indicate the sludge level variations in secondary clarifiers. In the critical floc structure changing periods (see Fig. 5), the combined comparative measurement of both SVI and DSVI could unambiguously, accurately and even earlier indicate the encouraged (or suppressed) growth of filaments, relative to microscopic observations which have limited quantifying capacity.

Laboratory experiments

For further laboratory experiments activated sludge samples were taken from the same WWTPs that were being investigated. Each of them contains an intermittently aerated bioreactor connected directly to the secondary clarifier. Figure 6 shows that the biochemical switch of the biomass (i.e. switching from anoxic to aerobic conditions by aeration) is crucial regarding the results of settling measurements. The impact of the available electron acceptor (oxygen or nitrate) proved to be more remarkable at higher abundance of filaments even at moderate temperature (20°C). At higher temperature (25°C) and higher nitrate concentration (25 mg·L⁻¹ NO₃N) the risk of sludge flotation increased considerably, particularly for non-pre-aerated samples. In the case of for 30 min pre-aerated samples, the settled sludge had a compact structure in the cylinder with a sharp horizontal upper surface and rising sludge was not detectable, whereas in the non-pre-aerated settled sludge samples a porous, flowing structure could be observed, and the upper surface was rough and irregular. While pre-aerated sludge followed the physical rules of settling, at 25°C settling of non-pre-aerated systems could rather be considered as a kind of complex bio-hydraulic process due to denitrification start-up (i.e. additional suspension flows and circulations). Independently of the biochemical switch of the samples (i.e. incubation carried out with or without aeration), at the start of settling measurement all of the cylinders contained considerable concentrations of dissolved oxygen (DO) due to intensive air input through filling up with sample as well as dilution with DO-rich (5–6 mg·L⁻¹ DO) treated effluent used for measuring DSVI. The DO level was still relatively high at the

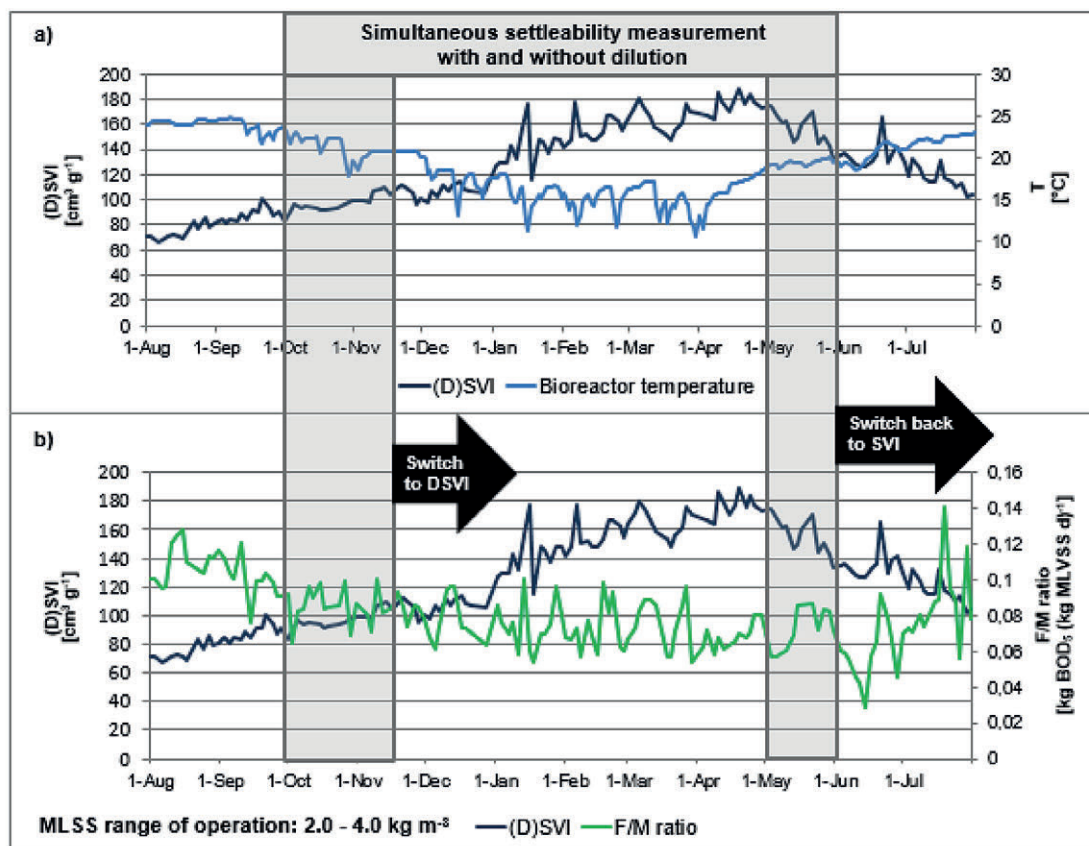


Figure 5. Sludge settleability results at BCWWTP between 01.08.2012 and 31.07.2013 coupled with a) the bioreactor temperature and b) the calculated F/M ratio

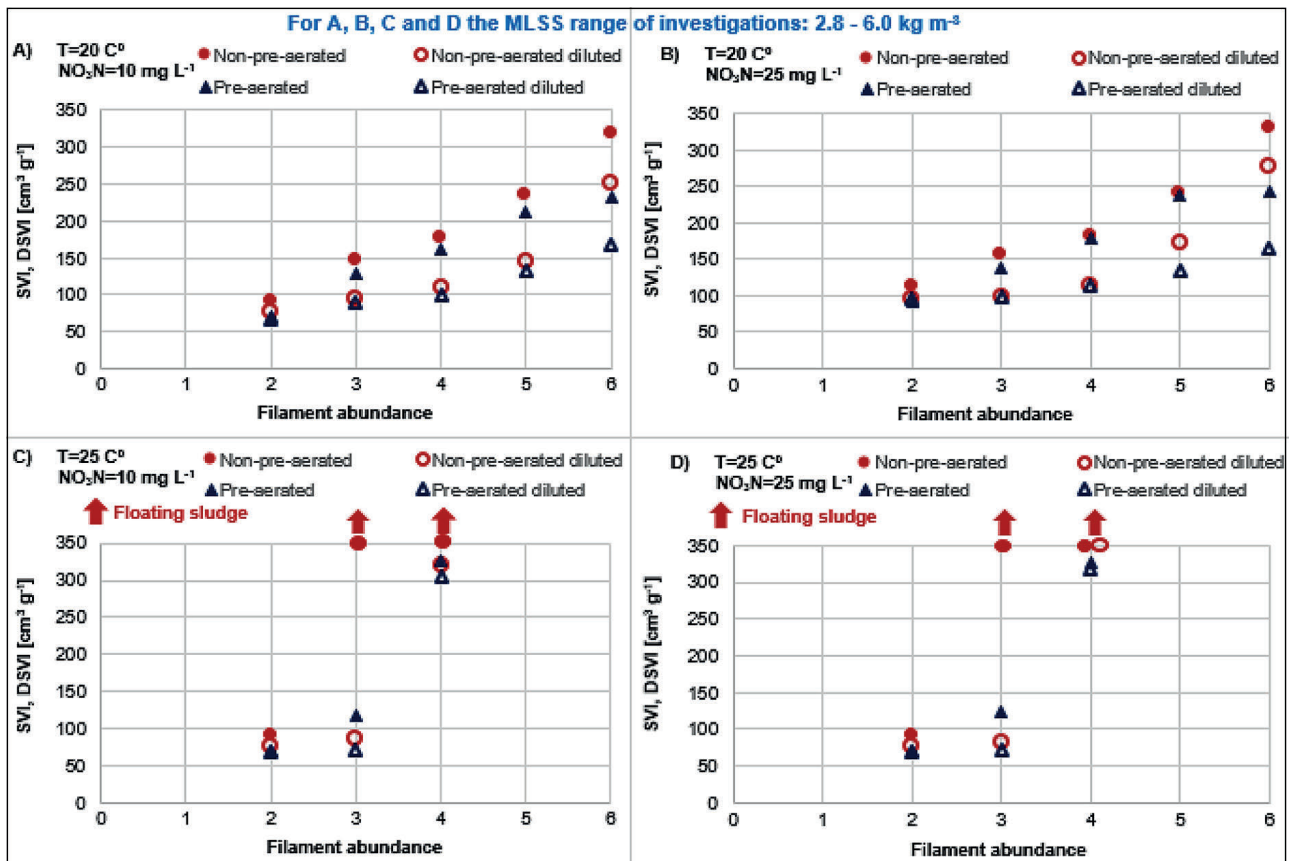


Figure 6. Results of SVI and DSVI measurements after aerobic (pre-aerated) or anoxic (non-pre-aerated) incubation of samples at different temperatures (20°C and 25°C) and nitrate levels (10 and 25 mg·L⁻¹ NO₃N); filament abundance was determined by the scoring method of Jenkins: 0–none; 1–few; 2–some; 3–common; 4–very common; 5–abundant; 6–excessive (Jenkins et al., 2004)

end of the measurement (after 30 min); even for non-pre-aerated sludge it could reach 2–3 mg·L⁻¹ DO at 20°C. It is well known that DO has strong inhibitory effect on denitrification (Jobbágy et al. 2000; Plósz et al., 2003); therefore intensive agitation of mixed liquor prior to the measurement may prevent sludge flotation during the 30 min of the settling test under low S (i.e. low substrate concentration – shortage of readily biodegradable COD) conditions. However, it may not be enough for appropriate anoxic-aerobic switch and to avoid slight swelling of the sludge in the cylinder in case of samples prepared without appropriate aeration. At 25°C, in the case of non-pre-aerated samples, DO concentration could fall to zero during the measurement and this was definitely coupled with detectable nitrate elimination and total flotation of the sludge.

Recommendations and methods for appropriate activated sludge sampling and sludge settleability measurement practice

In the case of intermittently aerated bioreactors it is recommended to take activated sludge samples in the aerated phase of operation (after at least 30 min aeration if possible), in order to avoid settling nuisances caused by denitrification in the measuring cylinder. (Similarly, in full-scale it is recommended to insert a small intensively aerated reaeration tank between the intermittently aerated basin and the secondary clarifier.) For ensuring data accuracy and comparability, it is recommended that samples are always taken in the same way, according to a fixed protocol. At WWTPs operating with several parallel

biological trains and applying intermittent aeration, in particular, the use of advanced automatic sludge volume and MLSS meters with a sampling pattern adjusted to the aeration cycles can be a good and powerful solution (with regular cleaning and maintenance required). At WWTPs where regular careful microscopic observations and floc structure follow-up cannot be carried out it is suggested to apply the simultaneous and comparative measurement of SVI and DSVI, at least in the crucial floc structure changing periods, for early indication of filament abundance changes. SVI indicates at an earlier time the increase of filament abundance; however, DSVI may better represent the conditions (e.g. sludge blanket level) of the full-scale secondary clarifier. It has been proven that filament abundance has a crucial effect on the calculated difference between measured SVI and DSVI values, regardless of the MLSS concentration of the samples in the investigated wide range of 2.8–6.0 kg·m⁻³ MLSS. Although SVI measurement standards definitely prescribe the dilution of the samples in cases when the sludge volume is above 200–250 cm³·L⁻¹ in the graduated 1 L cylinder after 30 min of settling, according to the results presented, it is not recommended to dilute automatically. In case of a relatively low MLSS concentration applied in the bioreactor (≈ 2.0–2.5 kg·m⁻³ MLSS or less), even a 2-times dilution may shift the diluted MLSS concentration down to approx. 1 kg·m⁻³ MLSS in the measuring cylinder, which is about the lowest limit of good flocculation and therefore the settling test may lead to erroneous DSVI data. Designers are advised to be careful with data supplied as design parameters; the background and accuracy should be carefully controlled before calculations.

Effect of BOD measuring method selection on BOD₅ results

BOD₅ concentration values of preclarified wastewater (i.e. wastewater sample taken after primary settling from the effluent of Sedipac 3D type compact primary clarifiers – incl. also grit and grease removal – at BCWWTP) measured by manometric method proved to be significantly higher than those obtained through electrochemical standard method, resulting in a difference of approx. 23% for unfiltered and 15% for filtered samples, on average. The results were statistically evaluated by ANOVA (analysis of variance) method, having chosen the measuring method (i.e. manometric vs. electrochemical) as fixed and sampling day as the random factor. At a significance level of 5% the differences, effects and interactions proved to be significant, as shown in Fig. 7. Although the original electrochemical BOD measurement method does not contain stirring, in the second step the method was applied both with and without stirring to investigate the cause of the difference measured. Results suggested that continuous stirring can be the main factor responsible for the difference detected between the results of the two methods. While the electrochemical method does not apply stirring, bottles equipped for BOD measuring by manometer head should be completely and continuously stirred during the measurement. When both methods were applied with stirring, the difference measured proved to be non-significant at a significance level of 5%. However, standard deviation of data from the electrochemical method was 2–4 times higher compared to that of manometric values. These results also provide an explanation for the higher difference measured for unfiltered samples (23%) compared to filtered ones (15%). Thus, without stirring, suspended solids settled down to the bottom of measuring flasks resulting in remarkably poorer microbial availability of oxygen and substrates.

Simulation results estimating denitrification efficiency using BOD based COD fractionation

For simulation calculations the quality of influent wastewater entering the bioreactors was characterized for BCWWTP and BIAWWTP. The majority of WWTPs in Hungary face carbon deficiency or marginal availability of biodegradable organics (Tardy et al., 2012). At BCWWTP BOD₅/NH₄N ratio of the typical inlet wastewater entering the biological step (i.e. preclarified wastewater) varied in the range of 4–6 which refers to moderate nitrogen removal ability. However, headworks and Sedipac 3D type lamella primary settlers keep the preclarified wastewater quality relatively balanced (compared to changes in the raw influent quality), and due to the mitigating effect of TSS removal by primary treatment, only moderate TSS, COD and BOD₅ fluctuations may happen. At BIAWWTP the inlet wastewater entering the biological step (i.e. raw wastewater leaving the screens) had high ammonium content (typically in the range of 65–95 mg·L⁻¹ NH₄N). Since only screens represented the primary treatment prior to the biological step, huge variations could occur regarding the inlet BOD₅/NH₄N ratio (from 3.5 up to 10), resulting in appropriately poor or moderate, or even good or excellent, carbon availability for denitrification.

Results of simulation calculations presented in Fig. 8 show that the method applied for BOD₅ analysis significantly affects the calculated denitrification efficiency through its strong influence on BOD-based COD fractionation for inlet

wastewater entering the biological step. As highlighted in Fig. 8b, the effects are significant, particularly in the case of marginal availability or severe deficiency of inlet organic carbon source. Modelling results shown in Fig. 8 proved that the best match between calculated and measured effluent concentrations of N-forms could be provided by applying manometric inlet BOD₅ concentrations for biodegradability-based COD fractionation. This fractionation is crucial for generating model input components. By applying BOD concentration measured by the electrochemical method, the slowly biodegradable COD fraction was underestimated, resulting in remarkably lower apparent denitrification efficiency.

In international wastewater laboratory practice the application of manometry is increasingly dominant due to its numerous advantages; however, most laboratories are accredited for and use both methods. Differences presented above may carry a latent risk for designers, particularly in the case of smaller or regional wastewater treatment plants where inlet C/N ratio may typically not be favourable for denitrification due to the large catchment area, long sewer network and enormous load fluctuations.

Recommendations and methods for appropriate BOD measurement practice

The general use of the manometric method (HSI, 2004, No. MSZ E 21420-9:2004; APHA, 1999, No. 5210D) is unambiguously privileged in order to ensure accurate results with low standard

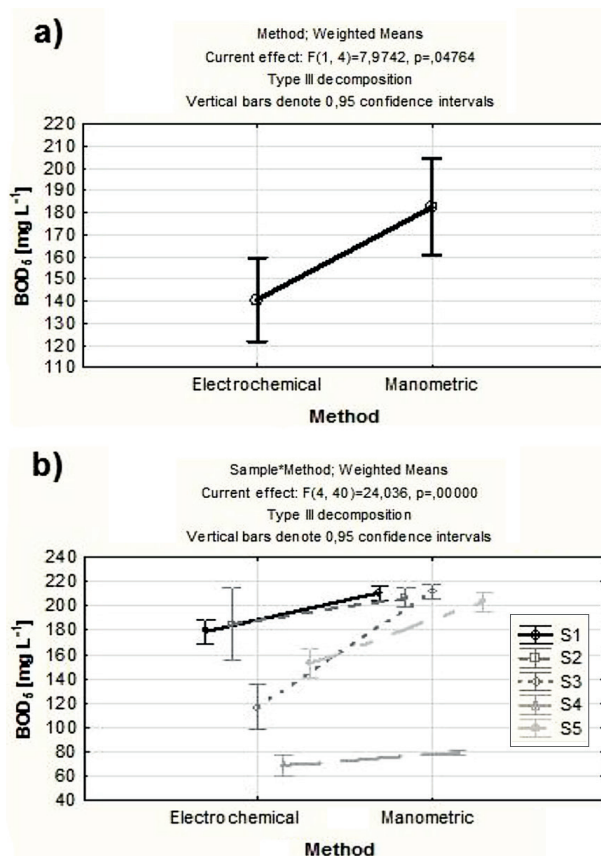


Figure 7. Results of ANOVA for unfiltered samples for a) method effect and b) sample – method interaction for 5 investigated samples (S₁₋₅). ($p_{\text{sample}} = 0.026$; $p_{\text{sample*method}} \approx 0$; $p_{\text{method}} = 0.048$, in case when method effect was analysed as planned comparison: $p_{\text{method}} \approx 0$)

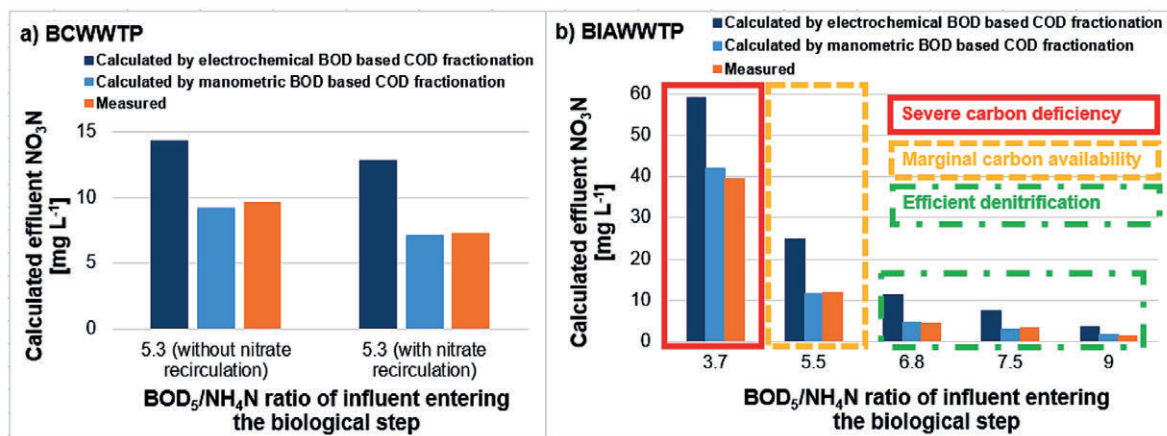


Figure 8. Impact of different BOD₅ measuring methods on calculated denitrification performance as a function of inlet C/N ratios (a) at BCWWTP for typical inlet wastewater quality with and without internal nitrate recirculation; (b) at BIAWWTP for the broad spectrum of possible inlet C/N ratios. (BOD₅/NH₄N ratios indicated on the axis of the graphs refer to values calculated and expressed according to manometric BOD₅)

deviation and provide representative input parameter values for bio-kinetic models. Laboratories are advised to precisely indicate their applied measurement methods in data-supplying documents. If only electrochemically measured BOD concentrations are available, application of additional continuous stirring is recommended and more careful calculations may be needed, since BOD concentration might be underestimated. Supplementary simultaneous measurements of inlet BOD by both methods can be useful for corrective calculations, at least several times if possible.

CONCLUSIONS

This paper points out that carefully designed sampling and analysis as well as clear description of data origin and critical data evaluation are crucial for providing an appropriate basis for calculations leading to best technological solutions. Useful recommendations, as well as powerful and easily implemented methods have been presented which can be important for both designers and practitioners.

It has been illustrated that wastewater quality can significantly change in the sampling tubing and vessels; even nitrification may happen during the sampling processes.

The study highlighted that besides activated sludge floc structure, temperature and dilution, electron acceptor (i.e. DO and/or nitrate) availability as well as biochemical switch of the biomass investigated are also significant influencing factors having strong and synergistic effects on activated sludge settling results. These emphasized findings are recommended to be taken into special consideration in the case of sampling intermittently aerated bioreactors. For early indication of activated sludge floc structure transformations a combined, comparative method of using SVI and DSVI measurements simultaneously has been developed and successfully applied.

Significant differences have been found between BOD₅ values of preclarified wastewater measured by electrochemical and manometric analysis methods, reaching 23%, on average, for unfiltered and 15% for filtered samples, respectively. Simulation studies of Budapest Central and Biatorbágy WWTPs proved that the measured differences in BOD₅ concentrations of influent entering the bioreactors remarkably affected the calculated denitrification efficiency, particularly in the case of low inlet BOD₅/NH₄N ratio.

Recommendations elaborated for upgrading activated sludge settling and wastewater BOD measurements ensure higher transparency and comparability, as well as appropriate integration of wastewater and mixed liquor analysis methods into state-of-the-art design and operation practice in order to avoid modelling failures, and prevent expensive operation and costly corrections of misdesigned technologies.

ACKNOWLEDGEMENTS

The cooperation of Budapest Waterworks Pte. Ltd. facilitating on-site measurements, full-scale process modelling and supplying data is gratefully acknowledged, with special thanks to Csaba Haranghy, Bence Márialigeti, László Váci and Szabolcs Pintér. The valuable help of engineer colleagues at North Budapest WWTP (Budapest Sewage Works Pte. Ltd.) and János Csóvári at Kecskemét WWTP (Bácsvíz Pte. Ltd.) is gratefully acknowledged. The helpful contribution of Dr Zsuzsanna Nagy (DHI Hungary Ltd.) enabled the access to WEST software for simulation calculations. Useful consultations with and technical support of Enrico Remigi (DHI Belgium) in modelling are also greatly appreciated. Authors are thankful to Prof. Sándor Kemény for his kind help in statistical evaluations. Authors express their special thanks to the laboratory and operational staff of the investigated wastewater treatment plants.

REFERENCES

- APHA (AMERICAN PUBLIC HEALTH ASSOCIATION) (1999) *Standard Methods for the Examination of Water and Wastewater* (20th edn). Methods: No. 5210B., 5210C., 5210D. APHA, Washington DC.
- ATV Design Standard A131 (2000) Dimensioning of single-staged activated sludge plants. ATV-DWA, Germany, Hennef.
- DREWNOWSKI J and MAKINIA J (2014) The role of biodegradable particulate and colloidal organic compounds in biological nutrient removal activated sludge systems. *Int. J. Environ. Sci. Technol.* **11** (7) 1973–1988. <https://doi.org/10.1007/s13762-013-0402-1>
- EKAMA GA, PITMAN AR, SMOLEN M and MARAIS GVR (1984) Secondary settling tanks. Chapter 8. In: *Theory, Design and Operation of Nutrient Removal Activated Sludge Processes*. Water Research Commission, Pretoria. 8.1–8.14.
- EKAMA GA, DOLD PL and MARAIS GVR (1986) Procedures for determining influent COD fractions and the maximum specific

- growth rate of heterotrophs in activated sludge systems. *Water Sci. Technol.* **18** 91–114.
- EKAMA GA, BARNARD JL, GÜNTHERT FW, KREBS P, MCCORQUODALE JA, PARKER DS and WAHLBERG EJ (1997) Secondary settling tanks: theory, modelling, design and operation. IAWQ Scientific and Technical Report No. 6. IAWQ, London. 41–65.
- HAUDUC H, RIEGER L, OEHMEN A, VAN LOOSDRECHT MCM, COMEAU Y, HÉDUIT A, VANROLLHEGEM PA and GILLOT S (2013) Critical review of activated sludge modeling: state of process knowledge, modeling concepts, and limitations. *Biotechnol. Bioeng.* **110** (1) 24–46. <https://doi.org/10.1002/bit.24624>
- HENZE M, GRADY JR CPL, GUJER W, MARAIS GVR and MATSUO T (1987) A general model for single-sludge wastewater treatment systems. *Water Res.* **21** (5) 505–515. [https://doi.org/10.1016/0043-1354\(87\)90058-3](https://doi.org/10.1016/0043-1354(87)90058-3)
- HSI (Hungarian Standards Institution) (2000) Standard methods for biochemical oxygen demand determination from water and wastewater – electrochemical method, No. MSZ EN 1899-1:2000; MSZ EN 1899-2:2000. Hungarian Standards Institution, Budapest.
- HSI (Hungarian Standards Institution) (2004): Standard methods for biochemical oxygen demand determination from water and wastewater – manometric method, No. MSZ E 21420-9:2004. Hungarian Standards Institution, Budapest.
- HSI (Hungarian Standards Institution) (2006): Standard methods for sludge volume index measurement, No. MSZ EN 14702-1,2:2006. Hungarian Standards Institution, Budapest.
- JENKINS D, RICHARD MG and DAIGGER GT (2004) *Manual on the Causes and Control of Activated Sludge Bulking, Foaming and other Solids Separation Problems*. (3rd edn). CRC Press LCC, Boca Raton, Florida. 10, 23–24, 52–53.
- JOBBÁGY A, SIMON J and PLÓSZ BGY (2000) The impact of oxygen penetration on the estimation of denitrification rates in anoxic processes. *Water Res.* **34** (9) 2606–2609. [https://doi.org/10.1016/S0043-1354\(00\)00013-0](https://doi.org/10.1016/S0043-1354(00)00013-0)
- LEE S-E, KOOPMAN B, BODE H and JENKINS D (1983) Evaluation of alternative sludge settleability indices. *Water Res.* **17** (10) 1421–1426. [https://doi.org/10.1016/0043-1354\(83\)90273-7](https://doi.org/10.1016/0043-1354(83)90273-7)
- OLSSON G, CARLSSON B, COMAS J, COPP J, GERNAEY KV, INGILDSSEN P, JEPSSON U, KIM C, RIEGER L, RODRÍGUEZ-RODA I and co-authors (2014) Instrumentation, control and automation in wastewater – from London 1973 to Narbonne 2013. *Water Sci. Technol.* **69** (7) 1373–1385. <https://doi.org/10.2166/wst.2014.057>
- PLÓSZ BGY, JOBBÁGY A and GRADY JR CPL (2003) Factors influencing deterioration of denitrification by oxygen entering an anoxic reactor through the surface. *Water Res.* **37** 853–863. [https://doi.org/10.1016/S0043-1354\(02\)00445-1](https://doi.org/10.1016/S0043-1354(02)00445-1)
- MAMAIS D, JENKINS D and PITT P (1993) RAPID COMMUNICATION: A rapid physical-chemical method for the determination of readily biodegradable soluble COD in municipal wastewater. *Water Res.* **27** (1) 195–197. [https://doi.org/10.1016/0043-1354\(93\)90211-Y](https://doi.org/10.1016/0043-1354(93)90211-Y)
- RIEGER L, LANGERGRABER G, THOMANN H, FLEISCHMANN NG and SIEGRIST HR (2004) Spectral in-situ analysis of NO₂, NO₃, COD, DOC and TSS in the effluent of a WWTP. *Water Sci. Technol.* **50** (11) 143–152.
- RIEGER L, LANGERGRABER G and SIEGRIST H. (2006) Uncertainties of spectral in situ measurements in wastewater using different calibration approaches. *Water Sci. Technol.* **53** (12) 187–197. <https://doi.org/10.2166/wst.2006.421>
- RIEGER L, TAKÁCS I, VILLEZ K, SIEGRIST H, LESSARD P, VANROLLEGHEM PA and COMEAU Y (2010) Data reconciliation for wastewater treatment plant simulation studies – planning for high-quality data and typical sources of errors. *Water Environ. Res.* **85** 426–433. <https://doi.org/10.2175/106143009X12529484815511>
- ROELEVELD PJ and VAN LOOSDRECHT MCM (2002) Experience with guidelines for wastewater characterization in The Netherlands. *Water Sci. Technol.* **45** (6) 77–87.
- STOWA (1996) *Methoden voor Influentkarakterisering – Inventarisatie en Richtlijnen*. STOWA Report 80-96. STOWA, Utrecht.
- TARDY GM, BAKOS V and JOBBÁGY A (2012) Conditions and technologies of biological wastewater treatment in Hungary. *Water Sci. Technol.* **65** (9) 1676–1683. <https://doi.org/10.2166/wst.2012.062>
- WANNER J and JOBBÁGY A (2014) Activated sludge solids separations. Chapter 10. In: Jenkins D and Wanner J (eds.) *Activated sludge – 100 Years and Counting*. IWA Publishing, Glasgow. ISBN 9781780404936. 171–194.

Physico-chemical properties and bacterial community structure dynamics during the mesophilic anaerobic digestion of pit latrine faecal sludge

MC Changara^{1*}, WT Sanyika², C Bangira³ and S Misi⁴

¹Department of Environmental Sciences and Technology, Chinhoyi University of Technology, P.O. Box 7724, Chinhoyi, Zimbabwe

²Department of Biotechnology, Chinhoyi University of Technology, P.O. Box 7724, Chinhoyi, Zimbabwe

³Department of Agricultural Engineering, Chinhoyi University of Technology, P.O. Box 7724, Chinhoyi, Zimbabwe

⁴Department of Civil Engineering, University of Zimbabwe, P.O. Box MP167, Mount Pleasant, Harare, Zimbabwe

ABSTRACT

The study characterized the changes in physico-chemical properties and bacterial community structure during mesophilic anaerobic digestion (AD) of pit latrine sludge. The sludge was sampled from six different pits six times at an interval of 40 days. Standard techniques were used to assess the changes in pollution indicators including COD and faecal coliforms. Metagenomic DNA from a composite sample from the six pits' sludge was then extracted at Days 0, 14 and 35 and directly sequenced followed by analysis of the microbial structure using the Ribosomal Database Project tools. Multivariate analyses were used to identify the main determinants of microbial community structure during the digestion process. AD significantly reduced the levels of pollution indicators ($p < 0.05$). Total solids, volatile solids and COD were reduced by 17–27%, 52–79%, and 42–63%, respectively. The indicator pathogenic microorganisms FC and *E. coli* were reduced by 34–54% and 35–60%, respectively. The reduction in terms of COD and BOD were, however, not sufficient to satisfy the standards for safe disposal into the environment. Proteobacteria were the most dominant bacterial phylum in the undigested sludge (24.1%) and were significantly reduced to 2.5% at the peak of the AD (Day 14) up until Day 35. Firmicutes significantly increased ($p < 0.05$) from 22.4% to 28.8% at Day 14 before being reduced to 11.6% at Day 35. This study contributes to our understanding of AD of pit latrine faecal sludge through mesophilic AD as a baseline study, and helps to inform future research on mesophilic AD.

Keywords: anaerobic digestion, bacterial community structure, next generation sequencing, pit latrine sludge, pollution indicators

INTRODUCTION

Globally one of the targets of Sustainable Development Goal (SDG) No. 6 is for all nations to provide access to adequate and equitable sanitation and hygiene for all, and to end open defecation by 2030 (WHO/UNICEF, 2017). The drive of improving access to clean drinking water and safe sanitation, according to Montgomery and Elimelech (2007), is one of the least expensive and most effective ways to improve public health and save lives. Hutton et al. (2007) states that improved sanitation has significant economic benefits, for example, the return on 1 USD spent on water and sanitation improvements in low-income countries is 5–46 USD depending on the intervention. The traditional approach when urban and peri-urban houses are being built in developing countries is that every community shall be able to have access to sewer-based systems and the 'flush and forget' stance shall remain forever. Sewer-based systems, however, are usually too costly, too complex and sometimes use too much energy to be implemented in poor and less-developed countries (Lalander et al., 2013; Mara, 2013). Pit latrines have been constructed in peri-urban settlements of developing countries like Zimbabwe as a faecal sludge management technology to substitute sewer-based systems. Pit latrines are among the affordable on-site sanitation facilities (Torendel et al., 2016) in developing countries, but will eventually fill-up at some stage and must be replaced or emptied. Replacement or emptying causes inconvenience is expensive and a health risk (Torendel et al., 2016).

Colon et al. (2015) highlights that one of the greatest challenges with sanitation is developing and managing innovative, user-friendly and easy to adapt low-cost pit faecal sludge disposal systems that are relevant to the local context. In addressing the management of faecal sludge from full latrine pits, there is a growing interest in technologies that not only treat human waste but also enable the recovery of nutrients for agricultural productivity and energy (Gijzen, 2002; Katukiza et al., 2012). However, the pathogenic microorganisms commonly found in human faeces are a cause for concern regarding human health if human excreta are used for crop production. These pathogenic microorganisms include *Campylobacter jejuni/coli*, *E. coli*, *Salmonella typhi/paratyphi*, *Salmonella* spp., *Shigella* spp., and *Vibrio cholera* (Drangert, 1998; Schonning and Stenstrom, 2004). Parasitic microorganisms and worms commonly found in human faecal waste include *Cryptosporidium parvum* and helminths; *Ascaris lumbricoides* (roundworm), *Taenia solium/saginata* (tapeworm), *Trichuris trichiura* (whipworm), *Ancylostoma duodenale/Necator americanus* (hookworm) and *Schistosoma* spp. (blood flukes) (Schonning and Stenstrom, 2004; WHO, 2006).

The process of AD has been found to reduce the levels of pathogens and pollution indicators while retaining plant nutrients (Sahlstrom et al., 2004). AD is a well-established process in which bacteria convert organic wastes to methane and carbon dioxide, a gas mixture (~60% methane and 40% carbon dioxide) called biogas (Lukehurst, 2010). According to Gijzen (2002) and McCarty et al. (2011), the process of AD has a lot of advantages compared to other treatment options, mainly due to low operational costs, production of energy in the form of biogas and potential reuse of the nutrients present

*Corresponding author, email: chrischangara@gmail.com

Received 6 June 2018; accepted in revised form 25 June 2019

in the digested waste. In developing countries such as Ghana the interest in AD is shifting from energy to onsite sanitation systems associated with faecal sludge management, especially in relation to some challenges in emptying, transportation and disposal of pit latrine sludge (Boot and Scott, 2008) and the technology can as well be explored in Zimbabwean context.

Little reliable data is available on the effectiveness of AD of human excreta, both at laboratory and pilot scale (Colon et al., 2015), for agricultural re-use and safe disposal of the digested sludge. Of the little available information on AD of human excreta, much of the emphasis is on energy and biogas production (Park et al., 2001; Colon et al., 2015; Onabanjo et al., 2016). Forbis-Stokes et al. (2016) did a pilot study at two different locations on onsite faecal sludge treatment through AD and recorded 85 and 89% reduction in terms of COD in the effluent in relation to the estimated faecal sludge input. Also log reductions of faecal coliforms recorded in the sterilization tank due to AD were greater than 5. Nwaneri et al. (2008) found that COD values for faecal material present in the first layer of pit latrine sludge were significantly lower than those measured for fresh faeces. The reduction in pollution indicators should, however, meet regulatory standards, for example, the WHO guidelines of 2006 (WHO, 2006), to ensure environmental safety and agricultural utilization of the treated sludge.

In the process of AD, there is also a need to increase the existing knowledge on the dynamics of the complex interacting microbial community that drives the process (Venkiteshwaran et al., 2016). Diverse microbial communities degrade organic matter within pit latrines, and little is known about the specific communities present in pit latrines and their association with faecal decomposition within the pit environment (Torendel et al., 2016). One key area requiring new knowledge involves understanding the effect of AD on the changes in bacterial community structure associated with the pit latrine faecal sludge. The understanding of the microbial community structure and ecology will help in the development of strategies to maintain and improve treatment efficiency (Narihiro et al., 2015).

The aim of this research was to assess the changes in physico-chemical properties and bacterial community structure during the mesophilic AD of pit latrine faecal sludge, and, secondly, to assess the effectiveness of the digestion to treat the faecal sludge to the quality requirements expected for agricultural re-use and safe environmental disposal, for example, the WHO (2006) guidelines for use in crop production and safe disposal. The bacterial community structure and dynamics were investigated through the analysis of the bacterial 16S ribosomal RNA (rRNA) gene diversity generated through Next Generation Sequencing (NGS).

MATERIALS AND METHODS

Sampling of pit latrine sludge

Samples were collected from 6 pit latrines in Shackleton (30.03°E and 17.30°S), located in Chinhoyi, Zimbabwe, during the period February–October 2016, 6 times from each pit at 40-day intervals. Pit latrines were selected purposively, being of approximately 2 years in use, constructed to a depth of approximately 2 m, accessed by more than one family of at least 8 adults (above 16 years) and without connection to water sources. Sampling was conducted using a modified auger which was graduated to show sampling depth. After determining the depth of each pit latrine, triplicate faecal samples were taken

from each pit (about 50–60 cm pit slab to contents depth). The sludge was sampled from the top surface about (10 cm deep) of the pit latrine sludge and the modified auger that was used for sampling collected about 135 g of sludge at each time. Composite samples of approximately 275 g were placed in 300 mL clear polythene bottles for AD, physico-chemical characterisation and microbial composition analysis. A total of 2 kg moist sludge was collected per each pit at each sampling time. The samples were immediately placed on ice and later transported to the laboratory for physico-chemical, microbial analysis and AD. Physico-chemical and microbial analysis excluding metagenomic DNA analyses were conducted after addition of 10% (w/v) rumen fluid inoculums before and after the AD for the 6 sampling times per pit. A composite sample was then collected from all 6 pits for which metagenomic DNA and physico-chemical properties were determined at Day 0, both without addition of rumen and after addition of rumen, at the climax of biogas production (Day 14) and at the end of the AD (Day 35). The sampling points for the metagenomic analysis were determined from preliminary experiments.

Inoculum preparation

A 150 mL volume of rumen fluid with a total solid content of 7% was obtained from a nearby abattoir (The Cold Storage Commission, Chinhoyi) and used as inoculum as prescribed by Ojolo et al. (2007). The rumen fluid was first mixed with distilled water in a ratio 1:1 (v/v). A 5-mm sieve was used to separate solid content from liquid in the preparation of the inoculum.

Sample preparation and anaerobic digestion

A 1.5 kg moist faecal pit latrine sludge sample from each of the 6 pits (total solids content ranging from 3 to 7%, depending on different pits and different sampling times) was fed into the 2-L high-density polyethylene (HDPE) plastic container reactors (Fig. 1). Each reactor was fitted with a flexible 12 mm gas delivery pipe (Fig. 1) and jacketed with cotton wool to insulate the internal working temperature from ambient fluctuations. An airtight seal was then used for anaerobic conditions to prevail. The reactors were shaken by hand at least 2 times a day for 30 s. The biogas produced in the reactors was collected in an inverted measuring cylinder in a plastic beaker (Fig. 1) containing 5% sodium hydroxide solution. The amount of biogas produced was measured through the displacement of the sodium hydroxide in the inverted measuring cylinder. At the end of the measuring cylinder there was a 12 mm gas delivery pipe linked to a gas valve then a Bunsen burner (Fig. 1) at the end of the anaerobic digester set up. At the end of each day the cumulative biogas in the inverted measuring cylinder in each reactor was released via the gas valve and a burning test was conducted, checking if the gas produced was burning and producing a blue flame on the Bunsen burner. The cumulative production of burnable biogas was used as a measure of better performance of the reactors in addition to the measurement of changes in physico-chemical properties during the digestion. The experiment was carried out in a batch reaction mode for 35 days at mesophilic temperature (18–35°C).

Analytical methods

Laboratory analyses of total solids (TS), volatile solids (VS), chemical oxygen demand (COD), biological oxygen demand (BOD), total coliforms (TC), faecal coliforms (FC),

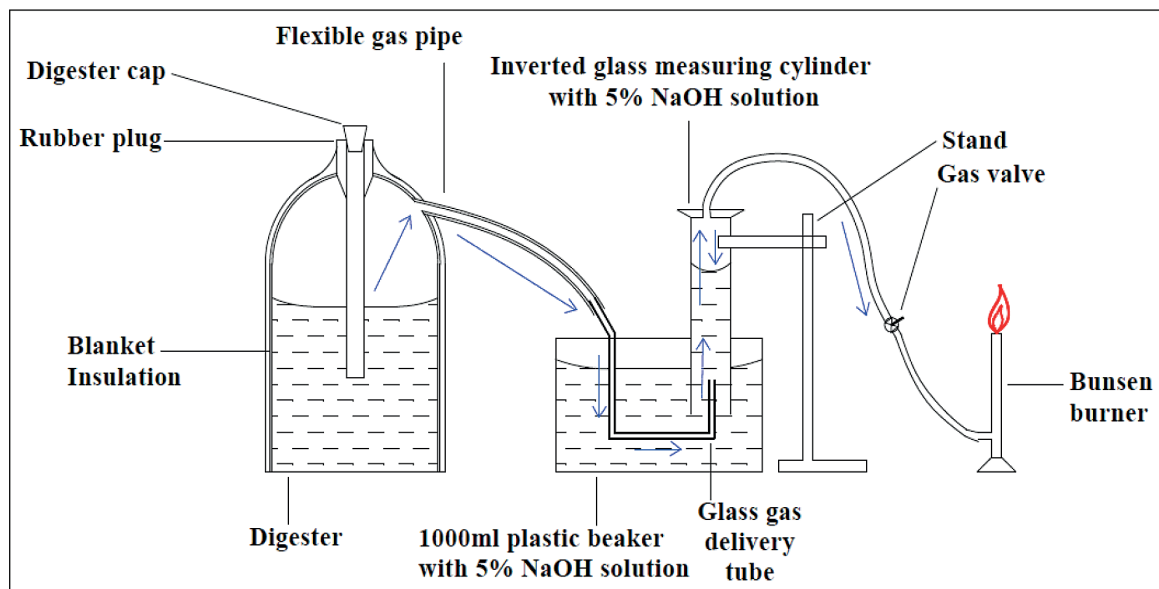


Figure 1. Laboratory set-up for the anaerobic digestion of pit latrine sludge through the displacement method. The biogas produced in the digester was collected in the inverted measuring cylinder containing sodium hydroxide.

Escherichia coli (*E. coli*), helminth eggs, total nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg) and sodium (Na) were measured in triplicate before and after a 35-day mesophilic AD in all the digesters for the 6 sampling times. The total chemical nutrients (N, P, K, Ca, Mg and Na) and the physico-chemical parameters (TS, VS, COD and BOD) were determined according to APHA/AWWA/WEF (2005) and Reddy (2013). Total coliforms, faecal coliforms and *E. coli* were determined using a Most Probable Number assay (MPN) method 8001A (USEPA, 2012). Helminth egg determination was done through the modified Bailenger method (Ayres et al., 1996).

DNA extraction

The DNA was extracted from 150 mg of each sample using the ZR Fecal DNA MiniPrep kit (Zymo research) following the manufacturer's instructions. The DNA was extracted from Day 0, Day 14 and Day 35 sludge samples during AD.

Polymerase chain reaction (PCR) and next generation sequencing (NGS)

The V1 and V3 variable regions were amplified in a 25 μ L reaction using Q5 Hot start High-Fidelity 2x Master Mix (New England Biolabs, USA). Amplicon library PCR was performed on all replicate extractions separately. The DNA primers used were 27F-16S AGAGTTTGATCMTGGCTCAG, 518R-16S ATTACCGCGGCTGCTGG. Thermal cycler setting for PCR amplification were as follows: (i) initial denaturation at 95°C for 2 min, (ii) 30 cycles of 95°C for 20 s, (iii) 55°C for 30 s, (iv) 72°C for 30 s and final elongation at 72°C for 5 min. The amplicon libraries were purified using the Agencourt AMPure XP bead protocol (Beckman Coulter, USA). Library concentration was measured using Nebnext Library quant kit (New England Biolabs, USA) and quality validated using Agilent 2100 Bio analyser (Agilent Technologies, USA). The samples were pooled in equimolar concentrations and diluted to 4nM based on library concentrations and calculated amplicon sizes. The library pool was sequenced on a MiSeq

(illumina, USA) using a MiSeq Reagent kit V3 600 cycle PE (illumina, USA). The final pooled library was at 12 pM with 15% PhiX as control. The illumina MiSeq automatically detected the barcodes and subsequently demultiplexed the clusters that passed on-board quality filter checks and the output was formatted as fastq files.

Analysis of pyrosequencing-derived data

Sequence quality checks were analysed using the Ribosomal Database Project (RDP) pipeline. The sequences were cleaned using the RDP GL FLX software Release 11 (<http://rdp.cme.msu.edu/>) (Cole et al., 2014), with the following filters minimum read Q score of 30, minimum sequence length of 150 bp and maximum sequence length of 500 bp. Raw sequence reads were filtered before subsequence analyses to minimize the effect of random sequencing errors. The pipeline initial process was used to remove sequences of low quality.

After the initial processing, the high-quality sequences were then assigned to the bacterial taxonomy using the RDP Classifier (Wang et al., 2007) at 80% confidence interval. Classification of the newly generated sequences was based on the trained RDP Naive Bayesian rRNA Classifier (Version 2.2, March 2010) of the RDP (Wang et al., 2007). Sequence alignment of the effective bacterial sequences of each sample was done using fast, secondary-structure aware Infernal aligner version 1.1rc4 (Nawrocki et al., 2009). After sequence alignment an aligned file was obtained from each sample. The aligned files were used to cluster the sequences of each sample using the RDP Complete Linkage Clustering tool and a cluster file obtained. The comparison of the microbial communities between Day 0, Day 14 and Day 35 was conducted using the RDP library compare tool at 80% confidence interval.

Statistical analysis

The software, Primer 7 (Primer-E) version 7.0.13 (Quest Research Limited, Auckland, New Zealand) and Multivariate Statistical Package (MVSP) version 3.1 (KOVACH Computing

Services, United Kingdom) were used for statistical analyses. Hierarchical clustering and multidimensional scaling (MDS) were based on a similarity matrix generated from standardized and normalized data. The Euclidean distance was used for environmental data. The hierarchical clustering was conducted using the group average method. Multidimensional scaling ordinations were based on 10 iterations and cluster overlays were based on cluster analysis. ANOSIM (analysis of similarity) was used to test the strength of separation and for significant differences between the different clusters and pit latrine sludge samples. Principle component analysis (PCA) was preceded by a distance matrix of the dataset. MVSP was used for canonical correspondence analysis (CCA). A paired t-test was performed to test for significant differences between undigested and anaerobically digested pit latrine sludge physico-chemical and microbial parameters per pit per cycle of digestion.

RESULTS AND DISCUSSION

Bioreactor performance

The temperature in all the digesters ranged between 23°C and 34°C (Fig. 2a). The highest temperature was recorded at Days 14 and 15 and the lowest temperature in all digesters was recorded at the first and the last days of the digestion process in all the digesters. The temperature remained at mesophilic range throughout the study while pH fluctuated between 8.1 and 5.1 in all the digesters (Fig. 2c). Generally, the pH decreased from the beginning of the digestion and the lowest

pH values were recorded at Days 13 to 16 before increasing again until Day 35 (Fig. 2c). The COD values decreased at a faster rate, mainly between Day 0 and Day 14 and thereafter continued to decrease up to Day 35 but at a much lower rate (Fig. 2b). The general trend of COD was similar to that observed for BOD, TS, and VS.

The cumulative volume of biogas produced from the different pit latrine sludge digesters ranged from 0.001 to 0.003 m³/KgVSa with maximum daily production observed at Day 14 (Fig. 2d). The biogas ignited in all reactors with the earliest combustion recorded on Day 6.

General trends of physico-chemical properties on net effect of AD on pit latrine sludge quality

Based on nMDS analysis (data not shown) the undigested sludge samples were classified separately from treated samples, irrespective of the space (pit) and time (sampling period) in which the samples were collected. The physico-chemical properties of treated and untreated sludge were distinct enough to be classified into these separate categories irrespective of pit or sampling period. The PCA confirmed the nMDS result that undigested sludge samples classified separately from treated samples and showed that TS, VS, COD and BOD mostly accounted for the variability between the digested and undigested sludge treatments (Fig. 3).

The process of AD resulted in changes in physico-chemical properties as highlighted by the changes in pH and COD (Fig. 2). Cluster analysis (data not shown), nMDS and ANOSIM (data

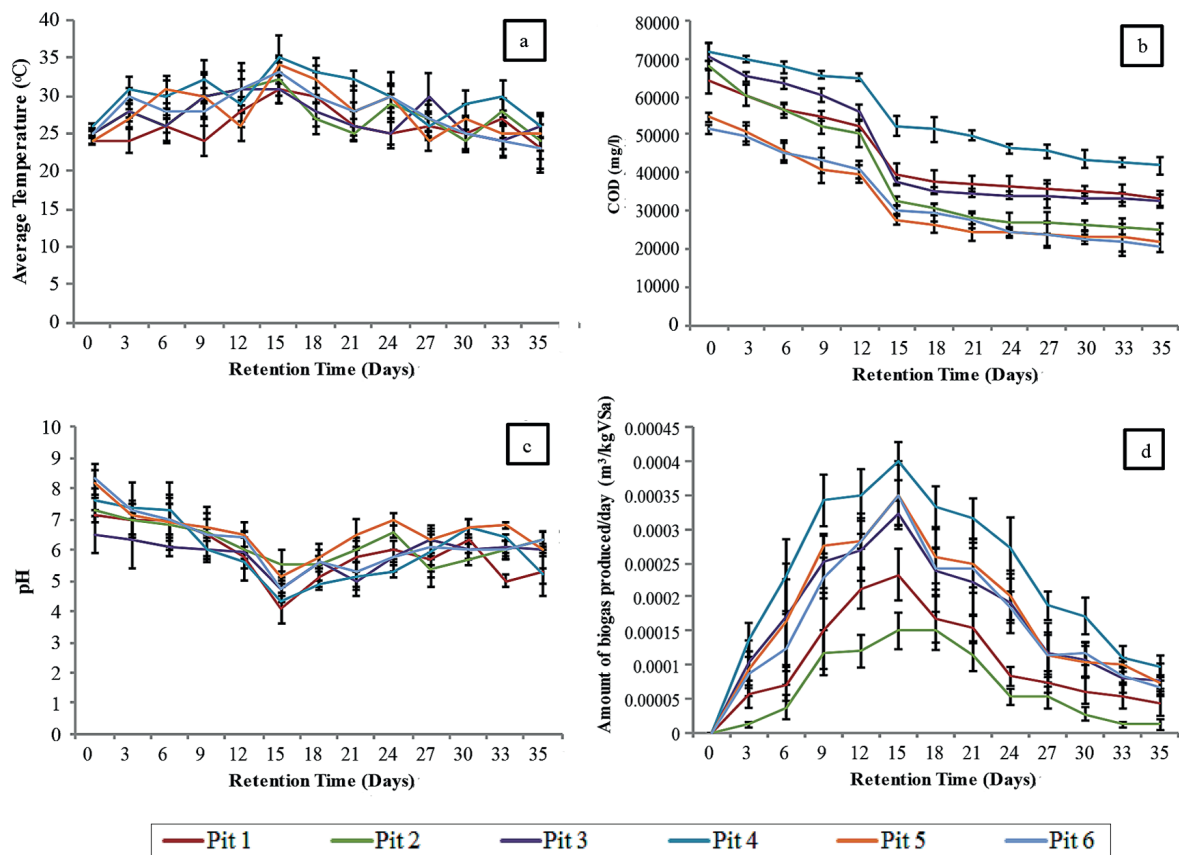


Figure 2. Physico-chemical changes during 35-day mesophilic anaerobic digestion of pit latrine faecal sludge (a) temperature, (b) COD, (c) pH and (d) biogas production

not shown) results also indicated that the mesophilic AD of pit latrine sludge resulted in significant changes in physico-chemical characteristics of the sludge. Biogas production (Fig. 2d) confirmed that the process of AD was complete in each digester.

Effect of anaerobic digestion on selected individual pollution indicators

At the level of individual pollution indicators, AD significantly reduced COD and BOD ($p < 0.05$) in all the pit latrine sludge samples (Table A1, Appendix). The percentage reduction for the COD and BOD ranged from 42–63% and 34–52%, respectively (Table A1, Appendix). These are comparable to the values of 58–61% reported by Issah et al. (2012) and Marti et al. (2008). The COD concentrations in the digested sludge from all the digesters, which ranged from $38\,733 \pm 10\,452$ to $53\,563 \pm 6\,875$ mg/L (Table A1, Appendix) were, however, still above the environmental safety disposal requirements of the Zimbabwe Regulations (Government of Zimbabwe, 2007) of < 60 mg/L of COD by approx. 3 orders of magnitude. Thus the digested pit latrine sludge cannot be disposed into the environment without further treatment. The BOD strength of the anaerobically digested pit latrine sludge, which ranged from $8\,048 \pm 1\,179$ to $10\,916 \pm 743.9$ mg/L, also remained higher than the WHO BOD₅ (1989) standard of 20–100 mg/L for agricultural use and disposal to the environment.

Feecal coliforms (FC) and *E. coli* were also significantly reduced ($p < 0.05$) by the process of AD (Fig. A1). The percentage reductions of FC and *E. coli* ranged from 33–61% in

all the samples (Table B.1). Helminth eggs were not detected in any of the pit latrine sludge samples, both before and after AD. The changes in ecological conditions during the AD process, including temperature and pH, eliminate the sensitive bacteria, including coliforms (Pandey and Soupir, 2011). In comparison, other studies have reported 50–70% reduction in pathogens using AD under mesophilic conditions (Sunders and Harrison, 2011; Issah et al., 2012), which was slightly higher than the percentage reduction in this study.

Although significant reduction of FC occurred in each digester during AD, the number of FCs in the digested sludge, which ranged from $1.0 \times 10^6 \pm 2.7 \times 10^4$ to $2 \times 10^6 \pm 1.8 \times 10^5$ MPN/100 mL, still exceeded the environmental safety disposal limit of $< 1\,000/100$ mL imposed by the Zimbabwe Regulations (Government of Zimbabwe, 2007).

Overall, the process of AD, as expected, significantly reduced the concentration and the counts of pollution indicators (COD, BOD, FC and *E. coli*) in the undigested sludge in comparison to the digestate. The concentration and the counts of the pollution indicators (COD and BOD) in the digestate were, however, still above the WHO (2006) guidelines for re-use of the waste in agriculture. It has to be acknowledged that the process of mesophilic AD, in as much as it significantly reduces pollution indicators in the pit latrine sludge, did not reduce these to levels acceptable for re-use or safe disposal. The process of mesophilic AD of pit latrine sludge on its own cannot thus be recommended for treatment of pit latrine sludge for safe disposal. The main thrust of this research was, however, to determine the extent to which (in terms of percentage reduction) the process of mesophilic AD reduces pollution indicators in pit latrine sludge, as a first step (baseline research) towards treating pit latrine sludge and to inform further studies.

Bacterial community structure changes during AD

Estimates of diversity

A total of 15 087, 7 683 and 9 715 pyrosequences were generated at Day 0, Day 14 and Day 35, respectively (Table 1). After sorting and trimming processes, a total of 11 478, 5 880 and 7 690 sequences of acceptable quality were obtained for Day 0, Day 14 and Day 35, respectively. These sequences, with an average length of 293–295 bp, were used to construct 16S rRNA gene libraries for subsequent analysis.

Diversity indices provide more information about species representation in community composition than just species richness. In this study the Shannon–Weaver Index (*H*) and the Chao 1 estimator were used to estimate the statistical parameters that described each community structure. The analysis of evenness (*E*) was based on the OTU data set. In general, all the samples had balanced populations that had high diversity indices, evenness and species richness. It was difficult to compare the diversity parameters in this study with other studies because of the differences in the reads generated from the pyrosequencing.

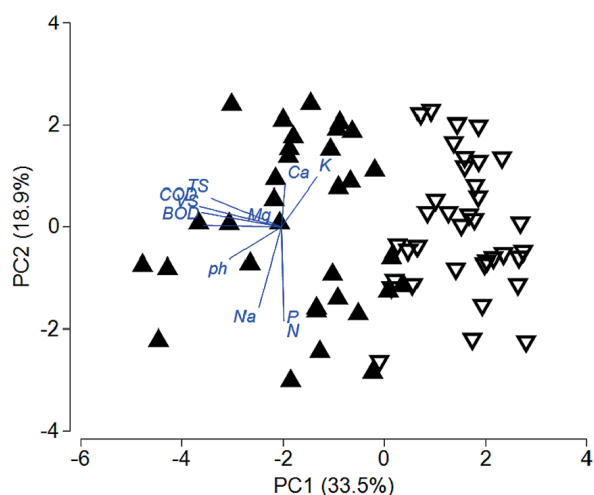


Figure 3. Principal component analysis (PCA) for undigested (▲) and digested (▼) sludge, mapping the undigested and digested samples over sampling periods based on the sludge physico-chemical characteristics. The blue lines show the extent of variability based on eigenvalues in relation to digested and undigested samples as data points. Analysis was based on the Euclidean distance matrix of the sludge physico-chemical characteristics.

Table 1. Estimation of bacterial diversity within gene libraries at 0.15 cluster distance

Library	Bacteria						
	Evenness Index (E)	Shannon Index (H)	Chao 1 Estimator	Richness (no. of OTUs)	Total sequences generated	Sequences after trimming	Average sequence length after trimming
Day 0	0.83	6.54	3 832.6	2 681.6	15 087	11 476	293
Day 14	0.80	5.86	2 203.5	1 441.2	7 683	5 880	293
Day 35	0.79	6.41	2 336.6	1 594.8	9 715	7 690	295

A total of 9 bacterial phyla (Fig. 4) were identified in the undigested pit latrine sludge that was inoculated with 10% (v/v) cow rumen fluid. The dominant bacteria in the undigested pit latrine sludge inoculated with rumen were of the phylum Proteobacteria 26.5%, then Firmicutes 25.1% followed by Spirochaetes (4.7%), Bacteroidetes 2.2% and Actinobacteria 0.9%. The rest of the bacterial phyla included the Planctomycetes, Synergistetes, Verrucomicrobia and Lentisphaerae, which were underrepresented (Fig. 4).

Comparison of microbial community composition clearly showed how this varied between the anaerobically digested sludge and the undigested samples. The Bacteroidetes increased from 2.2% at Day 0 to 16.3% at Day 14 and 32.1% at Day 35. Firmicutes increased from 22.4% at Day 0 to 28.8% at Day 14 before decreasing to 11.6% at Day 35. The Spirochaetes increased from 4% at Day 0 to 5.5% at Day 14 and further increased to 9.2% at Day 35. Proteobacteria decreased significantly to 2.8% at Day 35 from 24.1% at Day 0 to 2.5% at Day 14. All the changes were significant at $p < 0.05$.

It must be acknowledged that the bacterial composition reported here was after the addition of 10% cow rumen which could affect sludge bacterial composition. However, a comparison of the pit latrine sludge alone and that mixed with 10% cow rumen (Table 2) shows that the percentage distribution of the dominant phyla was the same.

A comparison of the bacterial phyla in this study after the addition of rumen with other pit latrine sludge samples and with rumen bacterial composition from literature suggest that the trends of the bacterial composition in this study follow that of the pit latrine sludge rather than cow rumen, though the contribution of the rumen fluid cannot be overruled. For example, Proteobacteria represented 26.5% of the bacterial community in our study, and, in another study, 25% in pit latrine sludge samples from Vietnam (Torendel et al., 2016), whereas a study by Ozbayram et al. (2018) found no Proteobacteria in in cattle rumen. Ozbayram et al. (2018) highlighted that the dominant phyla in cattle rumen were Bacteroidetes (54%); in this study Bacteroidetes contribution was 2.2%. In the same

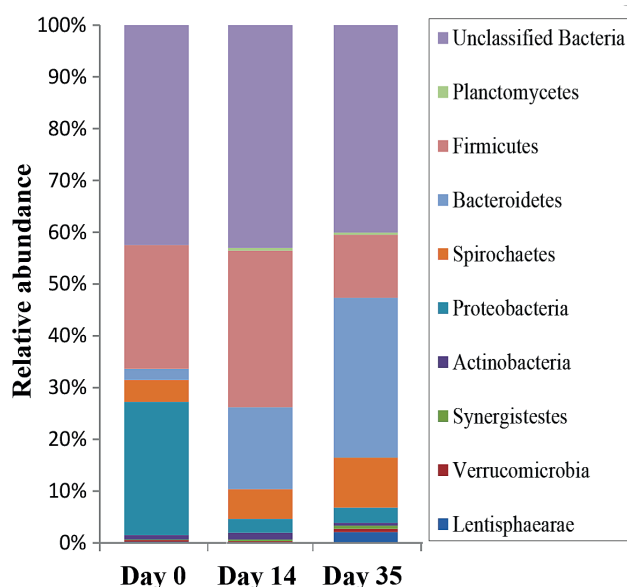


Figure 4. Phylum-level classification of bacterial diversity and community composition during anaerobic digestion of pit latrine faecal sludge inoculated with cow rumen

Table 2. Comparison of dominant bacterial phyla composition (%) in pit latrine sludge alone and pit latrine sludge mixed with 10% cow rumen

Bacterial phyla	Pit latrine sludge alone bacterial phyla composition (%)	Pit latrine sludge mixed with 10% cow rumen bacterial phyla composition (%)
Proteobacteria	27.3	26.5
Firmicutes	26.4	25.1
Spirochaetes	3.9	4.7
Bacteroidetes	1.4	2.2
Actinobacteria	1.2	0.9

study of cattle rumen by Ozbayram et al. (2018), Fibrobacteres and Lentisphaerae contributed 12% and 8%, respectively, but these phyla were not present or underrepresented in our study as well as in other pit latrine sludge samples from Tanzania and Vietnam (Torendel et al., 2016).

The process of AD was characterised by changes in community structure that were associated with biochemical and physico-chemical changes. These changes are usually attributed to the processes of hydrolysis, acidogenesis, acetogenesis and methanogenesis. The microbial diversity observed in the undigested sludge (Day 0) are associated with the early stages of AD and minimal activity in terms of biogas production. The activity (biogas production) increased reaching a peak at Day 14 which also corresponded with the highest temperature reached on this day (Fig. 2). Day 14 was therefore associated with maximum activity or maximum cumulative activity and the turning point towards the decline of activity (biogas production and temperature) during the anaerobic digestion (Fig. 2). Day 35 presented the latter stages of the process, which were characterised by declining and almost stationary biogas production (Fig. 2).

Our findings indicated that Proteobacteria, Firmicutes, Spirochaetes, Bacteroidetes and Actinobacteria were the major bacterial phyla commonly found in fresh faecal sludge in the early stages of digestion that occurs in the uppermost layer of pit latrines. The bacterial composition can be altered by a number of environmental factors, including anal cleansing and bathing material, and the design of the latrine, amongst other factors (Torendel et al., 2016). The Firmicutes, Bacteroidetes, Proteobacteria and Actinobacteria have been reported to be the most abundant phyla in pit latrine sludge bacterial communities (McLellan et al., 2010; Torendel et al. 2016). Sewage sludge profiles were also found to be dominated by the same phyla (Shanks et al., 2013). Firmicutes are common inhabitants of the human gut (Gill et al., 2006) and environments where AD occurs like activated sludge systems and anaerobic digesters (Garcia-Pena et al., 2011). One of the dominant genera, *Treponema* (Table A3, Appendix), is from the phylum Spirochaetes, which have previously been shown to be important components of the gut microbiota in individuals from rural, less developed regions (Schnorr et al., 2014), which is relevant to this study as the sludge was collected from a peri-urban area which is considered to be a poor area, for example, in terms of diet. Proteobacteria are important microbes in the AD process because most of the Alpha-, Beta-, Gamma-, and Deltaproteobacteria are well known glucose, propionate, butyrate, and acetate-utilizing microbial communities (Ariesyady et al., 2007).

As the AD process occurred, there were notable changes in numbers of the inherently emerging bacterial populations and disappearance of those involved in the early stages of

AD, showing that the process involves microbial succession and specific groups of microorganisms drive the specific biotransformations of sludge material towards completion. A shift in this succession could therefore be detrimental to the efficient production of biogas or effective treatment of the faecal sludge. Understanding these changes in microbial community structure is therefore very important for monitoring and improving the efficiency of AD of faecal sludge and derivation of maximum value. The turning point of activity, Day 14, was characterised by drastic shifts of bacteria at phylum level, with an increase in Firmicutes of the class Clostridia, Spirochaetes of the class Spirochaetea and Bacteroidetes of the classes Bacteroidia and Flavobacteria. These results suggest that an increase in numbers occurred in those bacterial groups actively involved in the turning point of the process or simply those adaptable to the emerging conditions. Most members of Firmicutes are syntrophic bacteria that can degrade various volatile fatty acids (Garcia-Pena et al., 2011). The Bacteroidetes, which increased up to Day 14, are known to produce various lytic enzymes and acetic acid during the degradation of organic materials (Chen and Dong, 2005; Robert et al., 2007; Riviere et al., 2009).

At Day 35, which was considered the end of the digestion process, there was a reduction in Firmicutes of the class Bacilli and further reduction in Proteobacteria of classes Alphaproteobacteria, Gammaproteobacteria and Betaproteobacteria. This reduction in numbers found in the undigested sludge was an indication of those bacteria that were reduced or eliminated by the process of AD or that did not favourably respond to the changes in conditions during the process.

The results obtained in this study in terms of bacterial shifts during AD was comparable to those of Alcantra-Hernandez et al. (2017) and Doloman et al. (2017), who reported that Proteobacteria diminished from 33% to 10–16% whilst Spirochaetes and Bacteroidetes increased as a result of AD. Their results were, however, different in relation to Firmicutes; we observed a significant decrease in Firmicutes as a result of AD, whilst Alcantra-Hernandez et al. (2017) and Doloman et al., (2017) reported an increase of Firmicutes as a result of AD. The waste material for the study by Alcantra-Hernandez et al. (2017) was organic urban residues and for that of Doloman et al. (2017) was microbial biomass, both being operated at mesophilic conditions.

Association between bacterial community structure and environmental properties at Days 0, 14 and 35 of AD

Using canonical correspondence analysis (CCA) at genus level. *Treponema*, *Alkaligenes* and *Celerinatantimonas* were associated with high COD, BOD and VS at Day 0. *Prevotella*, *Hellela*, and *Sarcina* were associated with high production of biogas at Day 14, which was the turning point for all the measured physico-chemical properties. *Mangroviflexus*, *Alkalitalea* and *Oligosphaera* were associated with high K and low COD, BOD and VS at Day 35 (Fig. 4).

It should be appreciated that no two pits are alike; even the same pit sampled at different depths is not alike (Bakare et al., 2012). What we have assessed here is a general trend from specific pit latrine sludge in Shackleton, but specific details could differ from pit to pit and area to area. Factors such as differences in gut microbiota, diet, anal cleansing material and pit latrine maintenance approaches, like different emptying patterns, can result in differing microbial compositions of raw pit latrine sludge (Torendel et al., 2016). The results presented here were obtained from small-scale reactors and thus may not exactly

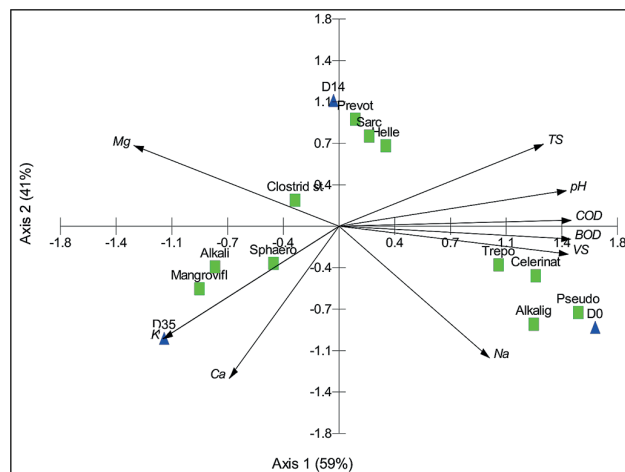


Figure 5: Canonical correspondence analysis ordination plots depicting the relationship between pit latrine sludge environmental characteristics at Days 0 (D0), 14 (D14) and 35 (D35) and bacterial populations at genus level. Key: Pit latrine sludge physico-chemical parameters – K (potassium), Ca (calcium), Na (sodium), VS (volatile solids), BOD (biological oxygen demand), COD (chemical oxygen demand), pH, TS (total solids) and Mg (magnesium). Bacterial genera: *Pseudo* (*Pseudomonas*), *Celerinat* (*Celerinatantimonas*), *Prevot* (*Prevotella*), *Helle* (*Hellela*), *Alkalig* (*Alkaligenes*), *Clostrid* (*Clostridium sensu stricto*), *Sarc* (*Sarcina*), *Treponema* (*Treponema*), *Sphaero* (*Sphaerochaeta*), *Oligosphaera* (*Oligosphaera*), *Mangroviflexus* (*Mangroviflexus*), *Alkalitalea* (*Alkalitalea*), *Clostrid* XI (*Clostridium* XI). The triangles represent the cases (different sampling times during anaerobic digestion) and the squares, either the bacterial phylum or genera. Arrows represent the direction and strength of the variables in relation to physico-chemical characteristics shown as data points. Distance of arrows from the axis centre shows extent of variability based on PCA scores (eigenvalues).

replicate the much bigger digesters and as such the hydraulic retention times, reactor temperatures, and mixing methods will likely require modifications during scale up. However, the results obtained using this anaerobic digester model provides insights into how full-scale reactors may perform.

To the best of our knowledge this is the first study describing the changes in microbial community structure during AD of pit latrine sludge. Most studies have traditionally been based on culture-dependent techniques targeting specific bacterial groups such as *E. coli* and faecal coliforms, which are indicators of faecal origin. The use of next generation sequencing based on the 16S rRNA identification reveals diversity that would not have been shown by culture dependent techniques. The findings from this study indicate the key bacterial groups involved in the AD process, which could be useful for improving the efficiency of the AD process and developing bacteria-based technological solutions for the treatment of pit latrine sludge and energy generation from the sludge.

CONCLUSIONS

The study revealed that the process of mesophilic AD of pit latrine sludge reduces pollution indicators in the sludge. The reduction was, however, insufficient to meet the threshold for safe disposal. The process of mesophilic AD of pit latrine sludge on its own cannot therefore be recommended for treatment of pit latrine sludge for safe disposal. The study, however, has provided information on qualitative and quantitative changes in the levels of pollution indicators during the AD process. The study also showed that there

is wide diversity of bacteria in pit latrine sludge, which changes in community structure as the AD process progresses. The changes in community structure described by this study are considered to be the changes that occur during healthy mesophilic AD of faecal pit latrine sludge. The process may, however, be affected by a number of factors that may alter the succession reported here and subsequently affect the efficiency of the process. The process of mesophilic AD promotes proliferation of Bacteroidetes and Spirochaetes and significantly reduces Proteobacteria and Firmicutes. The information generated will form a basis for future research in understanding the process of mesophilic AD of pit latrine sludge. The information generated can also be used to compare mesophilic AD with other proposed treatment options for pit latrine sludge and could be used to develop microbial-based treatment technologies and management solutions.

ACKNOWLEDGEMENTS

The authors would like to thank the Sanitation Research Fund for Africa (SRFA) for funding this research project through the Water Research Commission (SA), and Chinhoyi University of Technology for providing laboratories to carry out the experiments and allowing for the research to be carried out on campus.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

- ALCANTARA-HERNANDEZ RJ, TAS N, CARLOS-PINEDO S, DURAN-MORENO A and FALCON LI (2017) Microbial dynamics in anaerobic digestion reactors for treating organic urban residues during the start-up process. *Let. Appl. Microbiol.* **64** 438–445. <https://doi.org/10.1111/lam.12734>
- APHA/AWWA/WEF (2005) Standard Methods for the Examination of Water and Wastewater, 20th edition. American Public Health Association/American Water Works Association/Water Environment Federation, Washington DC, USA. <https://doi.org/10.12999/awwa.g510.13>
- AYRES RM, MARA DD and RACHEL M (1996) *Analysis of Wastewater for Use in Agriculture – A Laboratory Manual of Parasitological and Bacteriological Techniques*. World Health Organisation, Geneva, Switzerland. 1–35. URL: <http://www.who.int/iris/handle/10665/41832> (Accessed 3 February 2015).
- ARIESYADY HD, ITO T and OKABE S (2007) Functional bacterial and archaeal community structures of major trophic groups in a full-scale anaerobic sludge digester. *Water Res.* **41** 1554–1568. <https://doi.org/10.1016/j.watres.2006.12.036>
- BAKARE BF, FOXON KM, BROUCKAERT CJ and BUCKLEY CA (2012) Variation in VIP latrine sludge contents. *Water SA* **38** 479–486. <https://doi.org/10.4314/wsa.v38i4.2>
- BOOT NLD and SCOTT RE (2008) Faecal sludge management in Accra Ghana: strengthening links in the chain. In: *Proceedings of the 33rd WEDC International Conference – Access to sanitation and safe water global partnerships and local actions*, 7–11 April 2018, Accra, Ghana. <https://doi.org/10.3362/1756-3488.2008.007>
- CHEN S and DONG X (2005) Proteiniphilum acetatigenes gen. nov., sp. nov., from a UASB reactor treating brewery wastewater. *Int. J. Syst. Evol. Microbiol.* **55** (6) 2257–2261. <https://doi.org/10.1099/ijs.0.63807-0>
- COLE JR, WANG JA, FISH B, CHAI DM, MCGARRELL Y, SUN CT, BROWN A, PORRAS-ALFARO C, KUSKE R and TIEDJE JM (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucl. Acids Res.* **42** (Database issue) D633–D642. <https://doi.org/10.1093/nar/gkt1244>
- COLON J, FORBIS-STOKES AA and DESHUSSES MA (2015) Anaerobic digestion of undiluted simulant human excreta for sanitation and energy recovery in less-developed countries. *Energ. Sustain. Dev.* **29** 57–64. <https://doi.org/10.1016/j.esd.2015.09.005>
- DRANGERT JO (1998) Fighting the urine blindness to provide more sanitation options. *Water SA* **24** (2) 157–164.
- DOLOMAN A, SOBOH Y, WALTERS AJ, SIMS RC and MILLER CD (2017) Qualitative analysis of microbial dynamics during anaerobic digestion of microalgal biomass in a UASB reactor. *Int. J. Microbiol.* (2017) 1–12. <https://doi.org/10.1155/2017/5291283>
- FORBIS-STOKES AA, O'MEARA PF, MUGO W, SIMIYU GM and DESHUSSES MA (2016) On-site fecal sludge treatment with the anaerobic digestion pasteurization latrine. *Environ. Eng. Sci.* **33** (11) 898–906. <https://doi.org/10.1089/ees.2016.0148>
- GARCIA-PENA EL, PARAMESWARAN P, KANG DW, CANUL-CHAN M and KRAJMALNIK-BROWN R (2011) Anaerobic digestion and co-digestion processes of vegetative and fruit residues: Process and microbial ecology. *Bioresour. Technol.* **102** 9447–9455. <https://doi.org/10.1016/j.biortech.2011.07.068>
- GIJZEN HJ (2002) Anaerobic digestion for sustainable development: A natural approach. *Water Sci. Technol.* **45** (10) 321–328. <https://doi.org/10.2166/wst.2002.0364>
- GILL SR, POP M, DEBOY RT, ECKBURG PB, TURNBAUGH PJ, SAMUEL BS, GORDON JI, RELMAN DA, FRASER-LIGGETT CM and NELSON KE (2006) Metagenomic analysis of the human distal gut microbiome. *Science* **312** 1355–1359. <https://doi.org/10.1126/science.1124234>
- GOVERNMENT OF ZIMBABWE (2007) Environmental Management (Effluent and Solid Waste Disposal) Regulations (2007) Supplement to the Zimbabwean Government Gazette dated 5th of January, 2007. Government Printer, Harare.
- HUTTON G, HALLER L and BARTRAM J (2007) Global cost-benefit analysis of water supply and sanitation interventions. *J. Water Health* **5** (4) 481–502. <https://doi.org/10.2166/wh.2007.009>
- ISSAH A, AKLAKU ED and SALIFU T (2012) Comparative study of effluent for pollution indicators and indicator pathogenic organisms from anaerobic digesters from human and fruit wastes. *ARPN J. Agric. Biol. Sci.* **7** 416–419.
- KATUKIZA AY, RONTELTAP M, NIWAGABA CB, FOPPEN JW, KANSIIME F and LENS PN (2012) Sustainable sanitation technology options for urban slums. *Biotechnol. Adv.* **30** (5) 964–978. <https://doi.org/10.1016/j.biotechadv.2012.02.007>
- LALANDER CH, HILL GB and VINNERAS B (2013) Hygienic quality of faeces treated in urine diverting vermicomposting toilets. *Waste Manage.* **33** (11) 2204–2210. <https://doi.org/10.1016/j.wasman.2013.07.007>
- LUKEHURST CT, FROST P and AL SEADI T (2010) Utilization of digestate from biogas plants as bio-fertilizer. *IEA Bio-Energy* 5–16.
- MARA D (2013) Pits, pipes, ponds—And me. *Water Res.* **47** (7) 2105–2117. <https://doi.org/10.1016/j.watres.2013.01.051>
- MARTI N, BOUZAS A, SECO A and FERRER J (2008) Struvite precipitation assessment in anaerobic digestion processes. *Chem. Eng. J.* **141** 67–74. <https://doi.org/10.1016/j.cej.2007.10.023>
- MCCARTY PL, BAE J and KIM J (2011) Domestic wastewater treatment as a net energy producer—can this be achieved? *Environ. Sci. Technol.* **45** (17) 7100–6. <https://doi.org/10.1021/es2014264>
- MCLELLAN SL, HUSE S M, MUELLER-SPITZ SR, ANDREISHCHEVA EN and SOGIN ML (2010) Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environ. Microbiol.* **12** (2) 378–392. <https://doi.org/10.1111/j.1462-2920.2009.02075.x>
- MONTGOMERY MA and ELIMELECH M (2007) Water and sanitation in developing countries: including health in the equation. *Environ. Sci. Technol.* **41** (1) 17–24. <https://doi.org/10.1021/es072435t>
- NAWROCKI EP, KOLBE DL and EDDY SR (2009) Infernal 1.0: inference of RNA alignments. *Bioinformatics* **25** (10) 1335–1337. <https://doi.org/10.1093/bioinformatics/btp157>
- NARIHIRO T, KIM N. K, MEI R, NOBU MK and LIU WT (2015) Microbial community analysis of anaerobic reactors treating soft drink wastewater. *PLoS ONE* **10** (3) e0119131. <https://doi.org/10.1371/journal.pone.0119131>
- NWANERI CF, FOXON K, BAKARE BF and BUCKLEY C (2008) Biological degradation processes within a pit latrine. In:

Proceedings of WISA 2008 Conference, Sun City, South Africa, 19–22 May 2008.

- ONABANJO T, KOLIOS AJ, PATCHIGOLLA K, WAGLAND ST, FIDALGO B, JURADO N, HANAK DP, MANOVIC V, PARKER A, MCADAM E, WILLIAMS L, TYRREL S and CARTMELL E (2016) An experimental investigation of the combustion performance of human faeces. *Fuel* **184** 780–791. <https://doi.org/10.1371/journal.pone.0119131>
- OJOLO SJ, OKE SA, ANIMASAHUM K and ADESUYI BK (2007) Utilization of poultry, cow and kitchen wastes for biogas production: a comparative analysis. *Iran J. Environ. Health Sci. Eng.* **4** (4) 223–228.
- OZBAYRAM EG, INCE O, INCE B, HARMS H and KLEINSTEUBER S (2018) Comparison of rumen and manure microbiomes and implications for the inoculation of anaerobic digesters. *Microorganisms* **2018** (6) 15. <https://doi.org/10.3390/microorganisms6010015>
- PARK JA, HUR JM, SON BS and LEE JH (2001) Effective treatment of night soil using anaerobic sequencing batch reactor (ASBR). *Korean J. Chem. Eng.* **18** (4) 486–492. <https://doi.org/10.1007/bf02698295>
- PANDEY PK and SOUPIR ML (2011) *Escherichia coli* inactivation kinetics in anaerobic digestion of dairy manure under moderate mesophilic and thermophilic temperature. *AMB Express* **1** 18. <https://doi.org/10.1186/2191-0855-1-18>
- REDDY M (2013) Standard Operating Procedures. Howard College, School of Chemical Engineering; Population Research Fund. University of KwaZulu-Natal, Durban.
- RIVIERE D, DESVIGNES V, PELLETIER E, CHAUSSONNERIE S, GUERMAZI S, WEISSENBACH J, LI T, CAMACHO P and SGHIR A (2009) Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *ISME J.* **3** (6) 700–714. <https://doi.org/10.1038/ismej.2009.2>
- ROBERT C, CHASSARD C, LAWSON PA and BERNALIER-DONADILLE A (2007) *Bacteroides cellulosilyticus* sp. nov., a cellulolytic bacterium from the human gut microbial community. *Int. J. Syst. Evol. Microbiol.* **57** (7) 1516–1520. <https://doi.org/10.1099/ijs.0.64998-0>
- SAUNDERS O and HARRISON J (2011) Pathogen Reduction in Anaerobic Digestion of Manure. URL: <http://www.extension.org/pages/30309/pathogen-reduction-in-anaerobic-digestion-of-manure> (Accessed 5 February 2016).
- SAHLSTROM L, ASPAN A, BAGGE E, DANIELSSON-THAM ML and ALBIHN A (2004) Bacterial pathogen incidences in sludge from Swedish sewage treatment plants. *Water Res.* **38** (8) 1989–1994. <https://doi.org/10.1016/j.watres.2004.01.031>
- SCHNORR SL, CANDELA M, RAMPPELLI S, CENTANNI M, BASAGLIA G, TURRONI S, BIAGI E, PEANO C, SEVERGNINI M, FIORI J and co-authors (2014) Gut microbiome of the Hadza hunter-gatherers. *Nat. Commun.* **5** (3654) 1–12. <https://doi.org/10.1038/ncomms4654>
- SCHONNING C and STENSTROM TA (2004) Guidelines for the safe use of urine and faeces in ecological sanitation systems. *EcoSanRes Publications Series Report 2004-1*, Stockholm Environment Institute, Stockholm.
- SHANKS OC, NEWTON RJ, KELTY CA, HUSE SM, SOGIN ML and MCLELLAN SL (2013) Comparison of the microbial community structures of untreated wastewaters from different geographic locales. *Appl. Environ. Microbiol.* **79** 2906–2913. <https://doi.org/10.1128/aem.03448-12>
- TORONDEL B, ENSINK JH, GUNDOGDU O, IJAZ UZ, PARKHILL J, ABDELAHI F, NGUYEN VA, SUDGEN S, GIBSON W, WALKER AW and co-authors (2016) Assessment of the influence of intrinsic environmental and geographical factors on the bacterial ecology of pit latrines. *Microb. Biotechnol.* **9** (2) 209–223. <https://doi.org/10.1111/1751-7915.12334>
- USEPA (United States Environmental Protection Agency) (2012) Method 8001A: Total Coliforms, Faecal Coliforms and *Escherichia coli* in wastewater by Most Probable Number Method. URL: <http://www.epa.gov/microbes/1604sp02.pdf> (Accessed 23 January 2017).
- VENKITESHWARAN K, BOCHER B, MAKI J and ZITOMER D (2016) Relating anaerobic digestion microbial community and process function. *Water Microbiol.* **8** (S2) 37–44. <https://doi.org/10.4137/mbi.s33593>
- WANG QG, GARRITY M, TIEDJE JM and COLE JR (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* **73** (16) 5261–5267. <https://doi.org/10.1128/aem.00062-07>
- WHO (2006) *Guidelines for the Safe Use of Wastewater, Excreta and Grey Water*. World Health Organisation, Geneva.
- WHO (1989) Health guidelines for the use of wastewater in agriculture and aquaculture. Report of a World Health Organization Scientific Group, WHO Technical Report Series 778. World Health Organization, Geneva. <https://doi.org/10.1159/isbn.978-3-318-05724-9>
- WHO/UNICEF (2017) Progress on Drinking Water, Sanitation and Hygiene-2017 update and SDG Baselines. WHO/UNICEF, Geneva.

APPENDIX

Table A1: Effect of anaerobic digestion of pit latrine faecal sludge from Shackleton on COD, BOD, total solids and volatile solids ($n = 18$ per pit).

Pit Number	COD (mg/L)				BOD (mg/L)			
	Fresh	Digested	% reduction	<i>p</i> value	Fresh	Digested	% reduction	<i>p</i> value
1	84 049 ± 9 564	43 019 ± 4 593	48.8	***	14 987 ± 2 404	9 442 ± 1 470	37	***
2	74 713 ± 3 989	43 508 ± 3 640	41.7	***	14 736 ± 4 433	9 799 ± 1 211.5	34	***
3	115 754 ± 13 389	53 563 ± 6 875	53.7	***	18 329 ± 2 976	10 916 ± 743.9	40	***
4	128 212 ± 13 864	47 598 ± 11 121	63	***	22 249 ± 2 985	10 669 ± 461	52	***
5	99 523 ± 9 073	41 583 ± 8 911	58.2	***	16 595 ± 2 992	8 048 ± 1 179	52	***
6	97 341 ± 12 328	38 733 ± 10 452	60.2	***	15 055 ± 2 570	8 664 ± 735	43	***
	Total solids (mg/L)				Volatile solids (mg/L)			
1	39 537 ± 3 114	31 181 ± 3 698	21	***	20 175 ± 2 540	8 672 ± 1 873	57	***
2	28 212 ± 2 635	20 461 ± 1 964	27	***	11 163 ± 3 525	4 803 ± 1 213	55	***
3	55 345 ± 4 986	41 074 ± 6 454	24	***	24 646 ± 6 382	10 260 ± 2 241	52	***
4	62 698 ± 2 078	51 644 ± 1 568	17	***	33 910 ± 9 290	7 046 ± 5 553	79	***
5	51 687 ± 2 206	40 628 ± 2 480	21	***	22 565 ± 1 716	7 314 ± 1 607	68	***
6	46 276 ± 5 646	33 471 ± 5 307	27	***	21 944 ± 4 497	6 741 ± 1 994	69	***

± = standard deviation, p*** = significant difference, p** = no significant difference between undigested and digested pit latrine sludge

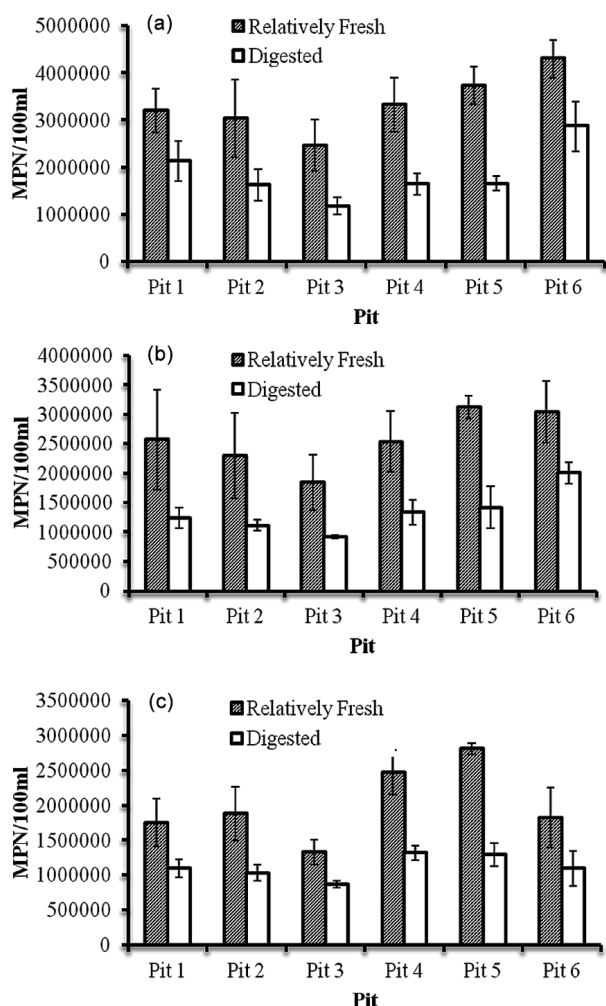


Figure A1: Effect of anaerobic digestion on (a) total coliforms, (b) faecal coliforms and (c) *E. coli* counts in pit latrine faecal sludge from the peri-urban settlement of Chinhoyi, Zimbabwe ($n = 18$ per pit), $p < 0.05$ in all samples before and after digestion.

Table A2: Total coliforms, faecal coliforms and *E. coli* percentage reductions due to the process of anaerobic digestion (*n* = 18 per pit)

Pit	Total coliforms (%)	Faecal coliforms (%)	<i>E. coli</i> (%)
1	33 ± 5	52 ± 12	37 ± 5
2	46 ± 8	52 ± 8	45 ± 10
3	48 ± 5	50 ± 10	35 ± 8
4	61 ± 5	47 ± 5	60 ± 5
5	55 ± 4	54 ± 5	54 ± 10
6	50 ± 10	34 ± 12	40 ± 8

Table A3: Taxonomic composition of bacterial communities at the genus level for the sequences retrieved from each sample at different stage of anaerobic digestion

Phylum	Genus	Relative abundance		
		Day 0	Day 14	Day 35
Actinobacteria	<i>Olsenella</i>	0.1%	0.2%	0.3%
	<i>Corynebacterium</i>	0.3%	0.3%	0.1%**
Firmicutes	<i>Kurthia</i>	1.9%	0%*	0%**
	<i>Planococcaceae- incertae-sedis</i>	0.3%	0%*	0%**
	<i>Lysinibacillus</i>	0.1%	0%	0%
	<i>Rummelibacillus</i>	1.6%	0%*	0%**
	<i>Anaerobacter</i>	0.3%	1.4%*	0.8%
	<i>Clostridium sensu stricto</i>	0.4%	2.9%*	1.7%**
	<i>Sarcina</i>	1.2%	6.8%*	0.1%**
	<i>Clostridium XIVa</i>	0%	0.1%	0.1%
	<i>Turcibacter</i>	0.1%	0.6%*	0.1%
	<i>Clostridium III</i>	0%	0.1%	0.1%
	<i>Bacillus</i>	0.2%	0%	0.1%
	<i>Clostridium XI</i>	0%	4.9%*	1.7%**
	<i>Sporacetigenium</i>	0.2%	0.5%	0.2%
Proteobacteria	<i>Anaerovorax</i>	0.1%	0.1%	0%
	<i>Ignatzschineria</i>	4.2%	0%*	0%**
	<i>Azomonas</i>	0.7%	0%	0%
	<i>Pseudomonas</i>	1.1%	0%*	0%**
	<i>Celerinatantimonas</i>	0.9%	0%*	0%**
	<i>Pussillimonas</i>	0.3%	0%	0%
	<i>Advenella</i>	0.5%	0.1%	0.1%
	<i>Paenalcigenes</i>	0.3%	0%	0%
Bacteroidetes	<i>Castellaniella</i>	1%	0.1%*	0.1%**
	<i>Alkaligenes</i>	1.1%	0%*	0.5%**
	<i>Petrimonas</i>	0.1%	0.5%	0.1%
	<i>Proteiniphilum</i>	0.2%	1%	0.6%
	<i>Prevotella</i>	0.1%	3.1%	0%
	<i>Hellela</i>	0%	1.1%	0%
	<i>Comamonas</i>	0.2%	0%	0.1%
	<i>Phocaeicola</i>	0%	0.1%	0.8%
	<i>Mangroviflexus</i>	0%	0%	4.9%**
	<i>Alkalitalea</i>	0%	0.2%	13.9%**
Spirochaetes	<i>Treponerma</i>	1%	0.5%	0.1%
	<i>Sphaerochaeta</i>	1.6%	3.9%*	7.5%**

* = significantly different at $p < 0.01$ between Days 0 and 14

** = significantly different at $p < 0.01$ between Days 0 and 35

Effects of nitrogen loading from domestic wastewater on groundwater quality

Tamás Mester^{1*}, Dániel Balla¹, Gergő Karancsi¹, Éva Bessenyei¹ and György Szabó¹

¹Department of Landscape Protection and Environmental Geography, University of Debrecen, Egyetem tér 1. H-4032 Debrecen, Hungary

ABSTRACT

In this study the effects of nitrogen effluent from a permeable constructed sewage tank on groundwater quality were investigated. Sampling took place before and 1.5 years after the closure. Using a 3D hydrogeological model, the spatial distribution of dissolved inorganic nitrogen (DIN), comprising the species NH_4^+ , NO_2^- and NO_3^- , was modelled in the saturated zone and the amounts and changes in the ratio of $\text{NH}_4\text{-N}$, $\text{NO}_2\text{-N}$, $\text{NO}_3\text{-N}$ were determined. The first part of our general hypothesis, that the groundwater was heavily contaminated in the area surrounding the sewage tank, was clearly verified, since every investigated nitrogen compound exceeded the contamination limit values; but the second part of our hypothesis, that the degree of contamination would significantly decrease after the sewage tank was taken out of use has not been confirmed, since the amount of nitrogen present in inorganic forms increased in the modelled zone. The increase in DIN and the relatively high concentration of NH_4^+ (35 mg/L) in the immediate vicinity of the tank can be explained by the fact that organic matter (OM) accumulated over the decades provided a constant supply of inorganic nitrogen forms.

Keywords: ammonium, nitrite, nitrate, groundwater quality, pollution, wastewater, nitrogen loading

INTRODUCTION

In inhabited areas the significant nitrogen (N) contamination of groundwater, generated mainly by domestic wastewater input, has now become a crucial issue (Drake and Brauder 2005; Kringel et al., 2016). Domestic sewage typically has total nitrogen concentrations of 20 to 100 mg-L⁻¹, primarily as NH_4^+ (Robertson et al., 2012). Nitrogen enters the shallow groundwater in numerous organic and inorganic forms and, due to their high solubility, is easily transported from septic and sewage tanks to the deeper aquifer zones, thus contaminating the aquifers (Simmons et al., 1992; Heatwole and McCray 2007, Moodley et al, 2017; Adams et al. 2019). Inappropriate wastewater treatment has led to widespread N pollution of aquatic systems, not only in the less-developed areas of the world, but in developed areas as well (Fantong et al., 2013; Goody et al., 2014; Benrabah et al., 2016, Robertson et al., 2016; Adhanom et al., 2018).

Based on investigations carried out in Tennessee, Hanchar (1991) concluded that septic tank effluent affected groundwater quality. The slightly elevated nitrite, nitrate and ammonium content can also be explained as a result of the effluent from the septic-tanks. Reay (2004) reached a similar conclusion after investigating the effects of septic tanks on groundwater. The nitrogen load to septic tanks is significant (5.7 to 10.7 kg-household⁻¹.yr⁻¹) and, as a consequence, the amount of the dissolved inorganic nitrogen (DIN) measured in the shallow groundwater can be 100 times higher than the DIN values measured in the nearby surface water. The research of McQuillan (2004) carried out in New-Mexico demonstrated that contamination originating from septic tanks has the largest effect on the groundwater, when compared to other sources of contamination.

Urbanization and intense agricultural and industrial activity have caused serious pollution of the environment and water resources on the African continent (Attoui et al., 2016). Pollution sources are related to raw sewage and domestic or

industrial wastewater discharges as well as agricultural runoff. In addition, the increase in the number of inhabitants leads to over-exploitation, increasing the vulnerability of groundwater (Nlend et al., 2018). Abdalla and Khalil (2018) investigated the surface and groundwater contamination in Qus City in Egypt. The sewage network is under construction, and the sewage is disposed in uninsulated underground sewage sites, which have direct contact with groundwater; therefore, wastewater infiltrates easily into the shallow aquifer. 94% of groundwater samples showed evidence of mixing between groundwater and wastewater. In the urban environment of Yaounde mass flow of nutrients from latrines and septic tanks was assessed to evaluate the groundwater quality. The rapid increase of electrical conductivity from the background to the built-up areas and the ammonium and nitrate concentrations above the WHO limit suggest anthropogenic effects (Kringel et al., 2016).

Similar problems can be identified in the countries of East-central Europe. In this region the most common contamination sources are pit latrines and uninsulated sewage tanks (Banks et al., 2002; Smoroń 2016; Mester et al., 2017). The N contamination of groundwater of anthropogenic origin is evidenced by several studies. Adumitroaei et al. (2016) investigated the inorganic nitrogen species, in the oxidized (NO_2^- , NO_3^-) and reduced (NH_4^+) forms, in groundwater in the northern and central part of Vaslui County, Romania. Both forms of nitrogen were above concentration limit values proving their anthropogenic origin. Based on investigations carried out in Serbia, Devic et al. (2014) concluded that NO_3^- of anthropogenic origin is one of the key factors impairing groundwater quality.

In the rural areas of Hungary, one of the most important sources of organic and inorganic (NH_4^+ , NO_2^- , NO_3^-) nitrogen compounds is wastewater originating from households, the collection of which remains an unsolved problem in many places (Mester and Szabó 2013; Szabó et al., 2016). Hungary with the accession to the European Union ratified the Water Framework Directive (WFD) (2000/60/EC) and the Urban Wastewater Treatment Directive (UWTD) (271/91/EEC), which regulate the issue of contamination originating from

*Corresponding author, email: mester.tamas@science.unideb.hu
Received 2 July 2018; accepted in revised form 2 July 2019

agriculture and domestic wastewater. The UWTD requires the establishment of a sewage system in every settlement with a population over 2 000. The establishment of the sewage system in Hungary has accelerated over recent years. While in 2004, 31.5% of the households with a public water supply system were not connected to the sewage system, this ratio decreased to 14.5% in 2016 (KSH, 2017).

Because of the expensive transportation costs, in many cases local inhabitants have chosen to build sewage tanks using permeable walls of concrete or brick (uninsulated sewage tanks) so that the wastewater would be able to seep into the soil, resulting in the contamination of groundwater as shown in Fig. 1 (Mester et al., 2016).

These sewage tanks pose a significantly higher risk of environmental damage than do septic tanks, since raw untreated sewage water flows into the groundwater from them, and consequently, in settlements without a sewage system these tanks are the most important sources of contamination of the groundwater.

After reviewing the literature, we concluded that even though papers have been published on the environmental problems caused by sewage effluent in settlements (Wolf et al., 2004; McArthur et al., 2012; Edo et al., 2014, Augustsson et al., 2016; Szabó et al., 2016), few studies have been carried out focusing on the direct environmental impact of a specific sewage tank (Mester et al., 2016, 2017). We are not aware of studies showing what types of changes occur in the quality of groundwater after eliminating a sewage tank. Since the elimination of these sources of contamination is a major goal throughout the world our studies could help to clarify the recovery processes of the groundwater.

In this study, investigations in the immediate area of a sewage tank were carried out, and the effects of nitrogen effluent from uninsulated sewage tanks on the groundwater were demonstrated. The sewage system in the studied settlement was constructed in October 2014; therefore, the sewage tank we examined was no longer in use. Given that we started the investigations before the sewage system was constructed, we were able to carry out comparative examinations involving an active and an out-of-use sewage tank. The nitrogen flowing out from sewage tanks migrates into the groundwater primarily in the form of NH_4^+ , NO_2^- , NO_3^- (Heatwole and McCray, 2007) and the ratio of the above compounds is an important index combination; that is why we chose these nitrogen compounds to evaluate the groundwater quality, and the recovery process after the elimination of sewage water emissions.

Our general hypothesis was that the groundwater was heavily contaminated in the vicinity of the uninsulated sewage tanks, and the degree of contamination would significantly decrease 1.5 years after the elimination of wastewater emissions.

In order to test our hypothesis, we performed the following tasks: (i) investigation of the spatial distribution of NH_4^+ , NO_2^- , NO_3^- concentrations in the immediate area of active and out-of-use sewage tanks, (ii) estimation of the volume of the nitrogen-contaminated groundwater using a 3D model, (iii) Identification of the amount of NH_4^+ , NO_2^- , NO_3^- present in the modelled saturated zone, (iv) evaluation of the changes in the ratio of the nitrogen forms investigated.

MATERIALS AND METHODS

Site location and characteristics

The investigated settlement – Báránd – is located in the eastern part of the Great Hungarian Plain, on the Nagy-Sárrét on

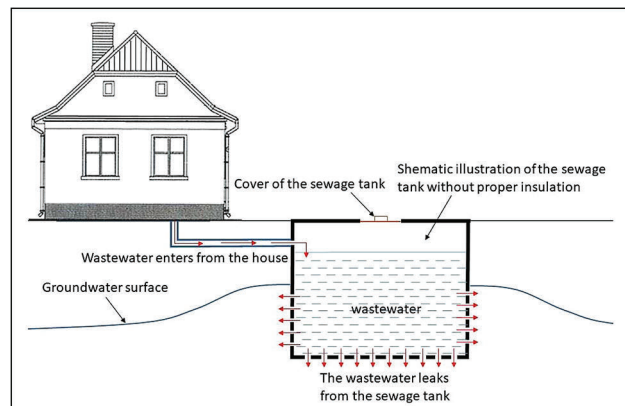


Figure 1. Schematic model of permeable constructed sewage tanks

the western part of the alluvial deposit of the Sebes-Körös River (Fig. 2), and has a population of 2 631 (KSH, 2015). The altitude of the Nagy-Sárrét is typically 85–89 m and the region is classified as a flat plain (relative relief 0–3 m·km⁻²). The groundwater level is close to the surface, at a depth of 1–2 m; consequently, all the soil types have been formed under the influence of water (Michéli et al., 2006). In the study area the most frequent soil types are Solonetz, Vertisol, Kastenzem, and Chernozem, and in the built-up area – as a result of anthropogenic effects – Technosol (Novák and Tóth, 2016).

Field sampling and laboratory analysis

In order to analyse the effect of sewage tanks located in the settlement on the environment, we selected a sewage tank located in the centre of the settlement (Fig. 3). Monitoring wells were established with a depth of 3 m in the immediate vicinity of the sewage tank (Fig. 3).

In this study an analysis of the water samples collected during the spring of 2013 and the spring of 2016 was performed. The well casing was made of PVC pipes with a diameter of 50 mm. For the water sampling a peristaltic pump was used. Before sampling the volume of water originally contained in the wells was extracted 3 times according to the

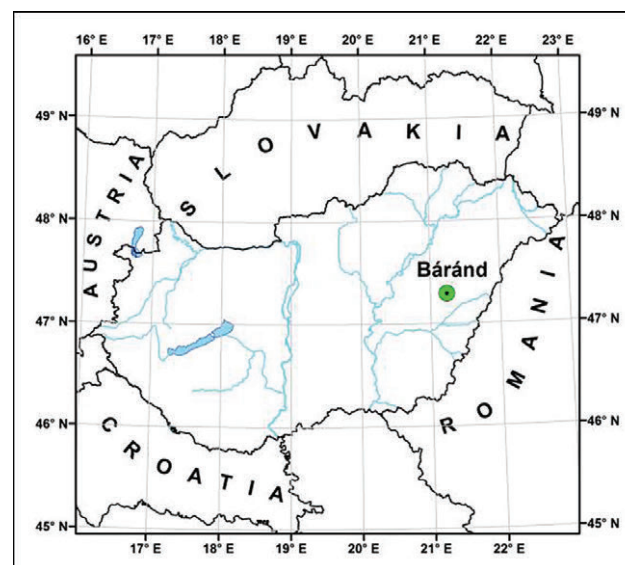


Figure 2. Location of the investigated settlement in Hungary

Hungarian Standard (MSZ 21464:1998). The groundwater levels in the monitoring wells were measured at the time of sampling. The hermetically closed samples were delivered to the Laboratory of Geography at the University of Debrecen, stored in a refrigerator and analysed within 24 h. Before the analysis samples were filtered using 1288 filter paper. The ammonium nitrite and nitrate concentrations were determined using UV-VIS spectrophotometer (Literathy, 1973). The chemical oxygen demand (COD_{Mn}) was determined in accordance with Hungarian Standard (MSZ 12750/21-71). Soil samples were collected from one of the monitoring wells at 20-cm intervals down to a depth of 3 m, and their texture was determined by the Köhn-pipette method (Müller et al., 2009).

In order to identify the elevation of the groundwater levels, a digital relief map was created, using two Trimble S9 dual-frequency, high precision geodesic GPS devices (accuracy 2 cm). The interpolation of the surface was completed with a free triangular mesh.

Model calculation

The spatial geological models were developed with Surfer 11 and RockWorks 14 modelling software. In order to demonstrate the NH_4^+ , and NO_3^- concentrations, isometric maps were created using the kriging geostatistical method in Surfer. Using the RockWorks software the 3D model of the area was compiled, during which kriging interpolation was used. In order to construct the 3D model of the distribution of the contaminant concentration (M), RockWorks applies the following formula:

$$M = V_{voxel} \sum_{i=1}^n n_{oi} c_i$$

where: M is the volume of the water body which exceeds the specified concentration, V_{voxel} is the volume of the voxel, n_{oi} is the effective porosity, and c_i is the concentration value measured in the monitoring wells.

The 3D models describe the distribution of the investigated contaminants down to 3 m below the surface level in the saturated zone, since data regarding the soil texture was available only to this depth.

Based on the models created with RockWorks, the volume of the water bodies contaminated with NH_4^+ , NO_2^- and NO_3^- in terms of the given concentrations was identified. Since the soil texture was loam in the investigated area, it was calculated based on a pore space of 45% (Stefanovits, 1981). Then, based on the results, the amount of inorganic nitrogen compounds which can be found in the given water body was identified in grams.

RESULTS

Soil texture

The soil texture plays a key role in the determination of groundwater flow. Figure 4 shows the granulometric composition of the soil samples from the BA5 monitoring well. Coarse sand (particle size of 0.2–2 mm) was not identified in any of the soil samples with depths of 0–300 cm.

It can be seen that the soil texture was loam in all of the investigated soil depths. The Zamarin filtration coefficient values identified on the basis of these results are very low, varying between $1.23 \times 10^{-7} \text{ m}\cdot\text{s}^{-1}$ and $5.20 \times 10^{-7} \text{ m}\cdot\text{s}^{-1}$ in the investigated layers (Zamarin, 1928).

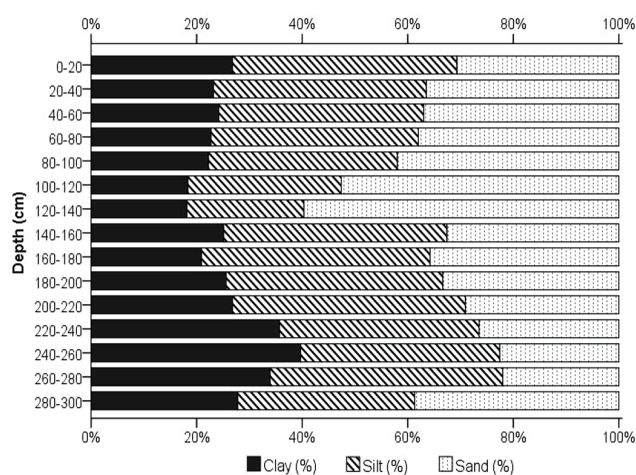


Figure 4. Distribution of the granulometric composition of soil samples from the BA5 monitoring well.

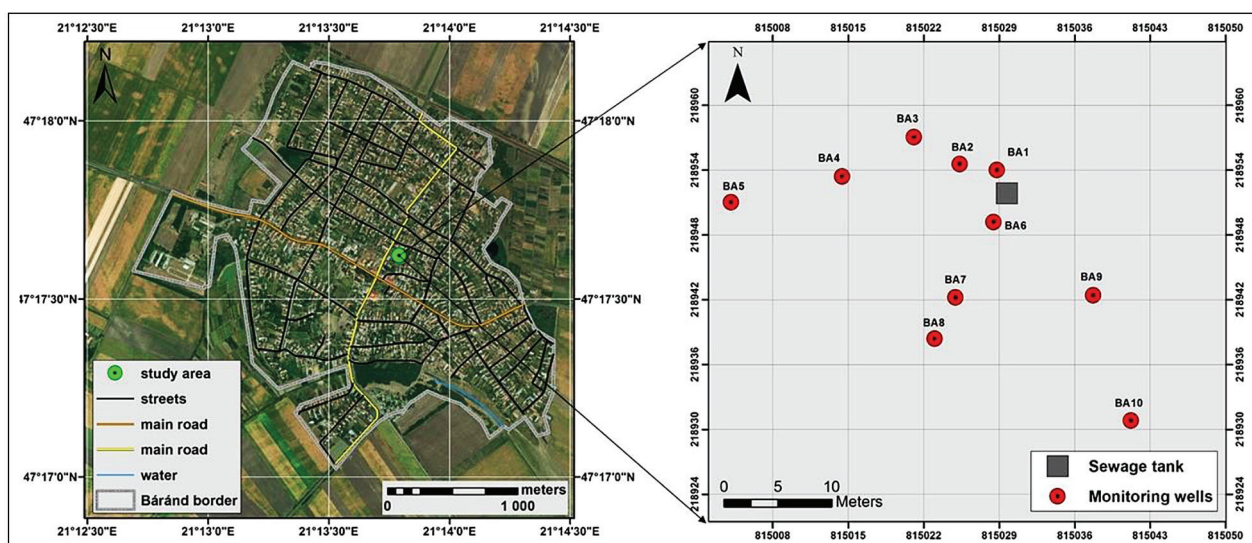


Figure 3. Locations of the monitoring wells and sewage tank in the study area

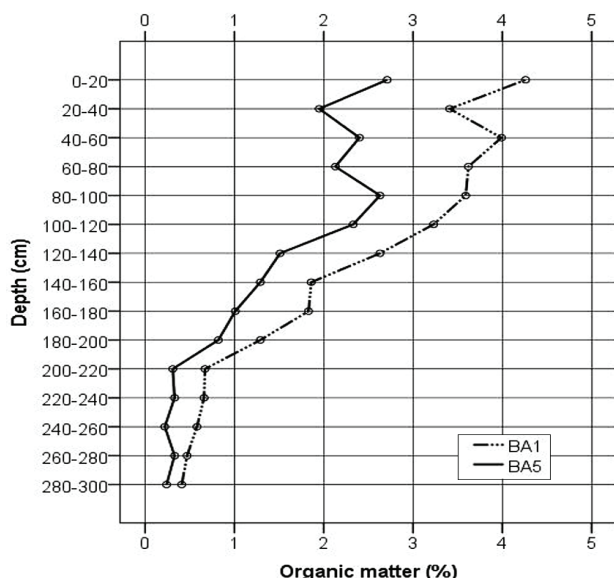


Figure 5. OM content of the soil samples from the BA1 and the BA5 monitoring wells

Organic matter content of the soil

In 2013 the organic matter content of the soil was studied in the monitoring wells in the sections which were closest to (BA1) to and furthest from (BA5) the sewage tank. An average of 40% higher OM content was measured in the section closest to the sewage tank. This significant difference between the sections within a 25 m distance is related to the accumulation of organic matter leaking from the sewage tank, since under natural conditions there was probably no significant difference between the two sections.

Organic matter content of groundwater

The highest COD values of the groundwater were measured in the vicinity of the sewage tank, similarly to the OM content of the soil (Fig. 6).

It can be stated that the spatial distribution of COD values 1.5 years after the closure of the sewage tank showed the same

pattern. These values are significantly reduced within a few meters from the sewage tank, but concentrations exceeded the contamination limit (6/2009 Government Decree) ($4.5 \text{ mg}\cdot\text{L}^{-1}$) in every monitoring well, indicating anthropogenic effects. Based on the measurements, it can be concluded that inorganic nitrogen forms are still being supplied.

Spatial distribution and amount of inorganic nitrogen forms (NH_4^+ , NO_2^- , NO_3^-)

Sewage water emission in the study area

In Hungary an average of $100 \text{ L}\cdot\text{capita}^{-1}\cdot\text{day}^{-1}$ of wastewater is generated, with a significant distribution (Takács 2013). Based on the water consumption data in the 4-person household studied, $116 \text{ L}\cdot\text{capita}^{-1}\cdot\text{day}^{-1}$ of wastewater is generated on average, or a total of $464 \text{ L}\cdot\text{day}^{-1}$ of wastewater per household (Mester et al., 2016). Based on the above, $170 \text{ m}^3\cdot\text{yr}^{-1}$ of wastewater flowed into the sewage tank from the household investigated. According to our records of wastewater transportation by sewage suction trucks, in 2013, 90 m^3 of wastewater was transported from the sewage tank in total, indicating that the amount of wastewater effluent from the sewage tank was 80 m^3 in 2013, making up 47% of the generated wastewater (Mester et al., 2017).

Spatial distribution and amount of NH_4^+

The most important contaminants of domestic wastewater originate from the decomposition of organic matter. During the first step of the process NH_4^+ is created, producing a concentration above $90 \text{ mg}\cdot\text{L}^{-1}$ (Takács, 2013), which is characteristic of the raw wastewater found in used sewage tanks ($124 \text{ mg}\cdot\text{L}^{-1}$) (Fig. 7). Due to the high organic carbon content, conditions are anaerobic and nitrification of NH_4^+ is limited. Moving away from the sewage tank, as the organic carbon content decreases nitrification progressively occurs. Therefore, the concentration of ammonium rapidly decreases, such that at a distance of 15–20 m from the well it decreases to a concentration close to the contamination limit value of $0.5 \text{ mg}\cdot\text{L}^{-1}$ (6/2009 Government Decree) (Fig. 7).

Even 1.5 years after closure, a concentration above $90 \text{ mg}\cdot\text{L}^{-1}$ was not measured in the monitoring wells, in the BA1

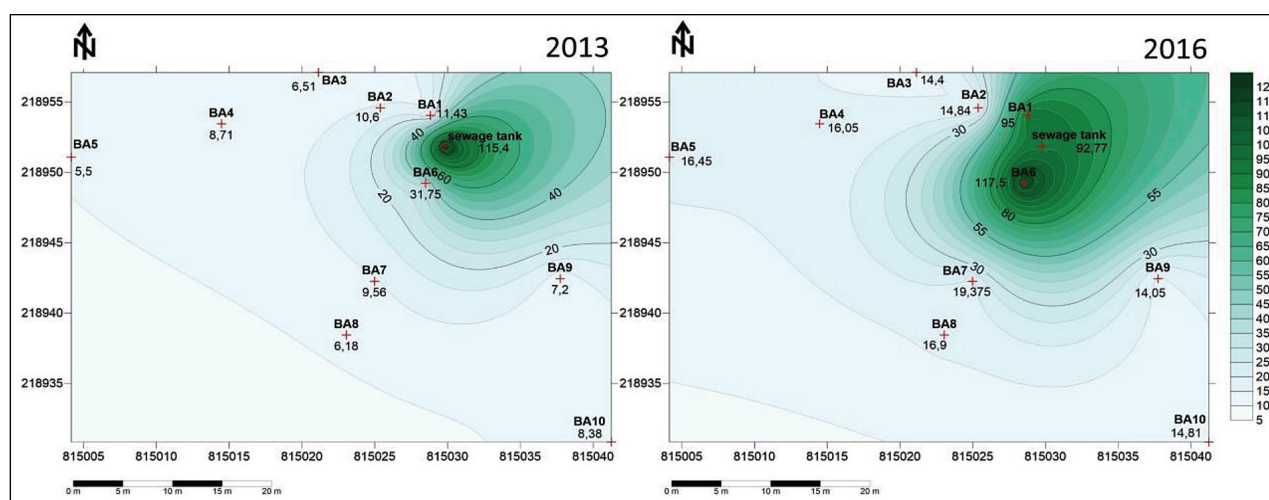


Figure 6. Spatial distribution of COD values in the groundwater in 2013 and 2016

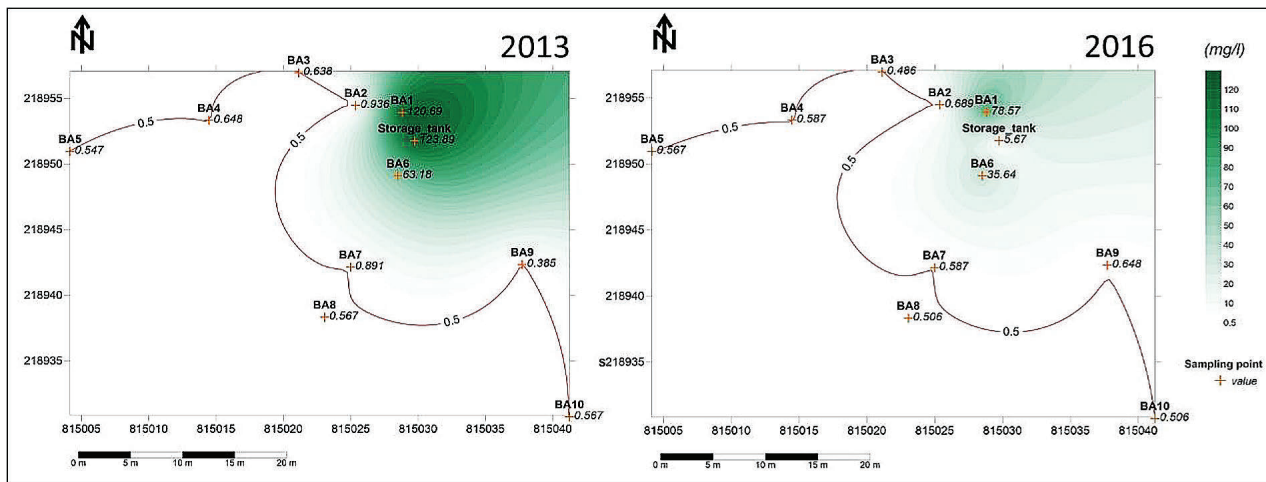


Figure 7. NH_4^+ concentrations in the groundwater in (a) 2013 and (b) 2016

well the concentration of NH_4^+ was still $78.57 \text{ mg}\cdot\text{L}^{-1}$, and it did not decrease to below the contamination limit in the case of more distant monitoring wells, which can be explained by two reasons (Fig. 7b): The very high concentrations detected in the vicinity of the sewage tank were caused by the considerable amount of organic material accumulated while the sewage tank was in use, the decomposition of which is still in progress 1.5 years after the sewage tank went out of use. Another possible cause was that in the modelled area the clay content of the soil varies between 18 and 40% (Fig. 4). According to Nieder et al. (2011), medium-textured soils have a fixed $\text{NH}_4^+\text{-N}$ content of $60\text{--}270 \text{ mg}\cdot\text{kg}^{-1}$. If the amount of NH_4^+ in the groundwater decreases by nitrification, a part of the fixed NH_4^+ can be released, which can be an additional source of NH_4^+ (Stefanovits et al., 1999).

Since the 3D kriging interpolation method provides us with a more accurate picture of the spatial distribution of

contaminants, a static hydrogeological model was created using the Rockworks software. In 2013 the groundwater dome caused by the sewage water flowing out from the tank could be observed (Fig. 8-I,II,III); by 2016, 1.5 years after the last sewage water emissions, the dome had completely disappeared (Fig. 8-IV,V,VI). In 2013, in a layer modelled with the software down to a depth of 3 m, a concentration above $90 \text{ mg}\cdot\text{L}^{-1}$ in 51 m^3 from a water body of 734.4 m^3 of the saturated zone was measured (Fig. 8-III), while the modelled saturated zone can be considered almost entirely contaminated in 2013 and 2016. (Fig. 8-I,IV).

Based on the water volumes established in the modelled zone, the amount of NH_4^+ in the study area in 2013 and 2016 was estimated (Table 1). Even though the groundwater level had increased slightly by 2016, causing the volume of the saturated zone to increase, the amount of NH_4^+ had decreased to almost a third of its former value.

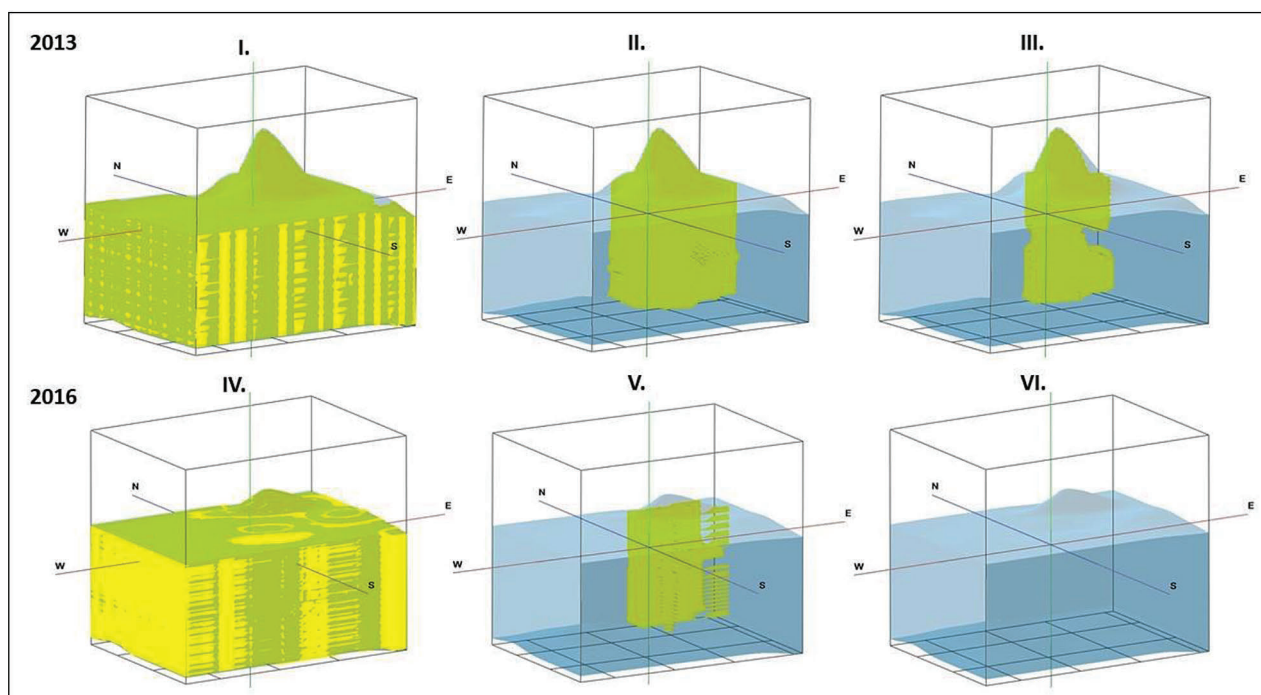


Figure 8. Spatial distribution of NH_4^+ concentrations down to a depth of 3 m from the surface. The extension of the water body with an NH_4^+ concentration: I, IV: higher than $0.5 \text{ mg}\cdot\text{L}^{-1}$, II, V: higher than $30 \text{ mg}\cdot\text{L}^{-1}$, III, VI: higher than $90 \text{ mg}\cdot\text{L}^{-1}$.

Table 1. Distribution and amount of NH_4^+ in 2013 and 2016 in the modelled zone

Concentration (mg·L ⁻¹)	2013		2016	
	Volume of water body in the saturated zone (m ³)	The amount of NH_4^+ which can be found in the groundwater (g)	Volume of water body in the saturated zone (m ³)	The amount of NH_4^+ which can be found in the groundwater (g)
90 <	51.0	5 380	0	0
60–89.9	48.6	3 645	8.0	600
30–59.9	48.4	2 160	33.7	1 516.5
5–29.9	168.0	2 940	167.6	2 933
2–4.9	76.3	267.1	77.9	272.7
0.5–1.9	339.5	424.4	493.7	617.1
0.49 >	2.6	1.3	4.2	2.1
In total	734.4	14 817.8	785.1	5 941.4

Spatial distribution and amount of NO_2^-

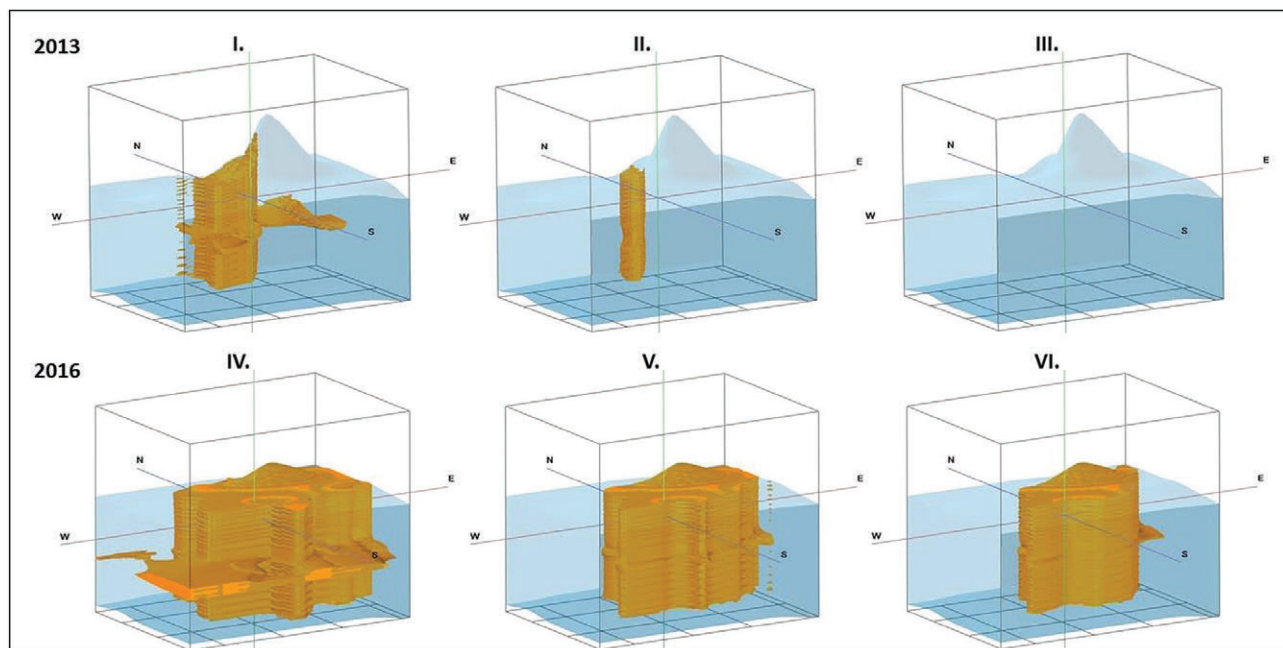
The NH_4^+ content originating from the decomposition of the organic materials is oxidized into nitrite through the activity of Nitrosomonas bacteria under aerobic conditions. In 2013, in the majority of the modelled saturated zone the amount of NO_2^- was under the contamination limit (0.1 mg·L⁻¹) (6/2009 Government Decree) in 677 m³ of a water body of 734 m³ (Fig. 9-I), and it was present in a concentration above

0.5 mg·L⁻¹ only in a water body of 10 m³ (Fig. 9-II).

In 2016, however, a significant increase in the concentration of NO_2^- could be identified, since after ending sewage water emissions the conditions for nitrification improved in the direct vicinity of the sewage tank. While in 2013 a concentration above 1 mg·L⁻¹ was not measured, in 2016 the concentration increased to a level of 1 mg·L⁻¹ in 149 m³ of a water body of 785 m³ (Fig. 9-VI, Table 2). The amount of NO_2^- increased to a value more than 2 times higher than that of the value measured in 2013 (Table 2).

Table 2. Distribution and amount of NO_2^- in 2013 and 2016 in the modelled zone

Concentration (mg·L ⁻¹)	2013		2016	
	Volume of water body in the saturated zone (m ³)	The amount of NO_2^- which can be found in the groundwater (g)	Volume of water body in the saturated zone (m ³)	The amount of NO_2^- which can be found in the groundwater (g)
5 <	0	0	8.6	50.5
1–4.9	0	0	149.1	447.3
0.5–0.99	9.9	7,425	66.3	49.7
0.1–0.49	56.6	16.9	140.1	42.03
0.02–0.09	649.5	357.2	420.3	231.2
0.019 >	18.4	17.5	0.7	0.7
In total	734.4	399.0	785.1	821.4

**Figure 9.** Spatial distribution of NO_2^- concentrations down to a depth of 3 m from the surface. The extension of the water body with an NO_2^- concentration: I, IV: higher than 0.1 mg·L⁻¹, II, V: higher than 0.5 mg·L⁻¹, III, VI: higher than 1 mg·L⁻¹.

Spatial distribution and amount of NO_3^-

Under aerobic conditions the amount of NO_2^- is not maintained in the water for a long period of time, because it is oxidized into nitrate through the activity of the *Nitrobacter* bacteria. In the case of NO_3^- significant changes can be identified as well. While in the vicinity of the sewage tank a concentration of 1–3 $\text{mg}\cdot\text{L}^{-1}$ was measured in 2013 (Fig. 10-A), the concentrations increased drastically after the sewage tank went out of use. The concentration value increased from the 2.02 $\text{mg}\cdot\text{L}^{-1}$ measured in 2013, to 2 341 $\text{mg}\cdot\text{L}^{-1}$ in the BA6 monitoring well, which is almost 50 times higher than the contamination limit of 50 $\text{mg}\cdot\text{L}^{-1}$ (6/2009 Government Decree), equivalent to the maximum permissible limit of nitrate concentration in drinking water required by the WHO (WHO, 2011) (Fig. 10-B). Significant changes can also be detected regarding the spatial distribution of concentrations. While in 2013 the NO_3^- values showed a gradual increase as one moved away from the well, in 2016 these values had decreased in the same direction.

While in 2013 no water bodies with a concentration above 500 $\text{mg}\cdot\text{L}^{-1}$ were found in the modelled saturated zone, in 2016, from a water body of 785 m^3 , 94 m^3 and 37 m^3 showed concentrations above 500 $\text{mg}\cdot\text{L}^{-1}$ and 1 000 $\text{mg}\cdot\text{L}^{-1}$, respectively (Fig. 11, Table 3).

The extension of the water body with an NO_3^- concentration: I, III: higher than 50 $\text{mg}\cdot\text{L}^{-1}$, II, V: higher than

150 $\text{mg}\cdot\text{L}^{-1}$, III, VI: higher than 500 $\text{mg}\cdot\text{L}^{-1}$.

Based on the estimation performed, the amount NO_3^- present in the modelled zone had increased 7.5 times from 19 049 g in 2013, to reach a value of 142 909 g by 2016 (Table 3).

Similarly to NO_2^- , in the case of NO_3^- the conditions for nitrification significantly improved as a consequence of the ending of fresh sewage water emissions, and this led to extremely high concentrations in the direct proximity of the sewage tank.

Ratio and mass balance of inorganic nitrogen forms

The ratio of $\text{NH}_4^+\text{-N}$, $\text{NO}_2^-\text{-N}$ and $\text{NO}_3^-\text{-N}$ is a very important index combination from the perspective of the cleaning process of waters. After the identification of inorganic nitrogen forms the values of $\text{NH}_4^+\text{-N}$, $\text{NO}_2^-\text{-N}$ and $\text{NO}_3^-\text{-N}$ were calculated in grams. It was concluded that in the modelled zone 1.5 years after the closure of the sewage tank the amount of nitrogen present in inorganic forms had increased from 19 533 g to 46 940 g, which is more than twice as high as the baseline value in 2013.

With regard to the nitrogen forms investigated it was not expected that inside the modelled zone the nitrogen concentration would increase after the closure of the sewage tank. This increase can be explained by the fact that the sewage tank had been in use for 27 years, and during this time a large amount of organic material accumulated in the vicinity of the sewage tank, the decomposition of which provided the source

Table 3. Distribution and amount of NO_3^- in 2013 and 2016 in the modelled zone

Concentration ($\text{mg}\cdot\text{L}^{-1}$)	2013		2016	
	Volume of water body in the saturated zone (m^3)	The amount of NO_3^- which can be found in the groundwater (g)	Volume of water body in the saturated zone (m^3)	The amount of NO_3^- which can be found in the groundwater (g)
2 000 <	0	0	0.4	868
1 000–1 999.9	0	0	37.4	56 100
500–999.9	0	0	55.4	44 550
300–499.9	0	0	46.0	18 400
150–299.9	36.3	6 352	61.6	10 780
50–149.9	357.4	3 574	200.0	2 000
25–49.9	194.6	7 297	216.3	8 111
24.9 >	146.1	1 826	168.0	2 100
In total	734.4	19 049	785.1	142 909

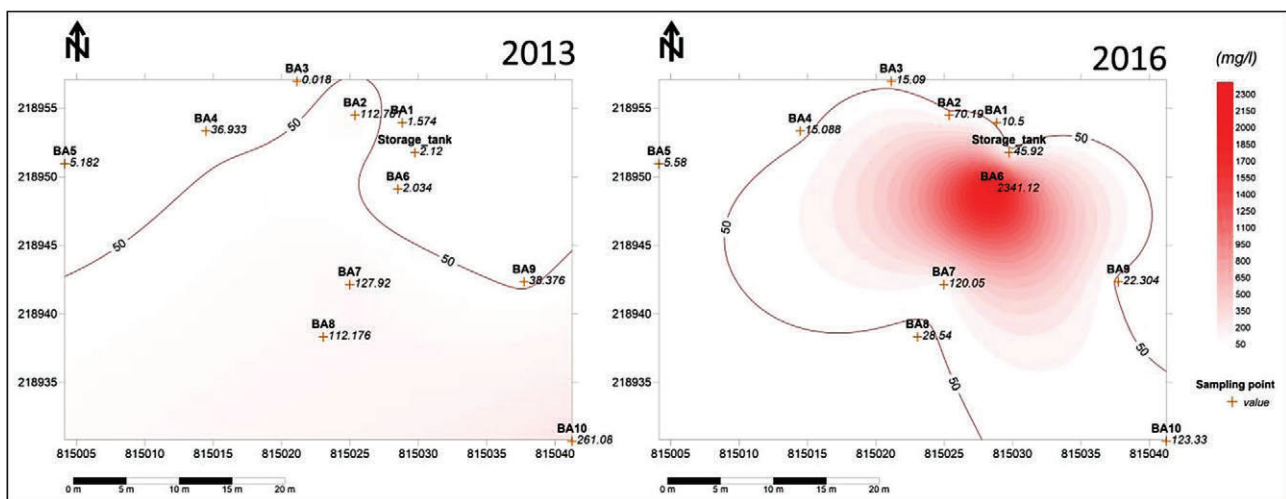


Figure 10. NO_3^- concentrations in the study area in 2013 and 2016

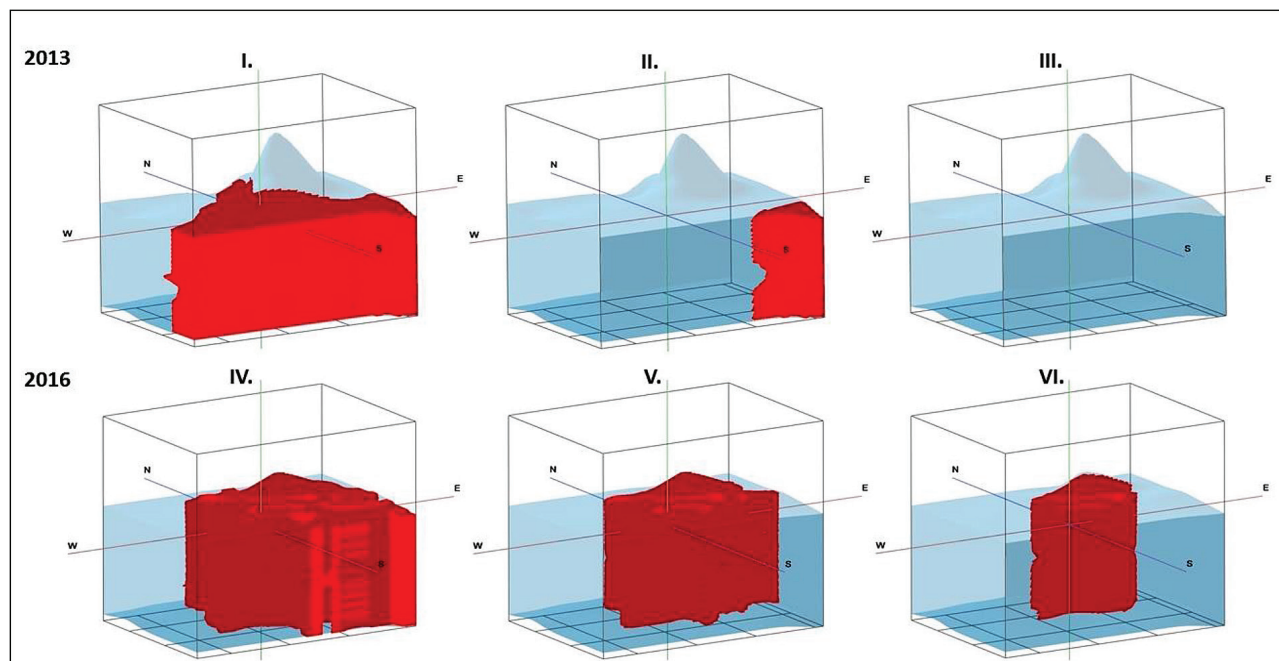


Figure 11. Spatial distribution of NO_3^- concentrations down to a depth of 3 m from the surface.

for the permanent creation of inorganic nitrogen forms, even after the sewage tank went out of use. This is proven by the NH_4^+ concentration of close to $80 \text{ mg}\cdot\text{L}^{-1}$ which was measured in the BA1 well found in the direct vicinity of the sewage tank (Fig. 5), which was measured 1.5 years after the sewage tank was closed.

During the period when the sewage tank was in use, the organic nitrogen forms were predominant in the direct vicinity of the sewage tank. This suggests that the increase in inorganic nitrogen forms after the ending of the sewage water emissions can only be explained by the alteration of the previously accumulated organic forms. Thus the amount of organic nitrogen constantly decreases, since the supply of organic material has stopped.

Significant changes also took place regarding the ratio between inorganic nitrogen forms. The ratio of NH_4^+ decreased from 70% to under 10%. Among the more oxidized forms the ratio of NO_2^- did not show a significant change, which can be explained by the fact that in the presence of oxygen NO_2^- oxidizes into nitrate at a faster rate. While in 2013, 29% of the inorganic nitrogen forms were nitrate, by 2016 this value had increased to close to 90% (Table 4). Based on the above it can be concluded that the degree of oxidation of nitrogen had significantly increased by 2016.

CONCLUSIONS

Our hypothesis that in the vicinity of uninsulated sewage tanks the groundwater is heavily contaminated was clearly verified. For each parameter investigated in the immediate vicinity of the sewage tank, concentrations characteristic of raw wastewater were measured (COD, NH_4^+ , NO_2^- , NO_3^-).

Our hypothesis, however, that 1.5 years after the ending of sewage water emissions the extent of contamination would significantly decrease, has not been confirmed. A decrease in the concentration of NH_4^+ could be detected, while the amount of NO_3^- increased, which can be partly explained by the fact

Table 4. Amount of N in grams in 2013 and 2016 in the modelled zone

	2013		2016	
Amount of $\text{NH}_4\text{-N}$ (g)	13 766	70.48%	4 634	9.87%
Amount of $\text{NO}_2\text{-N}$ (g)	176	0.9%	362	0.77%
Amount of $\text{NO}_3\text{-N}$ (g)	5 591	28.62%	41 944	89.36%
In total	19 533	100%	46 940	100%

that after the ending of sewage water emissions the conditions for nitrification were improved. The OM content of the soil near the sewage tank was significantly higher than at the border of the modelled area, which is the result of the accumulation of organic matter leaking from the sewage tank over decades. A similar spatial distribution in the COD values of the groundwater was detected, which remained 1.5 years after the closure of the sewage tank.

The rate of the cleaning process primarily depends on the amount of organic material accumulated in the vicinity of the source of contamination, therefore the amount of inorganic nitrogen forms can only be expected to decrease after the decomposition of organic materials. Based on our results it can be concluded that the cleaning process of the groundwater has already started, but could still take years to finish.

Even though our investigations focused on a specific sewage tank, the processes occurring in the study area can be generalized to any equipment from which sewage water can be emitted into the groundwater.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The research was financed by the Higher Education Institutional Excellence Programme (20428-3/2018/

FEKUTSTRAT) of the Ministry of Human Capacities in Hungary, within the framework of the 4.thematic programme of the University of Debrecen.

REFERENCES

- ABDALLA F and KHALIL R (2018) Potential effects of groundwater and surface water contamination in an urban area, Qus City, Upper Egypt. *J. Afr. Earth Sci.* **141** 164–178. <https://doi.org/10.1016/j.jafrearsci.2018.02.016>
- ADAMS JB, PRETORIUS L and SNOW GC (2019) Deterioration in the water quality of an urbanised estuary with recommendations for improvement. *Water SA* **45** (1) 86–96. <https://doi.org/10.4314/wsa.v45i1.10>
- ADHANOM G, HUGHES J and ODINDO A (2018) The effect of anaerobic baffled reactor effluent on nitrogen and phosphorus leaching from four soils in a laboratory column experiment. *Water SA* **44** (1) 1–12. <https://doi.org/10.4314/wsa.v44i1.01>
- ADUMITROAEI M V, GAVRILAOAIEI T, SANDU A V and IANCU GO (2016) Distribution of mineral nitrogen compounds in groundwater in Vaslui County (Romania). *Cancer* **12** 13.
- ATTOUI B, TOUMI N, MESSAOUDI S and BENRABAH S (2016) Degradation of water quality: the case of plain west of Annaba (northeast of Algeria). *J. Water Land Dev.* **31** 3–10. <https://doi.org/10.1515/jwld-2016-0031>
- AUGUSTSSON A, UDDH SÖDERBERG T, JARSJÖ J, ÅSTRÖM M, OLOFSSON B, BALFORS B and DESTOUNI G (2016) The risk of overestimating the risk-metal leaching to groundwater near contaminated glass waste deposits and exposure via drinking water. *Sci. Total Environ.* **566–567** 1420–1431. <https://doi.org/10.1016/j.scitotenv.2016.06.003>
- BANKS D, KARNACHUK OV, PARNACHEV VP, HOLDEN W and FRENGSTAD B (2002) Groundwater contamination from rural pit latrines: examples from Siberia and Kosova. *Water Environ. J.* **16** 147–152. <https://doi.org/10.1111/j.1747-6593.2002.tb00386.x>
- BENRABAH S, ATTOUI B and HANNOUCHE M (2016) Characterization of groundwater quality destined for drinking water supply of Khenchela City (eastern Algeria). *J. Water Land Dev.* **30** 13–20. <https://doi.org/10.1515/jwld-2016-0016>
- DEVIC G, DJORDJEVIC D and SAKAN S (2014) Natural and anthropogenic factors affecting the groundwater quality in Serbia. *Sci. Total Environ.* **468–469** 933–942. <https://doi.org/10.1016/j.scitotenv.2013.09.011>
- DRAKE VM, and BAUDER JW (2005) Ground water nitrate-nitrogen trends in relation to urban development, Helena, Montana, 1971–2003. *Groundwater Monit. Remediation* **25** (2) 118–130. <https://doi.org/10.1111/j.1745-6592.2005.0017.x>
- EDO FA, EJIJOGU CC, UZOIJE AP, NWACHUKWU MA and OKOLI CG (2014) Impact of open sewage dumpsites on groundwater quality in Igwuruta, Rivers State, Nigeria. *J. Glob. Biosci.* **3** (6) 919–930.
- FANTONG WY, FOUPE AT, DJOMOU SL, BANSEKA HS, ANAZAWA K, ADELANA SMA and NKENG GE (2013) Temporal pollution by nitrate (NO₃), and discharge of springs in shallow crystalline aquifers: Case of Akok Ndoue catchment, Yaounde (Cameroon). *Afr. J. Environ. Sci. Technol.* **7** (5) 167–174.
- GOODY DC, MACDONALD DMJ, LAPWORTH DJ, BENNETT SA and GRIFFITHS KJ (2014) Nitrogen sources, transport and processing in peri-urban floodplains. *Sci. Total Environ.* **494** 28–38. <https://doi.org/10.1016/j.scitotenv.2014.06.123>
- HANCHAR DW (1991) Effects of septic-tank effluent on ground-water quality in Northern Williamson County and Southern Davidson County, Tennessee. U.S. Geological Survey Water-Resources Investigations Report 91-4011. USGS, Nashville, Tennessee. <https://doi.org/10.3133/wri914011>
- HEATWOLE KK and MCCRAY JE (2007) Modeling potential vadose-zone transport of nitrogen from onsite wastewater systems at the development scale. *Contam. Hydrol.* **91** 184–201. <https://doi.org/10.3133/wri914011>
- HUNGARIAN STANDARDS INSTITUTION (1998) MSZ-21464:1998. Sampling of groundwaters. Hungarian Standards Institution, Budapest.
- KRINGEL R, RECHENBURG A, KUITCHA, D, FOUÉPÉ A, BELLENBERG S, KENGNE IM and FOMO MA (2016) Mass balance of nitrogen and potassium in urban groundwater in Central Africa, Yaounde/Cameroon. *Sci. Total Environ.* **547** 382–395. <https://doi.org/10.1016/j.scitotenv.2015.12.090>
- KSH (2015) Hungarian Central Statistical Office. URL: https://www.ksh.hu/docs/hun/hnk/hnk_2015.pdf (Accessed 24 March 2017).
- KSH (2017) Hungarian Central Statistical Office. URL: http://www.ksh.hu/docs/hun/xstadat/xstadat_eves/i_zrk006.html (Accessed 21 March 2018).
- LITERÁTHY P (1973) United Water Examination Methods I. Chemical methods (in Hungarian). Vol. 1, Department IV of Water Quality and Water Technology of the Water Management Scientific Research Institute, Budapest, Hungary. 233 pp.
- MCARTHUR JM, SIKDAR PK, HOQUE MA and GHOSAL U (2012) Waste-water impacts on groundwater: Cl/Br ratios and implications for arsenic pollution of groundwater in the Bengal Basin and Red River Basin, Vietnam. *Sci. Total Environ.* **437** 390–402. <https://doi.org/10.1016/j.scitotenv.2012.07.068>
- MCQUILLAN D (2004) Ground-water quality impacts from on-site septic systems. *Proceedings, National Onsite Wastewater Recycling Association, 13th Annual Conference*, Albuquerque, NM, 7–10 November 2004. 13.
- MESTER T, SZABÓ GY, BALLA D, KARANCSI G, SZABÓ G and TÓTH CS (2016) Egy nem zárt rendszerű szennyvízakra talajvízszintre gyakorolt hatásának vizsgálata (in Hungarian) [The investigation of the effects of an open sewage tank on the groundwater level]. Theory meets practice in GIS VII. 311–317.
- MESTER T, SZABÓ GY, BESSENYEI É, KARANCSI G, BARKÓCZI N and BALLA D (2017) The effects of uninsulated sewage tanks on groundwater. A case study in an eastern Hungarian settlement. *J. Water Land Dev.* **33** 123–129. <https://doi.org/10.1515/jwld-2017-0027>
- MESTER T and SZABÓ G (2013). Nitrate contamination in the groundwater wells of an Eastern Hungarian settlement. In: Ozdemir CC, Şahinkaya S, Kalıpcı E, and Oden MK (eds) *International Conference on Environmental Science and Technology*, Cappadocia. Conference CD. Digital Proceeding of THE ICOEST²⁰¹³. 1–16.
- MICHÉLI E, FUCHS M, HEGYMEGI P and STEFANOVITS P (2006) Classification of the major soils of Hungary and their correlation with the World Reference Base for Soil Resources (WRB). *Agrochem. Soil Sci.* **55** (1) 19–28. <https://doi.org/10.1556/agrochem.55.2006.1.3>
- MOODLEY KG, SOBANTU P, GERICKE G, CHETTY DK and PIENAAR DH (2017) Comparison of UV and ELS detectors in HSPEC analysis of natural organic matter in dam water. *Water SA* **43** (3) 520–528. <https://doi.org/10.4314/wsa.v43i3.17>
- MÜLLER HW, DOHRMANN R, KLOSA D, REHDER S and WOLF ECKELMANN W (2009) Comparison of two procedures for particle-size analysis: Köhn pipette and X-ray granulometry. *J. Plant Nutr. Soil Sci.* **172** 172–179. <https://doi.org/10.1002/jpln.200800065>
- NIEDER R, BENBI DK and SCHERER HW (2011) Fixation and defixation of ammonium in soils: a review. *Biol. Fertil. Soils* **47** (1) 1–14. <https://doi.org/10.1007/s00374-010-0506-4>
- NLEND B, CELLE-JEANTON H, HUNEAU F, KETCHEMEN-TANDIA B, FANTONG WY, BOUM-NKOT SN and ETAME J (2018) The impact of urban development on aquifers in large coastal cities of West Africa: Present status and future challenges. *Land Use Polic.* **75** 352–363. <https://doi.org/10.1016/j.landusepol.2018.03.007>
- NOVÁK TJ and TÓTH CSA (2016) Development of erosional microforms and soils on semi-natural and anthropogenic influenced solonchic grasslands. *Geomorphology* **254** 121–129. <https://doi.org/10.1016/j.geomorph.2015.11.018>
- REAY WG (2004) Septic tank impacts on groundwater quality and nearshore sediment nutrient flux. *Ground Water* **42** (7) 1079–1089. <https://doi.org/10.1111/j.1745-6584.2004.tb02645.x>
- ROBERTSON WD, MOORE TA, SPOELSTRA J, Li L, ELGOOD RJ, CLARK ID, SCHIFF SL, ARAVENA R and NEUFELD

<https://doi.org/10.17159/wsa/2019.v45.i3.6731>

Available at <https://www.watersa.net>

ISSN 1816-7950 (Online) = Water SA Vol. 45 No. 3 July 2019

Published under a Creative Commons Attribution 4.0 International Licence (CC BY 4.0)

- JD (2012). Natural attenuation of septic system nitrogen by Anammox. *Ground Water* **50** 541–553. <https://doi.org/10.1111/j.1745-6584.2011.00857.x>
- ROBERTSON WD, VAN STEMPVOORT DR, ROY JW, BROWN SJ, SPOELSTRA J, SCHIFF SL, RUDOLPH DR, DANIELESCU S and GRAHAM G (2016) Use of an artificial sweetener to identify sources of groundwater nitrate contamination. *Ground Water* **54** 579–587. <https://doi.org/10.1111/gwat.12399>
- SIMMONS RC, GOLD AJ and GROFFMAN PM (1992) Nitrate dynamics in riparian forests: groundwater studies. *J. Environ. Qual.* **21** 659–665. <https://doi.org/10.2134/jeq1992.00472425002100040021x>
- SMOROŃ S (2016) Quality of shallow groundwater and manure effluents in a livestock farm. *J. Water Land Dev.* **29** 59–66. <https://doi.org/10.1515/jwld-2016-0012>
- STEFANOVITS P, FILEP GY and FÜLEKY GY (1999) Soil sciences (in Hungarian) Mezőgazda Kiadó, Budapest.
- SZABÓ GY, BESSÉNYEI É, HAJNAL A, CSIGE I, SZABÓ G, TÓTH CS, POSTA J and MESTER T (2016) The use of sodium to calibrate the transport modeling of water pollution in sandy formations around an uninsulated sewage disposal site. *Water Air Soil Pollut.* **227** (2) 1–13. <https://doi.org/10.1007/s11270-015-2742-6>
- TAKÁCS J (2013) Possibilities for reducing the nutrient content of domestic wastewaters. *hulladékOnline.* **4** (1).
- The joint decree nr. 6/2009 (IV. 14.) of KvVM-EüM-FVM [the Hungarian Ministries of Environment, Healthcare and Agriculture, respectively] about the limit values and standard procedures to assess the pollution level, in order to protect the geological medium and ground waters against pollution.
- WHO (2011) *Guidelines for Drinking-water Quality* (4th edn). World Health Organization, Geneva. 541 pp.
- WOLF L, HELD I, EISWIRTH M and HÖTZL H (2004) Impact of leaky sewers on groundwater quality, *Acta Hydrochim. Hydrobiol.* **32** (4–5) 361–373. <https://doi.org/10.1002/aheh.200400538>
- ZAMARIN JA (1928) Calculation of ground-water flow (in Russian). Trudey I.V.H. Taskeni.
-

The feasibility of wastewater recycling that includes residue from dissolved air flotation within a drinking water treatment plant: case study of Midvaal Water Company, South Africa

Shalene Janse van Rensburg¹, Sandra Barnard² and Marina Krüger¹

¹Midvaal Water Company, Farm Buffelsfontein 443 IP, District Klerksdorp, 2570, South Africa

²Unit for Environmental Sciences and Management, North-West University, Potchefstroom, 2520, South Africa

ABSTRACT

When purifying water for potable use, wastewater is generated, due to the class of the water treatment plant and the quality of the source water. Midvaal Water Company recycled wastewater that included residue from the dissolved air flotation (DAF), sedimentation and filtration processes in an attempt to save water and reduce costs. The aim of this study was to determine functionality and water quality of such a wastewater recycling system. Samples were collected for analysis, at the sections that contributed to the total wastewater system as well as after various treatment processes. The water quality of these samples was determined, as well as the incidences of water quality failures of the final water, to establish whether the recycle stream that enters the plant together with the source water had any impact on the water quality after the different treatment processes. Data were grouped into periods prior to, during and after recycling to enable comparisons. The water quality of the recycle stream was poorer than that of the source water from the Vaal River with regard to the mean values for total chlorophyll, suspended solids, turbidity and dissolved organic carbon, but the sedimentation process of the wastewater system improved the wastewater quality by drastically reducing total chlorophyll, suspended solids and turbidity. The risk-defined compliance for the final water was excellent ($\geq 95\%$), despite aluminium, turbidity and total chlorophyll failures of the final water quality during the recycling period. Total chlorophyll was identified as the largest risk during wastewater recycling, especially after the filtration process. It is evident from the data that wastewater recycling, which included wastewater from the DAF, into the main inlet stream of the water treatment plant proved to be effective, based on compliance with national legislation, and had no detrimental impact on overall treatment processes or final water quality.

Keywords: dissolved air flotation, sludge balancing dam, total chlorophyll, wastewater recycling, water treatment

INTRODUCTION

Recent droughts and associated conditions in South Africa have increased water users' awareness of current water demands, which are likely to increase in the future. Deteriorating source water quality together with population increases and high water quality standards have led to greater expenses in the production of drinking water. Drinking water treatment plants generate spent water at various stages depending on the source water's quality and the unit operations involved (Bourgeois et al., 2004). According to Reissmann and Uhl (2006), this has led to numerous efforts to implement water reuse systems in treatment plants all over the world. Sludge produced by water treatment plants is mainly intended for disposal in sanitary landfills but the recycling of spent water produced at different stages of the drinking water treatment process can be applied by water treatment plants to reduce water treatment expenses prior to exploring water reuse systems (Cremades et al., 2018; Wang et al., 2018). Clarified spent filter backwash water has frequently been returned to the inflow of a water treatment plant after a sedimentation process (Reissmann and Uhl, 2006). Many studies related to the generation of potable water treatment waste refer only to the treatment, utilisation or disposal of the residue (Cremades et al., 2018; Herselman, 2013; Zhou et al., 2018). In her Water Research Commission report Herselman (2013) acknowledged that there are still information gaps regarding the characteristics of South African water treatment residue and its beneficial use. To the authors'

knowledge, Midvaal Water Company ('Midvaal' hereafter) is the only water treatment plant in South Africa to also include sludge/wastewater from the dissolved air flotation (DAF) process for recycling. Other bulk water treatment plants recycle wastewater from sedimentation and filtration processes only, such as Virginia and Balkfontein water treatment plants in the Free State Province (Oosthuizen and Janse van Vuuren, 2014).

Midvaal purchases source water from the Department of Water and Sanitation, treats it and sells bulk potable water to the local municipality and surrounding mining industries. In an attempt to reduce operational expenses, Midvaal investigated and implemented an upgraded recycling system for wastewater produced in the water treatment plant (Table 1). Most studies report that the recycled treatment wastewater comprises 2 to 10% of the plant's throughput (Bourgeois et al., 2004; Curko et al., 2013). The mean daily wastewater flow from the different unit processes was 9.03% of the maximum operational plant capacity and justified a business case for the recycling of this wastewater. The challenge at the onset of the recycling system was to integrate the sludge from the DAF process, as this contributed to around 46% of the total wastewater flow per day (Table 1). If the quantity and quality aspects of wastewater can be managed, large quantities of river water can be saved by wastewater recycling, thus lowering input costs, reducing environmental pollution and contributing to water security.

The increased demand for freshwater due to continuous worldwide population increase, coupled with the scarcity of clean water, compels stakeholders to explore alternative water sources, especially in South Africa (Marais et al., 2018). Midvaal piloted the recycling of the waste generated from

*Corresponding author, email: sandra.barnard@nwu.ac.za
Received 8 June 2018; accepted in revised form 7 June 2019

the DAF, sedimentation and filtration processes in June 2013 as an initiative to promote cost-effective water utilisation. Herselman (2013 p. 1) defines water treatment residue (WTR) as 'the accumulated solids or precipitate removed from a sedimentation basin, settling tank, or clarifier in a water treatment plant' but WTR of Midvaal includes the waste from the DAF and filtration processes as well, which renders it more of a wastewater than a residue. However, this is a rare practice in South Africa and the efficacy thereof needs to be determined.

It is imperative that the recycling of wastewater should not contribute to the deterioration of the source water and/or impact on final water quality (with special reference to taste and odours), over and above the benefits of saving water and associated costs. The aims of this study were to determine the functionality and water quality of the recent wastewater recycling system at Midvaal and to determine its effect on the overall treatment processes and associated risks regarding wastewater recycling. The wastewater recycling system was partly placed on hold after February 2016 due to maintenance

Table 1. Wastewater volumes generated per day at the dissolved air flotation (DAF), sedimentation and filtration process units of Midvaal Water Company treatment plant

Process unit	Volume (m ³ /day)
DAF scum	5 208
DAF sludge	5 220
Sedimentation	973
Filtration	11 175
Total wastewater (daily flow)	22 576
As % of 250 ML/day (maximum operational plant capacity)	9.03%

needs, and this presented an opportunity for evaluating the failures, benefits and future considerations for this process.

METHODS

Study site

The study site was the Midvaal water treatment plant and its wastewater recycling system (Fig. 1). It consists of a network of gravity pipelines, pumping systems and sludge-handling infrastructure. A sludge-thickening plant and pond system has been in operation since 1994 but has limited capacity for treating the total wastewater, especially since the DAF process was implemented in 1997 to address the high algal load in the middle Vaal River and due to the inability of conventional water treatment to effectively remove algae (Janse van Rensburg et al., 2016). An upgrade to Midvaal's treatment plant commenced in 2013 and entailed the provision of infrastructure to transfer wastewater from the DAF process units to a series of dams and, after some retention, to the sludge balancing dam (SBD). The recycling system comprised two inlet dams and a collection dam. The two inlet dams are fed alternately and chlorine is dosed between the inlet dam in use and the collection dam. Chlorine is dosed again in the canal between the overflow of the collection dam towards the SBD by means of sodium hypochlorite tablets; dosing concentrations range from 6 to 10 mg/L. The wastewater from the sedimentation and filtration processes gravitates directly to the SBD without any treatment (Fig. 1). The SBD thus receives the total volume of wastewater produced during plant operation from where the recovered water is pumped to the source water inlet pipe, prior to pre-ozonation.

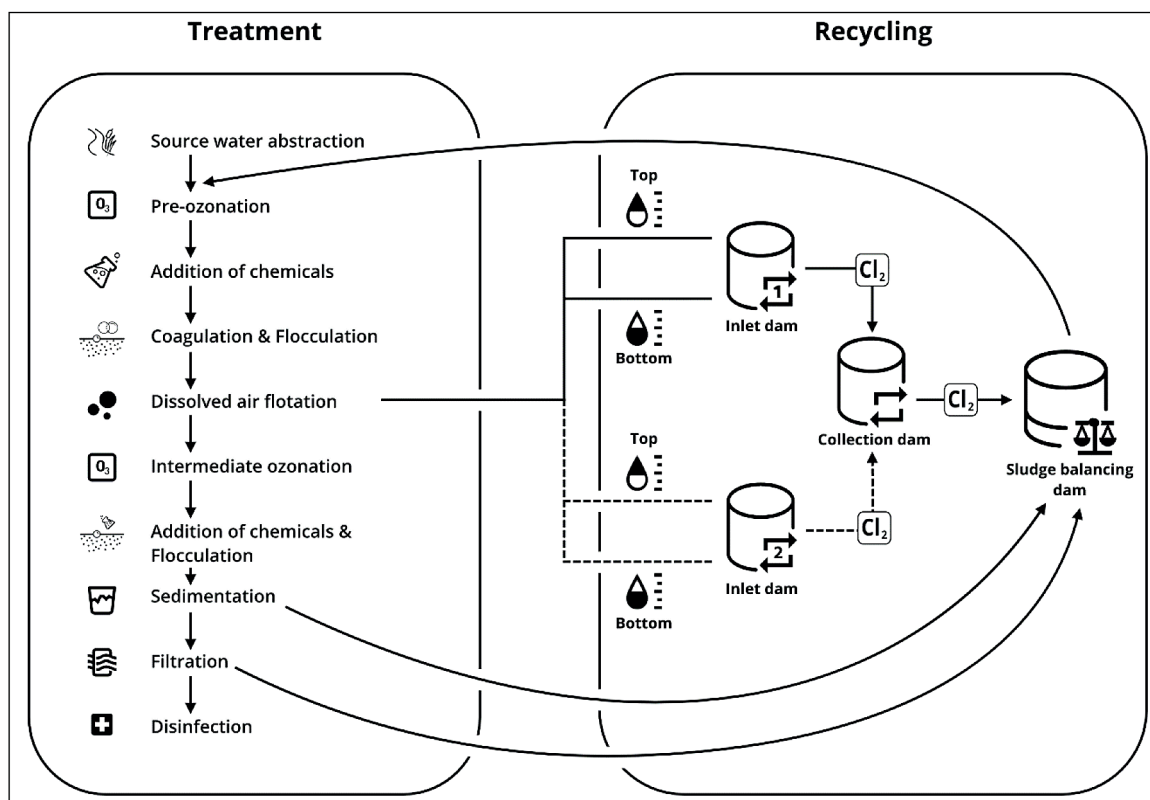


Figure 1. Sequence of the various treatment processes at the Midvaal Water Company and the flow of the wastewater recycling system

Seven sampling sites were identified at the onset of the wastewater recycling system in 2013 (Table 2, Fig. 1) to determine the water quality of the spent water generated during each of the targeted treatment processes.

To determine the impact of the recycled water (from the SBD) on the water treatment plant, water quality data from the river (source water), after DAF, after west sedimentation and after east sedimentation, were collected and statistically analysed. The study period included a year prior to the implementation of the recycling system (June 2012 to June 2013), the operational period of the recycling system (June 2013 to February 2016) and a year after the termination of the wastewater recycling system (February 2016 to February 2017).

Sampling regime

Midvaal's process controllers sampled the DAF top, DAF bottom, east wastewater, west wastewater and the collection dam overflow on a weekly basis from June 2013 to February 2016. Midvaal Water Company Scientific Services sampled water from the river, recycle stream, after DAF, after sedimentation and after filtration on a daily basis. Data reports for the daily samples were generated from the Scientific Services' Laboratory Information Management System from a year prior to the implementation of the recycling system to a year after its termination for statistical analyses (i.e. June 2012 to February 2017). The wastewater from the east and west wastewaters was recycled from February 2016 to February 2017. The samples were chemically and microbiologically analysed by Midvaal Water Company Scientific Services for the 10 determinants listed in Table 3. Samples of the final drinking water were submitted to the CSIR every second month for analyses of protozoan parasites (*Cryptosporidium* oocysts and *Giardia* cysts). Midvaal Water Company Scientific Services has been an accredited South African National Accreditation System (SANAS) testing laboratory (T0132) since 2002 based on the International Organisation for Standardisation 17025 (SANAS, 2018).

Statistical analyses

Microsoft Excel was used to compile column charts to illustrate differences in unit values of determinants for the different sampling sites. Statistica software (version 13) was used to determine descriptive statistics (mean, standard deviation, variance and confidence interval) for all determinants (Tibco Inc, 2017). The Shapiro–Wilks test for normality was used to determine whether the data were distributed parametrically. Since the data did not meet the assumptions of normality in the distribution of all variables, the Kruskal–Wallis analysis of variance (nonparametric statistics) for comparing multiple independent groups was used to determine differences between unit values of determinants measured after treatment process prior to, during and after the implementation of the wastewater recycling system. Results that were below the quantification limit were divided by two to be included in data processing, whereas those that were above the quantification limit were multiplied by two.

RESULTS

Water quality of wastewater

The mean total chlorophyll, suspended solids, turbidity, dissolved organic carbon, pH and electrical conductivity (EC) for each section of the wastewater recycling system are

Table 2. Sampling sites at the Midvaal Water Company wastewater recycling system

Site name	Abbreviation	Description
DAF top	DAF-T	Sludge is withdrawn at top of DAF units due to flotation and transferred to inlet dam
DAF bottom	DAF-B	Sludge is withdrawn at bottom of DAF units due to sedimentation and transferred to inlet dam
East sludge	ES	Combined wastewater from east side sedimentation and filtration of plant transferred to sludge balancing dam (SBD)
West sludge	WS	Combined wastewater from west side sedimentation and filtration of plant transferred to SBD
Collection dam overflow	CDO	Second retention dam that overflows into canal towards SBD
Recycle stream	RS	Water from SBD combined with river water prior to any treatment
Vaal River	R	Source water abstracted from middle Vaal River for treatment prior to introduction of recycle stream

compared in Figs. 2a and 2b. The wastewater from the DAF top had means of 9 825 µg/L total chlorophyll, 3 497 mg/L suspended solids, 2 457 NTU turbidity and 18 mg/L dissolved organic carbon, whereas the DAF bottom had means of 9 148 µg/L total chlorophyll, 4 693 mg/L suspended solids, 4 880 NTU turbidity and 33 mg/L dissolved organic carbon. The quality of the wastewater (sludge from DAF units) that

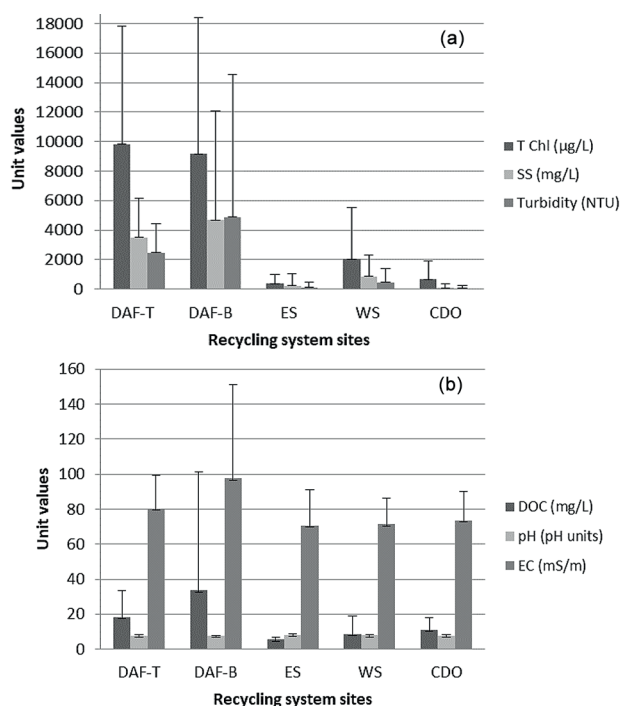


Figure 2. (a & b) Mean total chlorophyll (T Chl), suspended solids (SS), turbidity (NTU), dissolved organic carbon (DOC), pH and electrical conductivity (EC) concentrations with standard deviation error bars of the wastewater sampled at various sites for the water recycling system from 5 June 2013 to 3 February 2016

Table 3. Determinants in the datasets that were statistically analysed in this study; method number refers to the South African National Accreditation System (SANAS)-accredited method as indicated on the facility's schedule of accreditation

Determinant	Unit	Methods/Instruments	Method number
pH	pH units	Determined with a pH electrode	WL1
Electrical conductivity	mS/m	Determined with an electrical conductivity electrode	WL2
Turbidity	NTU	Determined with a turbidity meter	WL3
Suspended solids	mg/L	Gravimetric method	WL5
Aluminium	mg/L	Inductively coupled plasma optical emission spectroscopy	ICP1-A-1
Total organic carbon	mg/L	Determined by a persulfate–ultraviolet oxidation method	AAL5
Spectral absorbance coefficient 254	m ⁻¹	Absorption method	AL6*
Chlorophyll- <i>a</i>	µg/L	Extraction and absorption method	AL1*
Total chlorophyll	µg/L	Determined by means of the Sartory (Swanepoel et al., 2008) extraction method	AL2
<i>E. coli</i>	MPN/ 100 mL	Colilert	BL5-1
<i>Cryptosporidium</i> oocysts and <i>Giardia</i> cysts	Count/10L	Analyses performed by the CSIR	Outsourced

*indicates methods that are not SANAS-accredited

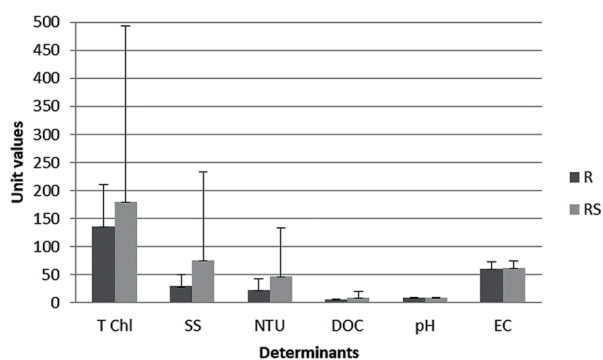


Figure 3. Mean total chlorophyll (T Chl), suspended solids (SS), turbidity (NTU), dissolved organic carbon (DOC), pH and electrical conductivity (EC) concentrations with standard deviation error bars for the Vaal River (R) and recycle stream (RS) from 5 June 2013 to 3 February 2016

overflowed from the collection dam to the SBD improved significantly, with means of 662 µg/L total chlorophyll, 93 mg/L suspended solids, 65 NTU turbidity and 11 mg/L dissolved organic carbon. The total chlorophyll, suspended solids and turbidity of the west wastewater were noticeably higher than those of the east wastewater. The pH and EC for each section's wastewater varied very little over the study period.

Figure 3 illustrates the water quality of the recycle stream compared to the water quality of the middle Vaal River. The mean pH and EC of the recycle stream and river differed significantly according to Kruskal-Wallis although these difference are small (8.4 vs. 8.7 pH and 62 vs. 60 mS/m electrical conductivity for recycle stream and river, respectively). The mean values of total chlorophyll (179 vs. 136 µg/L), suspended solids (75 vs. 29 mg/L), turbidity (46 vs. 22 NTU) and dissolved organic carbon (8.3 vs. 5.7 mg/L) of the recycle stream were higher than that of the river water but only suspended solids and turbidity differed significantly according to Kruskal-Wallis with *p*-values < 0.05. The total chlorophyll concentration in the recycle stream ranged from 11 to 6451 µg/L. The water quality of the recycle stream that originated from the SBD represents an improvement compared to that from the collection dam overflow in terms of total chlorophyll (662 vs. 179 µg/L), suspended solids (93 vs. 75 mg/L), turbidity (65 vs. 46 NTU) and dissolved organic carbon (11 vs. 8.3 mg/L). This observed improvement can be ascribed to a dilution effect when water from

the collection dam overflow was combined with higher-quality water from the east and west wastewater already present in the SBD.

Final water quality failures

To establish whether the recycle stream had any impact on the water quality of the final water produced by the plant, the incidences of water quality failures for the final water were determined. Failures were determined as noncompliance to South African National Standard (SANS) 241:2015. SANS 241:2015 requires aluminium concentrations to comply with a limit of ≤ 0.3 mg/L regarding operational risks and turbidity with limits of ≤ 1 NTU and ≤ 5 NTU for operational and aesthetic risks, respectively. Even though there is no national limit for total chlorophyll, Midvaal has an internal limit of ≤ 1.0 µg/L in the final water (Janse van Rensburg et al., 2016). Aluminium, turbidity and total chlorophyll failures occurred during the entire study period and most failures for all three of these determinants were recorded when the recycling system was in operation (Table 4). Despite these failures, the water quality of the final water during the addition of the recycle stream still complied at ≥ 95% for aluminium and turbidity.

Effect of wastewater recycling system on various treatment processes

Mean aluminium concentrations increased over time in the river and displayed slight increases after DAF and west and east sedimentation during the wastewater recycling process during the same time period (Fig. 4). However, all mean concentrations were below the limit of ≤ 0.3 mg/L at all times.

The increased turbidity levels in the river during the study period were also reflected in the increased turbidity measured after DAF, but the mean turbidity levels remained below 4 and 3 NTU after the west and east sedimentation processes, respectively (Fig. 5).

Mean total chlorophyll concentrations in the river decreased during the study period. This trend was also visible after the west and east sedimentation processes. Total chlorophyll removal was best achieved after DAF during wastewater recycling (Fig. 6).

Mean total organic carbon concentrations were consistently higher in the river and during the wastewater recycling process at all study sites (Fig. 7).

Table 4. Final drinking water failures from June 2012 to February 2017, considering that the recycle stream was in operation from June 2013 to February 2016 as well as the associated risk-defined compliances, as prescribed by South African National Standard 241:2015

	Pre-recycling	During recycling				Post-recycling
		From June 2013	2014	2015	To February 2016	
Aluminium						
Total number of analyses	149	82	147	149	14	143
Aluminium failures	0	1	1	4	0	1
% Compliance	100	99	99	97	100	99
Operational compliance	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent
Turbidity						
Total number of analyses	357	203	361	360	34	367
Turbidity failures	6	6	19	6	0	9
% Compliance	98	97	95	98	100	98
Operational compliance	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent
Aesthetic compliance	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent	(≥95%) – Excellent
Total chlorophyll						
Total number of analyses	242	138	224	237	21	235
Total chlorophyll failures	22	44	37	2	0	15
% Compliance	91	68	83	99	100	94

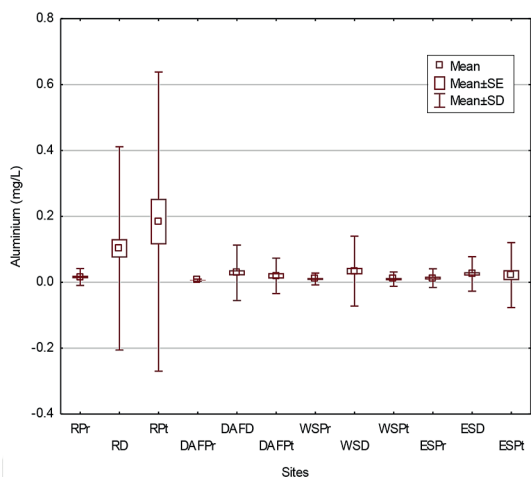


Figure 4. Aluminium concentrations of the Vaal River, after dissolved air flotation (DAF), after west and east sedimentation for the periods prior to, during and after implementation of wastewater recycling; RPr, RD and RPt: river pre, during and post recycling system; DAFPr, DAFD and DAFPt: DAF pre, during and post recycling system; WSPr, WSD and WSPt: west sedimentation pre, during and post recycle system; ESPr, ESD and ESPT: east sedimentation pre, during and post recycling system

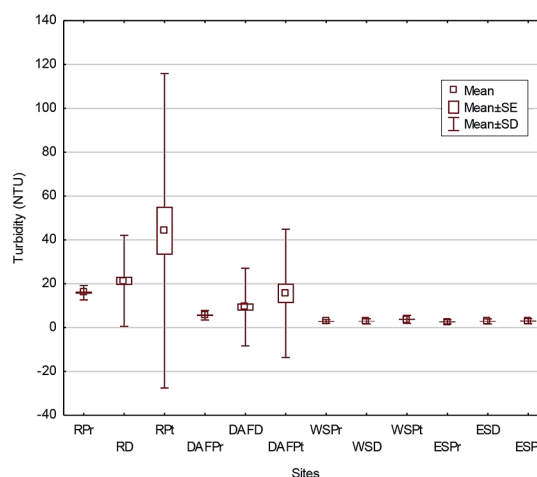


Figure 5. Turbidity levels of the Vaal River, after dissolved air flotation (DAF), after west and east sedimentation for the periods prior to, during and after implementation of wastewater recycling; RPr, RD and RPt: river pre, during and post recycling system; DAFPr, DAFD and DAFPt: DAF pre, during and post recycling system; WSPr, WSD and WSPt: west sedimentation pre, during and post recycle system; ESPr, ESD and ESPT: east sedimentation pre, during and post recycling system

Mean *E. coli* concentrations continuously increased over time in the river and after DAF but remained at ≤ 10 MPN/100 mL after the west and east sedimentation processes (Fig. 8).

The Kruskal–Wallis test was applied to data collected for pH, EC, spectral absorbance coefficient 254 and chlorophyll-*a*, but no significant differences between unit values of these determinants for the treatment processes prior to, during and after implementation of the wastewater recycling system were detected.

The water quality after the filtration process was not included in Figs. 4 to 8 as differences between periods prior

to, during and after implementation of wastewater recycling were negligible and this water represents the final drinking water, the quality of which is addressed in Tables 4 and 5.

In Table 5 the mean total chlorophyll was highest after west and east filtration during the recycling process but at levels and concentrations that complied with limits. Maximum pH levels and total chlorophyll concentrations exceeded limits throughout the study period except for the maximum total chlorophyll of the west filtration after recycling.

No *Cryptosporidium* oocysts or *Giardia* cysts were detected in the final water prior to, during or after recycling.

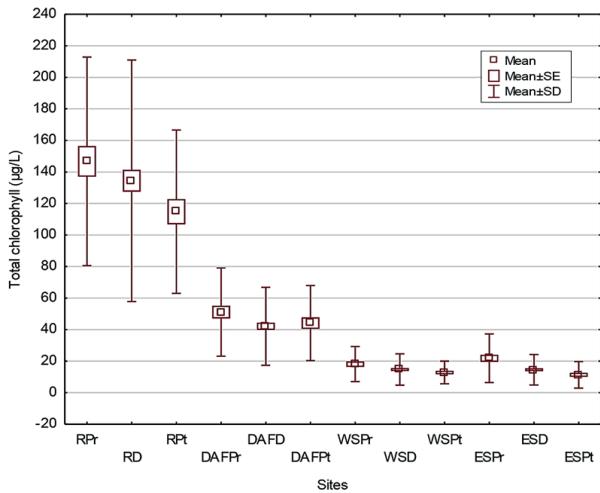


Figure 6. Total chlorophyll concentrations in the Vaal River, after dissolved air flotation (DAF), after west and east sedimentation for the periods prior to, during and after implementation of wastewater recycling; RPr, RD and RPt: river pre, during and post recycling system; DAFPr, DAFD and DAFPt: DAF pre, during and post recycling system; WSPr, WSD and WSPt: west sedimentation pre, during and post recycle system; ESPr, ESD and ESPt: east sedimentation pre, during and post recycling system

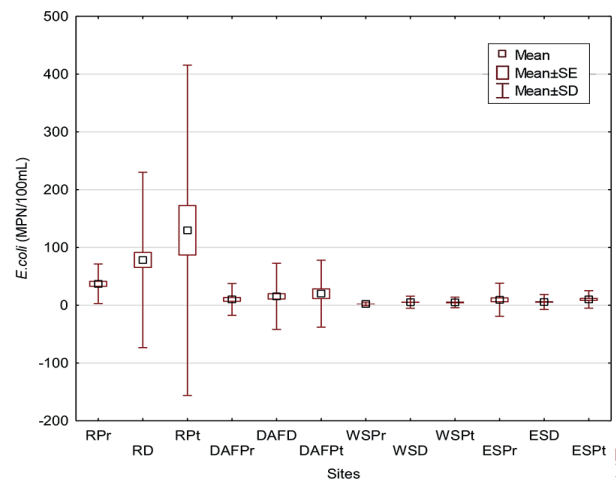


Figure 8. *E. coli* concentrations of the Vaal River, after dissolved air flotation (DAF), after west and east sedimentation for the periods prior to, during and after implementation of wastewater recycling; RPr, RD and RPt: river pre, during and post recycling system; DAFPr, DAFD and DAFPt: DAF pre, during and post recycling system; WSPr, WSD and WSPt: west sedimentation pre, during and post recycle system; ESPr, ESD and ESPt: east sedimentation pre, during and post recycling system

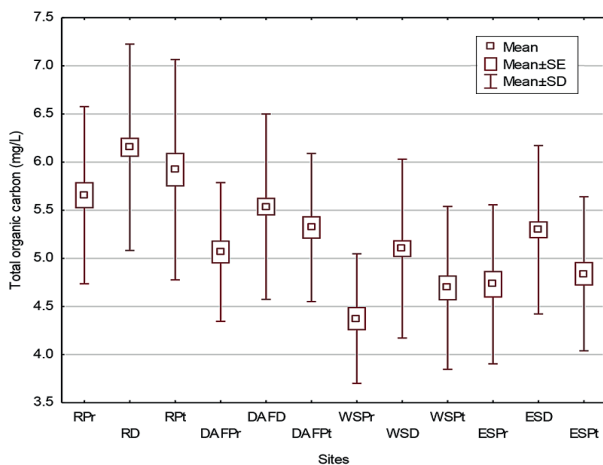


Figure 7. Total organic carbon concentrations of the Vaal River, after dissolved air flotation (DAF), after west and east sedimentation for the periods prior to, during and after implementation of wastewater recycling; RPr, RD and RPt: river pre, during and post recycling system; DAFPr, DAFD and DAFPt: DAF pre, during and post recycling system; WSPr, WSD and WSPt: west sedimentation pre, during and post recycle system; ESPr, ESD and ESPt: east sedimentation pre, during and post recycling system

DISCUSSION

Concerns about water recycling with regards to microorganisms (such as micro-algae and *E. coli*), heavy metals and increasing turbidity are some reasons why many water treatment plants have not returned wastewater to the water treatment process. Reissmann and Uhl (2006) were concerned about the recycling of precursors for disinfection by-products. During this investigation the sludge from DAF largely contributed to the poor water quality of the water recycling system. The total chlorophyll, suspended solids, turbidity and dissolved organic carbon of the

sludge transferred from the DAF top and bottom to the inlet dam were extremely high (Figs. 2a and 2b), but the concentrations of these determinants for the wastewater that overflowed from the collection dam to the SBD improved significantly. The effect of retention to allow for the settling of suspended matter in the holding and collection dams was noticeable. Haarhoff et al. (2001) also concluded that the turbidity of the supernatant on the sludge (sedimentation wastewater) and washwater (filtration wastewater) at the Vaalkop treatment plant was mostly lower than that of source water abstracted from the Vaalkop dam. Bourgeois et al. (2004) have shown that optimising the monovalent: divalent cation balance ratio can improve the quality of combined filtered backwash wastewater by sedimentation. The west and east sedimentation and filtration seemed to be equally effective when values were compared during recycling and did not verify the poorer water quality of the west sludge in Fig. 2a.

The total chlorophyll, suspended solids and turbidity of the recycle stream were identified as risks due to extreme concentrations (Fig. 3) together with the outcomes from Janse van Rensburg et al.'s (2016) study, which confirmed that increasing total chlorophyll concentrations and turbidity spikes were the main source-water quality challenges for Midvaal. Haarhoff et al. (2001) stated that the rate of solids production, associated with turbidity levels in wastewater, at treatment plants treating inland surface water, are highly variable and occasionally reach extremely high peaks (Haarhoff et al., 2001).

Aluminium is introduced as a water treatment chemical (aluminium sulfate) and concentrations were below the required limit for drinking water in the river despite the observed increase during 2012–2017 (Fig. 4). During the time that aluminium failures were recorded in the final water, neither aluminium concentrations nor turbidity levels were water quality concerns in the river water and it could be ascribed to the dosing of aluminium sulfate during the water treatment process. The turbidity levels of both the river and

recycle stream were not significantly high during times when turbidity failures were recorded and no pattern/correlation could be established. Moreover, the turbidity levels of the recycle stream were not necessarily higher than those of the river during times when turbidity failures were recorded. The mean pH of the river was 9.0 and the median 9.1 during times when turbidity failures occurred. The aluminium and turbidity failures only corresponded once (25 February 2015); the same was true for aluminium, turbidity and total chlorophyll failures (14 July 2014).

The total chlorophyll concentrations of the river prior to and after the wastewater recycling ranged from 57 to 373 µg/L and from 68 to 196 µg/L, respectively, during times when failures were recorded. Total chlorophyll concentrations of the river and recycle stream ranged from 57 to 393 µg/L and from 15 to 6 451 µg/L, respectively, during recycling at times when failures were recorded and no seasonal correlation could be confirmed. Total chlorophyll concentrations were highest after west and east filtration during the recycling process. The total chlorophyll concentrations increased from 0.6 to 0.8 µg/L after both west and east filtration during the wastewater recycling and is a cause for concern as it is close to the limit of ≤ 1.0 µg/L (Table 5). These total chlorophyll concentrations after west and east filtration do not correspond with the decreasing trends depicted in Fig. 6 after west and east sedimentation. The impact on the filtration process and irregular failures in terms of total chlorophyll have to be monitored and managed closely during wastewater recycling, especially since no particular correlation could be identified from this research. In the event of taste and odour episodes, recycling has to be terminated until the situation is resolved.

Total organic carbon concentrations in the river peaked during the recycling period, but all mean total organic carbon concentrations, including those for the river, were below the limit of ≤ 10 mg/L (Fig. 7) and complied with SANS 241:2015 requirements at all times (SABS, 2015). *E. coli* and turbidity in the river displayed similar increasing trends (Figs. 5 and 8, respectively) as opposed to the decreasing total chlorophyll concentrations (Fig. 6). Turbidity can inhibit photosynthesis by blocking sunlight and, as algae decay, the decomposition process allows organic particles to release as suspended solids and contribute to turbidity (Fondriest Environmental, Inc., 2014).

Supplementary benefits of wastewater recycling

The source water tariff for 2016 was R 3.22/kL which results in an average saving of approximately R 27 million per year when calculated:

$$22\,576 \text{ kL/day (Table 1)} \times 365 \text{ (days per year)} \times R\,3.22/\text{kL (2016 water tariff)} = R\,26\,533\,572.$$

Table 5. Descriptive statistics for turbidity levels and total chlorophyll concentrations after west and east filtration processes, where failures are shaded; WFPr, WFD and WFPt: west filtration pre, during and post recycle system; EFP, EFD and EFPt: east filtration pre, during and post recycle system

	Turbidity (NTU)				Total chlorophyll (µg/L)			
	Mean	Minimum	Maximum	Standard deviation	Mean	Minimum	Maximum	Standard deviation
WFPr	0.5	0.2	1.0	0.2	0.6	0.3	2.3	0.5
WFD	0.6	0.2	1.9	0.3	0.8	0.3	14	1.4
WFPt	0.6	0.4	1.0	0.2	0.4	0.3	0.7	0.2
EFP	0.6	0.3	2.0	0.3	0.6	0.1	3.0	0.6
EFD	0.7	0.2	2.4	0.3	0.8	0.3	3.7	0.6
EFPt	0.7	0.3	2.3	0.4	0.4	0.3	1.5	0.3

Wastewater recycling systems have several benefits in addition to reducing water treatment expenses. Such a system allows for suspended solids to settle out and thus contributes to turbidity reduction (Fig. 2a). Natural microbiological processes result in improved microbiological quality of the wastewater due to disinfection via sunlight. Furthermore, the SBD attenuates the flow rate of the recycled stream, which results in more consistent and controllable introduction of wastewater into the river water inlet stream. The health-related risks of recycling wastewater are considered manageable as the recycle stream is introduced at the beginning of the treatment process, prior to pre-ozonation and subsequent treatment processes to ultimately produce potable water for consumers. Furthermore, Cornwell and MacPhee (2001) reported that only when spent filter backwash water of up to 20% volume was recycled was *Cryptosporidium* removal affected.

Future considerations regarding wastewater recycling

In light of the findings of this study, several points regarding the implementation of wastewater recycling systems should be considered. Periodic blue-green algal blooms and poor source water quality require temporary termination of the wastewater recycling system, as this has a negative impact on the treated water quality if recycled, e.g., in terms of taste and odour. Water from the recycle dams at Virginia and Balkfontein water treatment plants was discharged into the Vaal River during prevalent algal blooms (Oosthuizen and Janse van Vuuren, 2014). Extension of the dam surface area is necessary as provision for the total wastewater flow to curtail recycling of poor-quality water plugs and prevent possible discharge into the environment. Anaerobic/anoxic conditions may develop in the dams and some form of mechanical agitation/aeration has to be considered for the future. Additional disposal sites are necessary for dry sludge, which would subsequently require further analyses and classification. Microbiological monitoring of the recycle stream could expand to include contaminants of emerging concerns, such as endocrine-disrupting compounds and persistent organic pollutants. Trihalomethane formation monitoring is required due to the high organic load (dissolved organic carbon) in the wastewater recycling system and the dosing of chlorine at two points (Fig. 1).

CONCLUSION

It is evident from the available data that wastewater recycling, which included wastewater from the DAF plant, into the main inlet stream of the water treatment plant proved to be effective based on SANS 241:2015 compliance and had no detrimental impact on overall treatment processes or final water quality.

The wastewater recycling system operated successfully and was cost-effective due to reduced river water purchases. Total chlorophyll was identified as the most prominent risk when wastewater is recycled due to the high concentration in both the river and recycle stream, borderline concentrations of 0.8 µg/L after filtration during recycling and failures during recycling. Water quality of the sludge from DAF units improved significantly after it was subjected to retention in the dam system and dilution with wastewater from the sedimentation and filtration processes. Final water quality failures recorded for aluminium, turbidity and total chlorophyll occurred mostly during the recycling period, but the risk-defined compliances were calculated and categorized as excellent ($\geq 95\%$) for both aluminium and turbidity in the periods prior to, during and after wastewater recycling. The total chlorophyll compliance was 94% in the year prior to the implementation of the wastewater recycling system, 88% during recycling and 96% in the year after recycling. The findings of this case study are, however, based on retrospective data evaluation and could not take all of the factors that contributed to water quality on the Midvaal treatment plant into account.

ACKNOWLEDGEMENTS

Midvaal Water Company is gratefully acknowledged for sharing information and data. We would like to thank Mr CF de Villiers (Plant Foreman – Midvaal Water Company) and Prof J Haarhoff for their valuable contributions and discussions, Mr A Marais for assistance with the design of the illustration and Ms E Harris for editorial input. We would like to thank the National Research Foundation (UID: 296036) for financial support.

REFERENCES

- BOURGEOIS JC, WALSH ME and GAGNON GA (2004) Treatment of drinking water residuals: comparing sedimentation and dissolved air flotation performance with optimal cation ratios. *Water Res.* **38** 1173–1182. <https://doi.org/10.1016/j.watres.2003.11.018>.
- CORNWELL DA and MACPHEE MJ (2001) Effects of spent filter backwash recycle on Cryptosporidium removal. *J. Am. Water Works Assoc.* **93** (4) 153–162. <https://doi.org/10.1002/j.1551-8833.2001.tb09185.x>.
- CREMADES LV, CUSIDÓ JA and ARTEAGA F (2018) Recycling of sludge from drinking water treatment as ceramic material for the manufacture of tiles. *J. Clean. Prod.* **201** 1071–1080. <https://doi.org/10.1016/j.jclepro.2018.08.094>.
- ĆURKO J, MIJATOVIĆ I, RUMORA D, CRNEK V, MATOŠIĆ M and NEŽIĆ M (2013) Treatment of spent filter backwash water from drinking water treatment with immersed ultrafiltration membranes. *Desalin. Water Treat.* **51** (25–27) 4901–4906. <https://doi.org/10.1080/19443994.2013.774142>.
- FONDRIEST ENVIRONMENTAL, INC. (2014) Turbidity, Total Suspended Solids and Water Clarity. Fundamentals of Environmental Measurements. URL: <https://www.fondriest.com/environmental-measurements/parameters/water-quality/turbidity-total-suspended-solids-water-clarity/> (Accessed 7 May 2018).
- HAARHOFF J, VAN HEERDEN P and VAN DER WALT M (2001) Sludge and washwater management strategies for the Vaalkop water treatment plant. *Water Sci. Technol.* **44** (6) 73–80. <https://doi.org/10.2166/wst.2001.0344>.
- HERSELMAN JE (2013) Guidelines for the utilisation and disposal of water treatment residues. WRC Report No. TT559/13. Water Research Commission, Pretoria.
- JANSE VAN RENSBURG S, BARNARD S and KRÜGER M (2016) Challenges in the potable water industry due to changes in source water quality: case study of Midvaal Water Company, South Africa. *Water SA* **42** (4) 633–640. <http://dx.doi.org/10.4314/wsa.v42i4.14>.
- MARAIS SS, NDLANGAMANDLA NG, BOPAPE DA, STRYDOM WF, MOYO W, CHAUKURA N, KUVAREGA AT, DE KOCK L, MAMBA BB, MSAGATI TAM and co-authors (2018) Natural Organic Matter (NOM) in South African Waters. WRC Report No. 2468/1/18. Water Research Commission, Pretoria.
- OOSTHUIZEN MGJ and JANSE VAN VUUREN S (2014) Changes in density and composition of algal assemblages in certain unit processes of two South African water purification plants. *Afr. J. Aquat. Sci.* **39** (3) 313–326. <https://doi.org/10.2989/16085914.2014.957638>.
- REISSMANN FG and UHL W (2006) Ultrafiltration for the reuse of spent filter backwash water from drinking water treatment. *Desalination* **198** (1–3) 225–235. <https://doi.org/10.1016/j.desal.2006.03.517>.
- SABS (South African Bureau of Standards) (2015) South African National Standard 241-1:2015 Drinking water, Part 1: Microbiological, physical, aesthetic and chemical determinants. 241-2:2015 Drinking water, Part 2: Application of SANS 241-1. SABS, Pretoria.
- SANAS (South African National Accreditation System) (2018) Directory of Accredited Facilities. URL: <http://www.sanas.co.za/schedules/testing/T0132-05-2018.pdf>. (Accessed 6 May 2018).
- SWANEPOEL A, DU PREEZ H, SCHOEMAN C, JANSE VAN VUUREN S and SUNDRAM A (2008) Condensed laboratory methods for monitoring phytoplankton, including cyanobacteria, in South African freshwaters. WRC Report No. TT 323/08. Water Research Commission, Pretoria.
- TIBCO SOFTWARE INC. (2017) Statistica (data analysis software system), version 13. URL: <http://statistica.io> (Accessed 20 January 2019).
- WANG C, WU Y, BAI L, ZHAO Y, YAN Z, JIANG H and LIU X (2018) Recycling of drinking water treatment residue as an additional medium in columns for effective P removal from eutrophic surface water. *J. Environ. Manage.* **217** 363–372. <https://doi.org/10.1016/j.jenvman.2018.03.128>.
- ZHOU Z, YANG Y, LI X, LI P, ZHANG T, LV X, LIUC L, DONG J and ZHENG D (2018) Optimized removal of natural organic matter by ultrasound-assisted coagulation of recycling drinking water treatment sludge. *Ultrasonics Sonochem.* **48** 171–180. <https://doi.org/10.1016/j.ultsonch.2018.05.022>.

Membrane surface properties and their effects on real waste oil-in-water emulsion ultrafiltration

Marjana Simonič¹

¹Faculty of Chemistry and Chemical Engineering, University of Maribor, Smetanova 17, 2000 Maribor, Slovenia

ABSTRACT

Membrane surface properties and their effect on the efficiency of ultrafiltration (UF) of real waste oily emulsions was studied. Experiments were performed in cross-flow operation at total recycle condition in a lab-scale system. The ceramic UF membrane in the tubular type module was employed. During the experiments permeate flux was measured. The most important influential factors, such as temperature, TMP, and pH, were considered during the experiments. Zeta potential was measured in order to explain the phenomena on the membrane surface. The isoelectric point of the fouled membrane was shifted to the alkaline range. COD removal efficiency reached 89%. Gas chromatography measurements were performed in order to determine the composition of waste emulsions. SEM micrographs showed the formation of calcite on the membrane, which contributed to membrane fouling. Chemical cleaning was examined using alkaline and acid solutions, and a cleaning strategy was determined.

Keywords: ceramic ultrafiltration, waste oily emulsions, cleaning, scaling

INTRODUCTION

Real oil-in-water emulsions contain up to 95% water and 5% solvents, such as alcohols (nonyl-phenols, ethoxylated alcohols), formaldehyde-based biocides, sequestrate media, corrosion inhibitors (benzotriazole, propylene glycole) and surfactants (amine propoxylate) (Lobo et al., 2006; Bennito et al., 2010). The contents are not known in detail due to confidential formulas of producers. The quantity of additives varies in different products, even if the purpose of the product is the same (Lin and Lan, 1998). In practice, different sources of emulsions are mixed together; therefore, such waste emulsions are very difficult to treat due to their diverse concentration and composition (Cheng et al., 2005; Gutiérrez et al., 2007). More than one process has to be adopted for efficient emulsion treatment (Bensadok et al., 2007). Pre-treatment is mostly done by breaking the emulsion using coagulants (Cañizares et al., 2008). In the next step, oil and organic additives are separated from the de-emulsified solution.

The application of ultrafiltration (UF) has been accepted as a highly effective process to treat oil-in-water emulsions. Membrane separation techniques are superior to conventional treatment (Lobo et al., 2006; Fu and Chung, 2011; Fenglin 2006). Razvanpour et al. (2009) and Hesampour et al. (2008) studied the effect of factors such as pH, temperature, transmembrane pressure (TMP), oil content, and membrane type on flux. pH value influences the membrane surface charge and consequently oil retention. If pH increases emulsion stability increases (Razvanpour et al., 2009). It was also found that flux increases with increasing pH, while oil retention decreases. Zeta potential is the main indicator of emulsion stability (Coca et al., 2011) and should be slightly negative. The same authors claim that non-ionic surfactants improve the stability of emulsion.

However, the efficiency of ultrafiltration is limited by membrane fouling and concentration polarisation, resulting in a permeate flux decline and, therefore, increasing energy

consumption of the process (Faibish and Cohen, 2001). Fouling can be divided into reversible fouling, which can be removed by physical cleaning, and irreversible fouling that must be treated with chemical agents.

In an oil-in-water emulsion the drops are smaller than 20 μm (Hu et al., 2002), thus the application of microfiltration, and its efficiency and cleaning, has been extensively studied (Silalahi et al., 2009; Hua et al., 2007; Zhou et al., 2010; Ghandashtani et al., 2015). The use of ultrafiltration is particularly interesting for the treatment of cutting oil used in various industries, since the value of oil rejection is high. Emulsion separations have been studied at laboratory scale by mixing a certain amount of carefully selected oils in water, optionally with surfactant (Križan et al., 2014; Ju et al., 2015). However, we did not find any studies reported in the literature on real waste emulsions particularly those collected from cutting fluid industries, mixed together followed by separation of oil from water. An enterprise in Slovenia is dealing with such emulsions containing various cutting fluids. Therefore, this research was initiated to investigate the feasibility of using ceramic UF membranes for oil removal from waste emulsions after the emulsion breaking process.

The aim of this study was to evaluate the performance of $\alpha\text{-Al}_2\text{O}_3/\text{ZrO}_2$ tubular ceramic membrane for oil removal from real waste emulsions and to study different types of cleaning agent for efficient flux recovery. The UF membrane has a pore size of 50 nm. UF performance using a model solution was reported for a previous study (Križan et al., 2014). In this study, UF performance was studied and eventually the membrane was fouled by waste emulsion. Chemical cleaning was done using alkaline and acid solutions as well as a combination thereof. After chemical cleaning the membrane was thoroughly washed with water and cleaning efficiency was determined by assessing flux recovery compared with the permeability of non-fouled membrane. Cleaning efficiency was studied at two different temperatures, concentrations and TMPs. The novelty of this study lies in the measurement of zeta potential, together with other techniques, such as SEM and GC/MS analyses, in order to explain membrane fouling phenomena during ultrafiltration of real waste emulsions.

*Corresponding author, email: marjana.simonic@um.si

Received 29 August 2017; accepted in revised form 27 May 2019

METHODS

Slovene enterprise is engaged in real oil-in-water emulsion transport and handling. Around 2 000 t/yr of emulsion from different industrial branches are mixed together, transported and collected in one large reservoir. Firstly, emulsion breaking is performed using $\text{Al}_2(\text{SO}_4)_3$ and anionic flocculant (PAM A-100). Samples were taken after this pre-treatment step in order to determine water quality. COD was determined to be between 18 000 and 30 000 mg/L O_2 . These values are far above the limit values for discharge into water resources, which for COD in Slovenia is determined at 120 mg/L O_2 (Official Gazette of Republic Slovenia, 2005). Secondly, ultrafiltration (UF) was chosen as the next step in the emulsion treatment procedure. Three real oil-in-water samples (denoted N1, N2 and N3) were taken after pre-treatment over a period of 1 month. The pre-treatment was directed towards reduction of COD and turbidity.

UF performance

Filtration experiments were performed in a lab-scale plant in cross-flow operation. The ceramic membrane, with a pore size of 50 nm and an active area 0.418 m², was integrated in a tubular stainless-steel module. Membrane characteristics are summarized in Table 1.

Figure 1 shows a schematic diagram of the filtration equipment used in this study. The oil-in-water solution was pumped by motor pump (PK) from a feed tank (Tank B1) into a tubular module (membrane module). VK4 valve allowed the regulation of TMP. Concentrate as well as permeate were returned to the feed tank.

Analyses

pH, electro-conductivity and COD were measured according to standard methods SIST ISO 10523, SIST EN 27888 and SIST ISO 6060, respectively. Preparation of samples, headspace-solid phase microextraction (HSME) and GC/MS conditions are reported elsewhere (Pickl et al., 2011).

Zeta potential measurements

The streaming current measurements were done using an electro-kinetic analyser (SurPASS, Anton Paar GmbH, Austria) equipped with a cylindrical cell, where pieces of membrane ($2r > 25 \mu\text{m}$) were mounted into the measuring cell. A 1 mM KCl solution was used as the background electrolyte and,

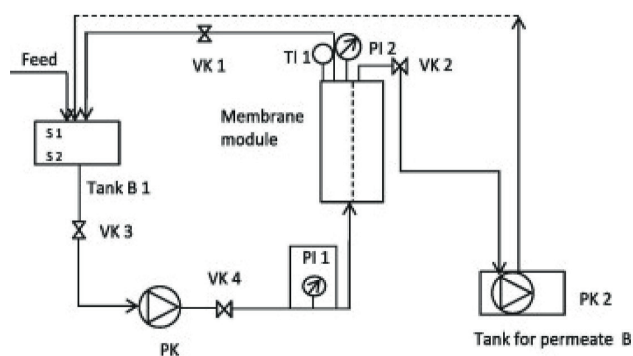


Figure 1. A schematic diagram of the filtration equipment with ceramic membrane module

Table 1. Characteristics of ceramic membrane

Characteristic	Description
Material	$\alpha\text{Al}_2\text{O}_3\text{-ZrO}_2$
Dimensions	40 mm × 1 000 mm
Outside diameter	40 mm
Channel diameter	3.6 mm
Quantity of channels	37
Membrane area	0.418 m ²
Pore size	50 nm
pH range	0–14
Operating pressure	200–1 000 kPa
Maximum temperature	100°C

prior to measurement, the given sample was rinsed with this aqueous solution. The pH dependence of the zeta potential within the range pH 6–9 was determined using 0.1 M NaOH as the titration liquid. The zeta potential was calculated from the measured streaming current using the Helmholtz–Smolouchowski equation, which takes into account any surface conductivity (Križan et al., 2014).

Membrane cleaning

Membrane cleaning was necessary after each experiment, due to membrane fouling during filtration. Two alkaline solutions, ASG (PRU 06-03, Gütling, Germany) and US73 (Ultrasil 73, Henkel-Ekolab Ltd., Germany), and two acid solutions, H_2SO_4 (pH = 3) and HCl (pH = 3), were applied for membrane cleaning. Solutions were chosen based on the previous experience of the research team, as reported in Križan et al. (2014), with some improvements. Cleaning was done at first by flushing the membrane with pure water for 5 min. Then alkaline solution was heated to 40°C during cleaning and the filtration run for 1 h at a pressure of 100 kPa. In the next step the whole system was flushed with pure water for 5 min followed by acidic cleaning at room temperature for 30 min. In the last step the membrane was flushed with pure water. Verification of the cleaning was done by comparing the water flux after filtration with the initial water flux measured on a virgin membrane under similar conditions. The sustainable value for the flux J_{rec} to be achieved was 220 L/(m²·h) at pressure of 200 kPa. Six different strategies for cleaning were implemented, as seen in Table 2.

Scanning electron microscopy

After the experiment, in order to visualize the structure of the deposits formed due to adsorption on the membrane surface, the effluent solution after cleaning of the fouled membrane was investigated by a scanning electron microscope (SEM), Quanta 200 3D (FEI Company).

Table 2. Cleaning strategies

Cleaning strategy	Description
1	ASG + HCl
2	ASG
3	US73
4	US73 + H_2SO_4
5	ASG + H_2SO_4
6	US73+HCl

RESULTS AND DISCUSSION

Zeta potential of clean membrane

The zeta potential versus pH data for the experiments performed is shown in Fig. 2. An iso-electrical point (IEP) for applied ceramic membrane was determined at pH = 5.3. For this membrane, the surface charge is positive at lower pH range, passes through the IEP and becomes negative in the higher pH range. The shape of the zeta potential curve is indicative of amphoteric surfaces. Varying values for ceramic membrane IEP, ranging from 5.5 to 8.2, have been reported (Wakily et al., 2010). According to Lin (2006), the maximum permeability was obtained as a result of pH values roughly corresponding to the IEP of the membranes. Therefore, the UF experiments took place at pH = 5, but also at original pH.

UF operation

The results with model emulsions were thoroughly investigated in a previous study (Križan et al., 2014). The pure water flux and that of waste emulsion were determined. Figure 3a shows the fluxes of pure water (J_w), emulsion at original pH (J_e) and pure water after back-flushing (J_{wb}), and Fig. 3b the fluxes of pure water (J_w), emulsion at pH = 5 (J_e) and pure water after back-flushing (J_{wb}), respectively.

The emulsion flux increases with TMP. If the UF was performed at original emulsion pH, the flux was close to zero at TMP below 150 kPa, while emulsion flux reached 22 L/(m²·h) in the same TMP region if the emulsion pH was adjusted to 5. The flux increase is in agreement with another study (Tanis-Canbur et al., 2018) which showed that permeate flux increased if the electrostatic interaction (determined by zeta potential) between membrane and oil emulsion was low. In our experiments there was low interaction between membrane (IEP at pH = 5.3) and emulsion with pH = 5. Consequently, the adsorption of surfactants and oil onto the membrane was limited, leading to a decrease in cake layer formation.

It is seen that the emulsion flux is higher at pH = 5 than at original emulsion pH, and increases with TMP. After the back-flush, pure water flux was measured in order to determine the fouling rate. The results are denoted in Fig. 3a and 3b as ' J_{wb} '. It is clearly seen that the flux after back-flushing is higher in Fig. 3b and the slope of J_{wb} in Fig. 3b is steeper than that in Fig. 3a. Therefore, it was confirmed that the emulsion pH near IEP (see Fig. 2) allowed higher emulsion fluxes than at original emulsion pH, due to a lower fouling rate, as reported for other studies (Wakily et al., 2010).

The same procedure as described above for the UF protocol was used with Sample N2. The results are not presented for

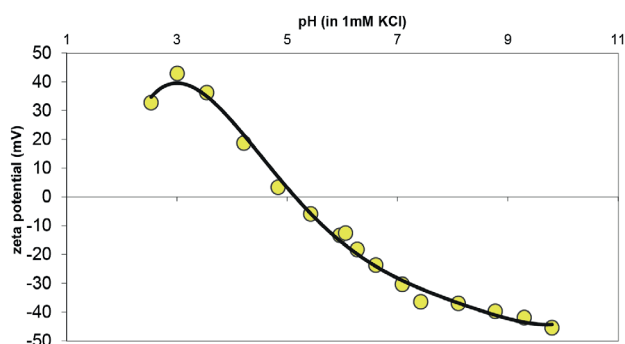


Figure 2. Effect of pH on zeta potential of clean membrane

Sample N2 as these were almost the same as for Sample N1 at original pH (Fig. 3a), and even if pH was adjusted to 5. The UF trials with Sample N3 were performed with an emulsion with pH = 5. The results for UF are shown in Fig. 4. The flux of emulsion at pH = 5, the pure water flux and the water flux after back-flushing of the membrane, were very similar to those presented in Fig. 3b. The permeate flux of the waste emulsion was a little higher compared with Sample N1 (see Fig. 3b). It was assumed that the slight Sample N3 flux increase was due to a slightly lower COD, shown in Table 3.

The COD varies widely across Samples N1, N2 and N3 (Table 3). The COD is lower for Samples N1 and N3 and differs by 8%, while the COD difference between N2 and N3 exceeded 23%. Higher organic pollutant content in Sample N2 caused an increase in cake layer formation (Tanis-Kanbur et al., 2018) and consequently lower emulsion flux. The limiting flux of 100 L/(m²·h) was reached with acidic samples N1 and N3, as shown in Fig. 3b and Fig. 4. Almost the same limiting flux value was previously reported by Matos et al. (2008).

Physico-chemical analyses

pH, electro-conductivity and COD were measured in untreated samples, denoted by N1, N2 and N3. UF-treated samples where pH was adjusted to 5 are denoted by N1b, N2b and N3b. Sample N1a is treated Sample N1 at original pH value. Results are presented in Table 3. As seen from Table 3 the electro-conductivity is very high, in all cases above 10 mS/cm. It is seen that COD decreased by up to 85% for Sample N1, by 84% for

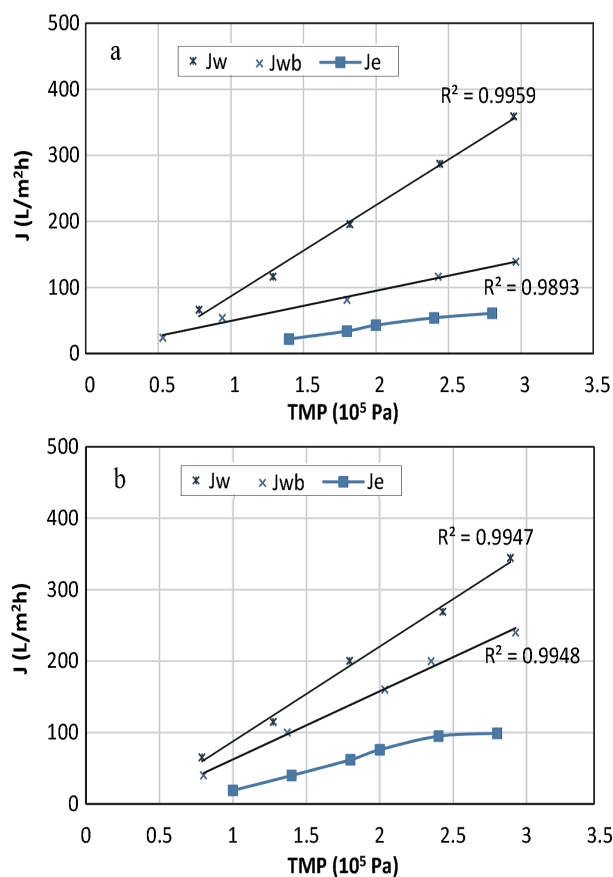


Figure 3. Pure water flux, pure water flux after back-flushing and emulsion N1 flux at (a) original emulsion pH and (b) at emulsion pH = 5

Table 3. Physico-chemical analyses

Sample	pH	χ (mS/cm)	Turbidity (NTU)	COD (mg O ₂ /L)
N1	7.6	13.91	350	19 490
N1a	7.6	10.97	4.2	4 450
N1b	5.0	7.49	3.6	2 840
N2	8.1	10.09	850	25 380
N2b	5.0	5.19	5.0	3 980
N3	7.1	7.20	160	17 950
N3b	5.0	2.56	3.6	3 770

Sample N2 and by 79% for Sample N3. pH of original samples varied from 7.1 to 8.1. It was discovered that pH has an important influence on COD removal. This could be attributed to the particle size distribution measurement, in accordance with the results of a previous study (Križan et al., 2014), where particles were determined to be larger at pH = 5 than in neutral range. COD of UF-treated samples was nearly 8% lower if the pH was adjusted to 5 (N1b) compared to samples treated at original pH (N1a). Turbidity was removed by 99.3% and the emulsion was within the drinking water quality limit for turbidity in all three treated samples.

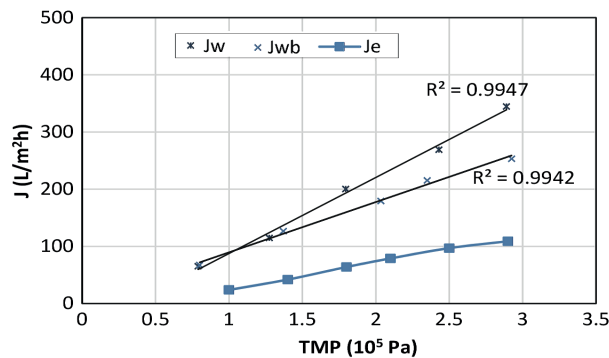
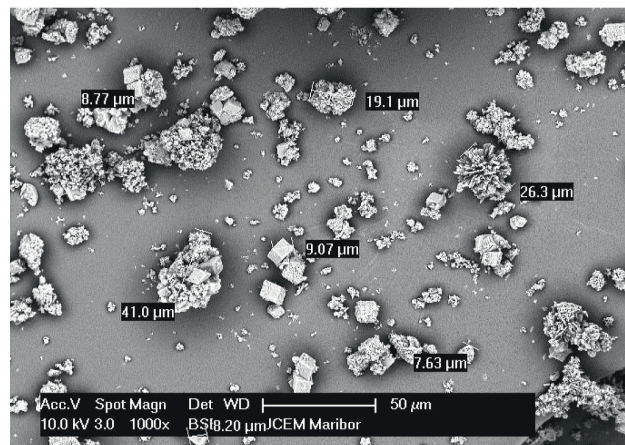
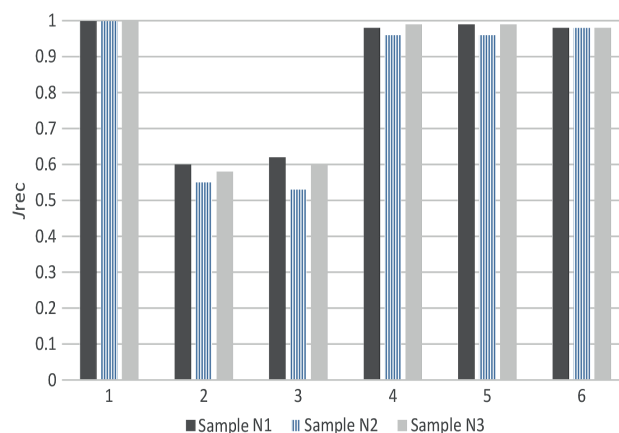
Membrane cleaning

Efficiency of membrane cleaning was determined by assessing flux recovery compared with the permeability of non-fouled membrane. Fouled membranes were cleaned using alkaline solution ASG at 40°C; however, the flux did not reach the initial flux of 220 L/(m²·h) (at 200 kPa) after cleaning. The SEM image of the effluent solution after cleaning is presented in Fig. 5. The presence of calcite, which has a typical rhomboidal structure, is seen in the SEM image, with a crystal size of 9.07 µm in diameter. The same calcite size was previously reported by Martos et al. (2010). One of the main mechanisms suggested to explain flux decline in a membrane system is filter cake formation (Shaefer et al., 2004). In this study, crystalline CaCO₃ particles were formed in the bulk of the emulsion and were deposited onto the membrane. If the emulsion pH was set to 5, solubility of CaCO₃ significantly increased with the shift from alkaline to acidic pH (Doberšek and Goričanec, 2014); consequently fewer crystals were formed in emulsion in comparison with the emulsion at original pH.

Therefore, the restoration of the initial flux after solely alkaline cleaning was not achieved (Fig. 6, Cleaning Strategy 2). Cleaning with acid was necessary to remove the calcite. HCl solution (pH = 3) was used and then flushed with distilled water. Calcite dissolved in acid (Doberšek and Goričanec, 2014) and the flux reached the initial value (99.9%) as seen in Fig. 6 (Cleaning Strategy 1).

Many combinations were tested, and those combinations with the first step alkaline cleaning plus the second step acid cleaning gave satisfactory results with flux recovery above 95%. Alkaline cleaning alone could not remove all organic and inorganic pollutants, which were present in waste emulsions (see Fig. 6, Cleaning Strategies 2 and 3): cleaning with ASG media (Cleaning Strategy 2) recovered the flux slightly better in comparison with US73 (Cleaning Strategy 3).

Based on these results, the best cleaning option was, firstly, sequential usage of ASG at 40°C for 1 h at TMP = 100 kPa, followed by flushing with distilled water (till pH of effluent was

**Figure 4.** Flux dependence of TMP for Sample N3 at pH = 5**Figure 5.** Typical scale particles of calcite**Figure 6.** Flux recovery (J_{rec}) at different combinations of cleaning agents (Cleaning Strategies 1 to 6, as described in Table 2)

neutralised) and, secondly, the usage of HCl (pH = 3) for 30 min followed by final flushing with distilled water. Comparison of flux recovery compared with the permeability of non-fouled membrane is shown in Fig. 7. Instead of HCl, sulphuric acid (pH = 3) was used after alkaline cleaning and flux recovery reached 98% of initial flux (Fig. 6, Cleaning Strategy 4).

Cleaning tests also showed that the temperature of alkaline media as well as acids should be at least 40°C. The strategy at higher TMP of 200 kPa was not successful. The recoveries were worse, well below 60%, and are not presented. The results are in agreement with those reported for another study (Silalahi

et al., 2009), where lower TMP during cleaning enabled better recoveries in comparison with higher TMP. Authors assumed that the foulant could be compacted into membrane pores at higher TMPs, e.g. 200 kPa.

Zeta potential of fouled membrane

The reduction in permeate flux of the ceramic membrane was observed during filtration of the oily wastewater. Membrane cleaning was not performed in order to foul the membrane, and eventually the flux stopped. The fouled membrane was removed from the filtration unit and, after careful cross-section slicing, placed in the streaming potential cell without rinsing, in order not to disturb any surface deposits on the membrane. Figure 8 demonstrates the zeta potential values of the fouled membrane as a function of pH. It can be observed that the zeta potential values of fouled membranes were just a little more negative when compared to those of the clean membrane (Fig. 2). The shape and slope of the fouled membrane curve differed from those of the clean membrane, in the region from pH 5 to 7, indicating contamination of the membrane surface due to fouling. However, the negative zeta potential remained for a great part very similar to that of the non-fouled membrane. Therefore, not only was the membrane surface fouled, but the membrane's pores were also narrowed, concurring with the literature (Faibish and Cohen, 2001).

Under acidic conditions (pH 5), the membrane gets a positive charge and electrical repulsion promotes the retention of positively charged cationic surfactant. Positive surface sites are thus not diminished, as seen from Fig. 2. In contrast, more positively charged solutes are attached to the surface and the IEP shifts to a higher pH = 7, as seen from Fig. 8. Most likely, negatively charged surfactants are attached to the membrane surface and influence its surface charge (Tanis-Kanbur et al., 2018). As the conductivity is very high, it is assumed that many charged ions are present in waste emulsion, and contribute to the membrane surface charge. The hydrophilic head of the surfactant attaches to the membrane surface and the hydrophobic tails are oriented towards the bulk phase (Lobo et al., 2006), increasing the hydrophobicity of the surface and reducing the aqueous flux through the membrane. It is known that a variety of non-ionic *silane surfactants* which are present in cutting oils have been synthesized recently (Cheng et al., 2005). The amino groups were present in silane which could contribute to the IEP shift.

At neutral pH, there is electrical repulsion between the anionic surfactants (and oils) and negatively charged membrane. Non-ionic surfactants fouled the membrane, reducing the flux due to adsorption, not only on the surface by cake formation, but also through narrowing of the membrane pores (Faibish and Cohen, 2001). The process is even faster at higher TMP. Therefore, the emulsion flux during UF was the same at TMP 250 kPa and 300 kPa, and fouling increased. As a result of attractive forces, ions may be strongly adsorbed onto the membrane surface as well as into the pores. In this way, the effective pore radius would be reduced. Since the membrane was positively charged in the acidic range, a higher quantity of oils may be attracted. As seen from Fig. 8, in the alkaline region the plateau in zeta potential is less pronounced due to the coverage of negatively charged surface groups, such as anionic surfactants (Matos et al., 2008). However, the coverage of non-polar organic molecules from emulsions, such as long-chain fatty acids (C20-C32), is possible (Pickl et al., 2011). Their presence in

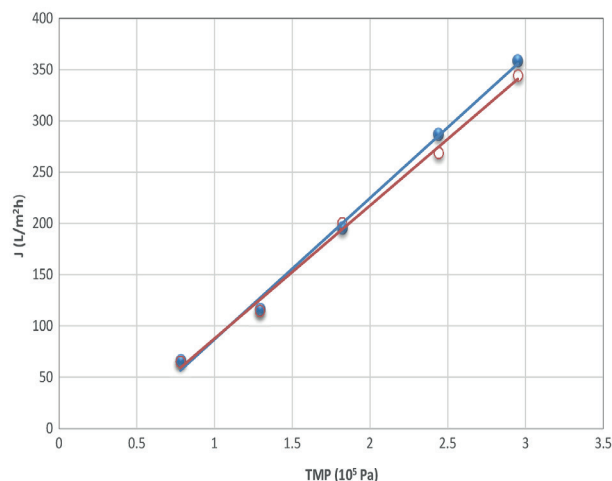


Figure 7. Flux recovery at combinations of ASG and HCl (o) in comparison with the virgin membrane water flux (shaded circles)

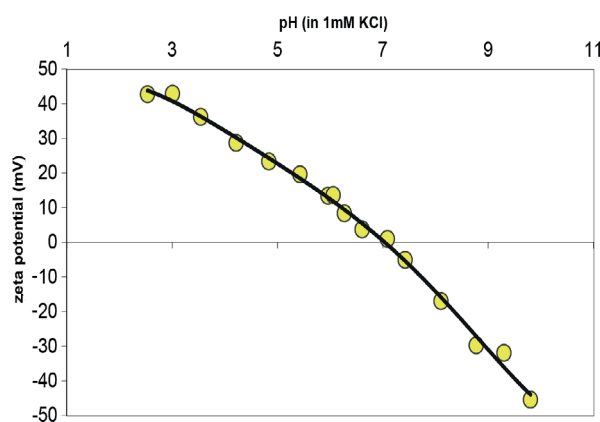


Figure 8. Zeta potential of fouled membrane in function of pH

emulsions was detected with GC-MS analysis (Fig. 9). Problems were experienced during quantitation of the GC-MS analysis; therefore, these results could only be used qualitatively.

CONCLUSIONS

The applicability of ultrafiltration was tested for COD reduction in waste emulsions. Real oil-in-water waste emulsions had already been pre-treated by an emulsion-breaking process. The results showed that the emulsion could be treated using UF. The highest efficiency was achieved with two different real waste emulsion samples at acidic pH, pH = 5, near isoelectric point. The COD decreased by 89% and 85% in the first and second emulsion samples, respectively. Although the removal efficiency is satisfactory, the values are still very high (above 2 000 mg/L O₂) and subsequent treatment in the wastewater treatment plant should be applied in order to reduce organic pollution. The efficiency of chemical cleaning of fouled ceramic membranes was examined. Low recoveries of fouled membrane were achieved using alkaline cleaning agent only, although most organic pollutants were removed. The fouled membranes could be fully restored in two steps: alkaline cleaning is necessary in the first step and acid cleaning to remove scale in the second step.

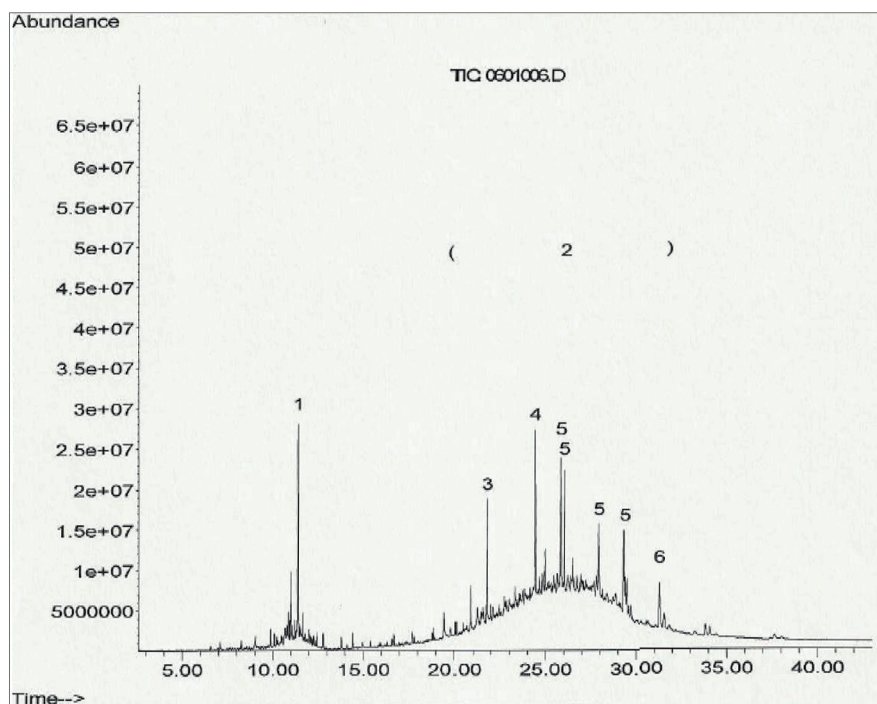


Figure 9. GC-MS analysis: 1 tri-decane, 2 fatty acids (C20-C32), 3 oleic acid, isopropyl ester, 4 non-identified, 5 fatty acid alkyl esters (FAAE)

ACKNOWLEDGEMENTS

The author would like to acknowledge the Slovenian Research Agency for their financial support (Project No. P2-0032).

REFERENCES

- BENITO JM, CAMBIELLA A, LOBO A, GUTIERREZ G, COCA J and PAZOS C (2010) Formulation, characterization and treatment of metalworking oil-in-water emulsions. *Clean Technol. Environ. Polic.* **12** 31–41. <https://doi.org/10.1007/s10098-009-0219-2>
- BENSADOK K, BELKACEM M and NAZZAL G (2007) Treatment of cutting oil/water emulsion by coupling coagulation and dissolved air flotation. *Desalination* **206** 440–448. <https://doi.org/10.1016/j.desal.2006.02.070>
- CAÑIZARES P, MARTÍNEZ F, JIMENEZ C, SAEZ C and RODRIGO M A (2008). Coagulation and electrocoagulation of oil-in-water emulsions. *J. Hazardous Mater.* **151** 44–51. <https://doi.org/10.1016/j.jhazmat.2007.05.043>
- CHENG C, PHIPPS D and ALKHADDAR R M (2005) Treatment of spent metalworking fluids. *Water Res.* **39** 4051–4063. <https://doi.org/10.1016/j.watres.2005.07.012>
- COCA J, GUTIÉRREZ G and BENITO J (2011) Treatment of oily wastewater. In: *Coca-Prados J and Gutiérrez-Cervelló G* (eds) *Water Purification and Management*. Springer Netherlands, Dordrecht. 1–55.
- DOBERŠEK D and GORIČANEČ D (2014) An experimentally evaluated magnetic device's efficiency for water-scale reduction on electric heaters. *Energy* **77** 271–278. <https://doi.org/10.1016/j.energy.2014.09.024>
- FAIBISH RS and COHEN Y (2001) Fouling resistant ceramic-supported polymer membranes for ultrafiltration of oil-in-water microemulsions. *J. Membr. Sci.* **185** (2) 129–143. [https://doi.org/10.1016/S0376-7388\(00\)00595-0](https://doi.org/10.1016/S0376-7388(00)00595-0)
- FU Y and CHUNG D D (2011) Coagulation of oil in water using sawdust, bentonite and calcium hydroxide to form floating sheets. *Appl. Clay Sci.* **53** 634–641. <https://doi.org/10.1016/j.clay.2011.05.014>
- GHANDASHTANI MB, ASHTIANI FZ, KARIMI M and FOULADITAJAR A (2015) A novel approach to fabricate high performance nano-SiO₂ embedded PES membranes for microfiltration of oil in water emulsions. *Appl. Surf. Sci.* **349** 393–402. <https://doi.org/10.1016/j.apsusc.2015.05.037>
- GUTIÉRREZ G, CAMBIELLA A, BENITO JM, PAZOS C and COCA J (2007) The effect of additives on the treatment of oil-in-water emulsions by vacuum evaporation. *J. Hazardous Mater.* **144** 649–654. <https://doi.org/10.1016/j.jhazmat.2007.01.090>
- HESAMPOUR M, KRZYZANIAK A and NYSTROM M (2008) Treatment of waste water from metal working by ultrafiltration, considering the effects of operating conditions. *Desalination* **222** 212–221. <https://doi.org/10.1016/j.desal.2007.01.155>
- HU X G, BEKASSY-MOLNAR E and VATAI G (2002) Study of ultrafiltration behaviour of emulsified metalworking fluid. *Desalination* **149** 191–197. [https://doi.org/10.1016/S0011-9164\(02\)00758-0](https://doi.org/10.1016/S0011-9164(02)00758-0)
- HUA F L, TSANG Y F, WANG Y J, CHAN SY, CHUAH and SIN S N (2007) Performance study of ceramic microfiltration membrane for oily wastewater treatment. *Chem. Eng. J.* **128** (2–3) 169–175. <https://doi.org/10.1016/j.cej.2006.10.017>
- JU J, WANG T and WANG Q (2015) Superhydrophilic and underwater superoleophobic PVDF membranes via plasma-induced surface PEGDA for effective separation of oil-in-water emulsions. *Colloids Surf. A: Physicochem. Eng. Aspects* **481** 151–157. <https://doi.org/10.1016/j.colsurfa.2015.01.041>
- KRIZAN J, PETRINIC I, GORSEK A and SIMONIC M (2014) Ultrafiltration of oil-in-water emulsion by using ceramic membrane: Taguchi experimental design approach. *Cent. Eur. J. Chem.* **12** (2) 242–249. <https://doi.org/10.2478/s11532-013-0373-6>
- LIN HF (2006) Ceramic membrane technology applied to oily wastewater separation. PhD dissertation, The Hong Kong Polytechnic University, China.
- LIN SH and LAN WJ (1998) Treatment of waste oil/water emulsion by ultrafiltration and ion exchange. *Water Res.* **32** (9) 2680–2688. [https://doi.org/10.1016/S0043-1354\(98\)00035-9](https://doi.org/10.1016/S0043-1354(98)00035-9)
- LOBO A, CAMBIELLA A, BENITO J M, PAZOS C and COCA J (2006) Ultrafiltration of oil-in-water emulsions with ceramic membranes: Influence of pH and crossflow velocity. *J. Membr. Sci.* **278** (1–2)

- 328–334. <https://doi.org/10.1016/j.memsci.2005.11.016>
- MATOS M, LOBO A, FERNANDEZ E, BENITO JM, PAZOS C and COCA J (2008) Recycling of oily ultrafiltration permeates to reformulate O/W emulsions. *Colloids Surf. A: Physicochem. Eng. Aspects* **331** (1–2) 8–15. <https://doi.org/10.1016/j.colsurfa.2008.06.004>
- MARTOS C, COTO B, PEÑA JL, RODRÍGUEZ R, MERINO-GARCIA D and PASTOR G (2010) Effect of precipitation procedure and detection technique on particle size distribution of CaCO₃. *J. Crystal Growth* **312** (19) 2756–2763. <https://doi.org/10.1016/j.jcrysgro.2010.06.006>
- NORDIN AK and JÖNSSON (2010) Influence of module configuration on total economics during ultrafiltration at high concentrations. *Chem. Eng. Res. Design* **88** 1555–1562. <https://doi.org/10.1016/j.cherd.2010.04.003>
- Official Gazette of Republic Slovenia (2005) Decree on Emission of Material and Heat by Wastewater Discharge from Cutting Oil Production, No47/2005, Ljubljana, Slovenia.
- PICKL KE, ADAMEK V, GORGES R and SINNER FM (2011) Headspace-SPME-GC/MS as a simple cleanup tool for sensitive 2,6-diisopropylphenol analysis from lipid emulsions ad adaptable to other matrices. *J. Pharm. Biometric. Anal.* **55** (5) 1231–1236. <https://doi.org/10.1016/j.jpba.2011.03.019>
- REZVANPOUR A, ROOSTAAZAD R, HESAMPOUR M, NYSTRÖM M and GHOTBI C (2009) Effective factors in the treatment of kerosene-water emulsion by using UF membranes. *J. Hazardous Mater.* **161** 1216–1224. <https://doi.org/10.1016/j.jhazmat.2008.04.074>
- SCHAEFER AI, ANDRITSOS N, KARABELAS AJ, HOEK EMV, SCHNEIDER R and HYSTROM M (2004) Chapter 8. Fouling in nanofiltration. In: Schäfer AI, Waite TD and Fane AG (eds) *Nanofiltration-Principle and Application*. Elsevier, New York.
- SILALAH SHD and LEIKNES T (2009) Cleaning strategies in ceramic microfiltration membranes fouled by oil and particulate matter in produced water. *Desalination* **236** (1–3) 160–169. <https://doi.org/10.1016/j.desal.2007.10.063>
- TANIS-KANBUR MB, VELIOGLU S, TANUDJAJA HJ, HU X and CHEW JW (2018) Understanding membrane fouling by oil-in-water emulsion via experiments and molecular dynamics simulation. *J. Membr. Sci.* **566** 140–150. <https://www.sciencedirect.com/science/article/pii/S037673881831562X?via%3Dihub>
- ZHOU J, CHANG Q, WANG Y, WANG J and MENG G (2010) Separation of stable oil-water emulsion by the hydrophilic nano-sized ZrO₂ modified Al₂O₃ microfiltration membrane. *Sep. Purif. Technol.* **75** (3) 243–228. <https://doi.org/10.1016/j.seppur.2010.08.008>
- WAKILY H, MEHRALI M and METSELAAR H (2010) Preparation of Homogeneous Dense Composite of Zirconia and Alumina (ZTA) using Colloidal Filtration. *World Acad. Sci. Eng. Technol.* **70** 140–145.

Removal of lead in water using activated carbon prepared from *Acacia catechu*

K Lakshmikandhan¹ and A Ramadevi^{2*}

¹Department of Chemistry, CMS College of Engineering and Technology, Coimbatore, India

²Department of Chemistry, Alagappa Chettiar Government College of Engineering and Technology Karaikudi, India

ABSTRACT

The adsorption of Pb (II) on bicarbonate-treated *Acacia catechu* carbon (BTACC) and commercial activated carbon (CAC) was investigated to assess the possible use of this adsorbent for Pb (II) removal from aqueous solutions. The results obtained from batch studies showed that 98% Pb (II) adsorption for BTACC occurs within 4 h of contact time, within a pH range of 4–10 and a carbon dosage of 100 mg with initial Pb (II) concentration of 10 mg·L⁻¹, whereas for CAC 52% Pb (II) adsorption occurred within 5 h of contact time and a narrow range of pH 5 and carbon dosage of 100 mg. Adsorption followed pseudo-second-order kinetics for both carbon sources, with the mechanism of adsorption being intra-particle diffusion and film-diffusion. The best fitting adsorption isotherm for both BTACC and CAC were the Langmuir, Freundlich and Temkin models. Surface characteristics were studied using FT-IR, SEM, EDX.

Keywords: *Acacia catechu*, commercial activated carbon, adsorption isotherms, kinetics

INTRODUCTION

Water contains impurities of various kinds which are dissolved as well as suspended. Heavy metals are major toxic pollutants with serious health effects on humans (Inglezakis et al., 2003; Demirbas, 2008). The toxicity of heavy metals depends on the concentration of metal ions, and the nature of the organism with which it interacts. The heavy metals are the most toxic metals of widespread use in industries such as tanning, electroplating, electronic equipment, manufacturing and chemical processing plants. Lead has been recognized for centuries as a cumulative poison (Hua et al., 2012; Kadirvelu et al., 2001; WHO, 1977).

Health studies done in Poland have linked elevated levels of lead in the environment with retardation and learning disabilities of children (Groffman et al., 1992). Acute lead poisoning in humans causes severe damage to kidneys, the reproductive system, the liver, the brain and central nervous system. The neurotoxicity of lead is well known but the exact mechanisms of its toxicity are not yet understood. Disturbances in glutamate homeostasis of neural tissue and interactions of lead with calcium metabolism have been considered as potential mechanism of neurotoxicity. (Raunio et al., 2001).

Removal of Pb (II) by adsorption using treated granular activated carbon in both batch and column studies has been studied (Goel et al., 2005). Activated carbon developed from tamarind wood by zinc chloride activation was examined for the removal of Pb (II) from wastewater (Jyotikusum et al., 2009). Activated carbon prepared from marine green algae was used for adsorption of Pb (II) ions (Suresh Jeyakumar et al., 2014). Treatment of lead-contaminated water using activated carbon adsorbent from locally available papaya peel bio-waste has also been investigated (Sahar Abbaszadeh et al., 2016).

In the present investigation, out of various non-conventional adsorbents, modified *Acacia catechu* carbon is studied for its adsorption capacity for lead (II) removal from aqueous solution. This work reports the results of batch

studies on the removal of lead (II) from aqueous solution by adsorption, using bicarbonate-treated *Acacia catechu* carbon (BTACC) as an adsorbent. Various parameters, such as equilibrium time, pH and dosage of adsorbent, and adsorption isotherms were studied. Commercial activated carbon (CAC) procured from the market was used for evaluation purposes.

MATERIALS AND METHODS

Preparation of bicarbonate-treated *Acacia catechu* carbon (BTACC)

Acacia catechu seeds were collected, washed with distilled water and dried at 110°C. Seeds were then broken mechanically into small particles and sieved to 20–50 ASTM mesh size. Then the material was treated with concentrated sulphuric acid with a weight ratio of 1:1 and kept in an air oven at 150 ± 5° C for 24 h. The carbonized material was washed well with distilled water to remove the free acid and dried at 105 ± 5°C. Then it was soaked in 1% sodium bicarbonate solution until the effervescence ceased, and finally soaked in sodium bicarbonate solution for 24 h to remove any residual acid. The material was then washed with distilled water, and dried at 105 ± 5°C (BTACC). Preliminary studies were carried out with raw *Acacia catechu* seed, sulphuric acid treated *Acacia catechu* seed carbon and sulphuric acid treated *Acacia catechu* seed modified by 1% sodium bicarbonate solution (BTACC) for Pb (II) removal. Based on removal efficiency, BTACC was chosen for further studies. The characteristics of BTACC and commercial activated carbon (CAC) were investigated and are summarized in Table 1. Subsequent experiments were carried out with BTACC and the performance of the carbon was evaluated by using CAC obtained from SD Fine Chemicals, considering the same particle size of 20–50 ASTM.

Batch experiment

A stock solution of Pb (II) was prepared by dissolving 0.3997 g of anhydrous lead nitrate in 100 mL distilled water containing 0.1 mL of concentrated nitric acid (to prevent hydrolysis) and

*Corresponding author, email: ramadeviramanujam61@gmail.com
Received 11 June 2018; accepted in revised form 4 June 2019

diluting it to 250 mL. This solution was diluted further to 100 mL of Pb (II) solution of desired concentration adjusted to a desired pH in reagent bottles of 300 mL capacity. A known amount of BTACC and CAC were added and pH was adjusted using dilute hydrochloric acid or sodium hydroxide solutions. All chemicals used were of analytical grade and were obtained from Ranbaxy, BDH. The solutions were agitated and the filtrate was analysed for lead content (APHA, AWWA, 1973). Maintaining the dosage of carbon at a constant level, the adsorption isotherm for Pb (II) with different initial

concentrations was studied. For pH effect, 10 mg·L⁻¹ of Pb (II) for BTACC and CAC with a dosage of 100 mg·100 mL⁻¹ were used. In order to correct for any adsorption of Pb (II) due to the container, control experiments were carried out without adsorbent and there was negligible adsorption by the container wall.

RESULTS AND DISCUSSION

Examination of carbon characteristics in Table 1 shows that the BTACC has higher moisture content than CAC. Higher moisture content of BTACC suggests that the acid treatment made the carbon more porous and BTACC has lower ash content than CAC, which indicates more carbon content. The surface area is found to be greater for BTACC than CAC.

Surface characterization of activated carbon

Fourier transform infrared spectroscopic analysis (FT-IR)

Fourier transform infrared spectroscopy (FT-IR) studies were used to identify the functional groups present on the surface of the adsorbent. FT-IR spectra of BTACC and CAC are shown in Fig. 1a–d. This shows the presence of poly functional groups. The strong absorption peak at 3 445 cm⁻¹, is due to the -OH stretching vibration due to inter- and intra-molecular

Table 1. Characteristics of bicarbonate-treated *Acacia catechu* carbon (BTACC) and commercial activated carbon (CAC)

Sl. No.	Control test	BTACC	CAC
1	Bulk density (g·L ⁻¹)	0.7656	0.7581
2	Moisture (%)	9.14	7.32
3	Ash (%)	2.15	20.94
4	Solubility in water (%)	1.21	0.453
5	Solubility in 0.25N HCl (%)	5.19	2.03
6	pH	4.84	8.35
7	Decolorizing power (mg·g ⁻¹)	9.0	4.5
8	Phenol number	68	51
9	Ion exchange capacity (m equiv·g ⁻¹)	0.6024	Nil
10	Surface area (m ² ·g ⁻¹)	326	205
11	Iron (%)	6.21	4.15
12	Porosity (%)	2.1	1.95

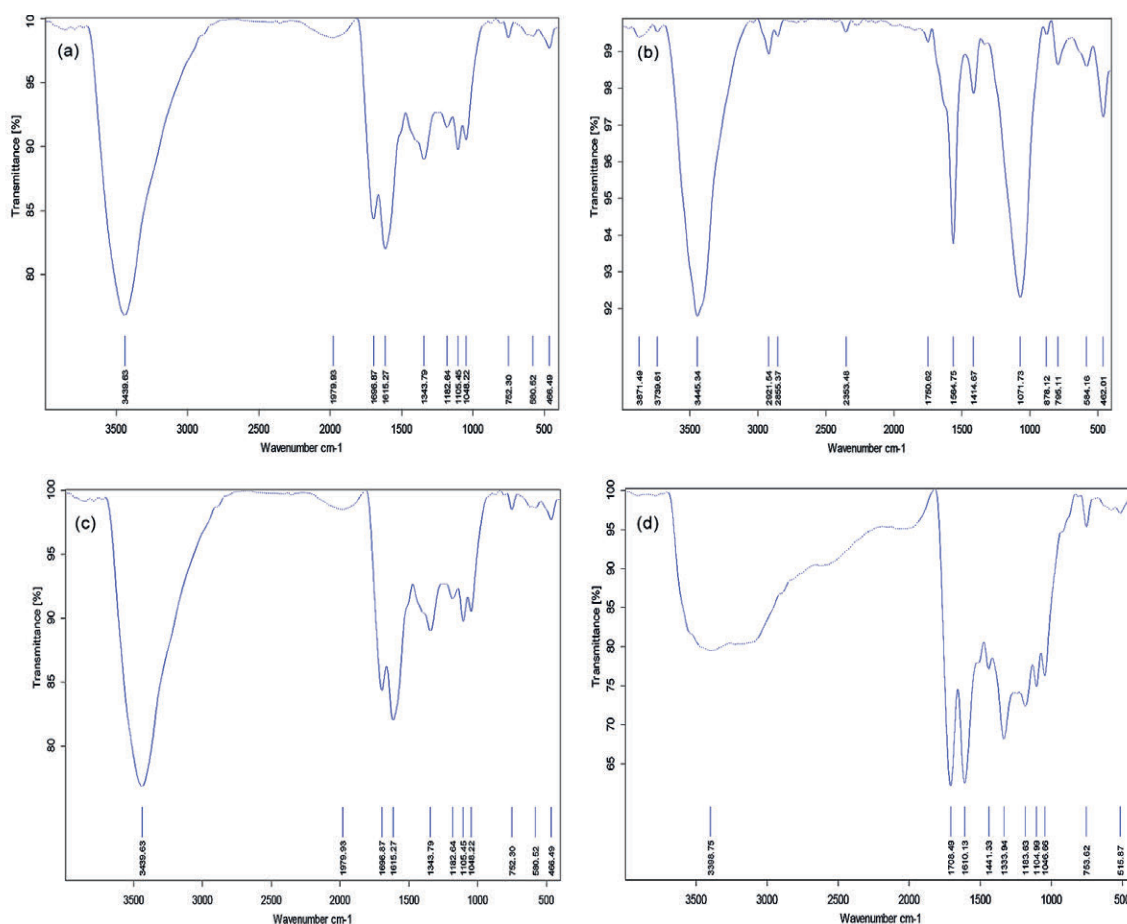


Figure 1. (a, b) FT-IR spectra of before and after adsorption of Pb (II) ions on BTACC; (c, d) FT-IR spectra of before and after adsorption of Pb (II) ions on CAC

hydrogen bonding of alcohols, phenols and carboxylic acids. The peaks at 2928 cm^{-1} , 2362 cm^{-1} and 1698 cm^{-1} are due to the C-H stretching, $\text{C}\equiv\text{C}$ stretching and -CO stretching vibration of ether. The presence of sulphonic acid groups is confirmed by the peak at 1343 cm^{-1} . From the observations it is evident that some of the peaks shift or become weak indicating the incorporation of heavy metal ion Pb (II) within the adsorbent through the interaction of the active functional group after adsorption.

Scanning electron microscopy (SEM)

SEM was used to study the morphology and surface characteristics of the adsorbent material (Tharanitharan et al., 2009). Adsorption capacity depends on the chemical composition of the adsorbent. The pore size, shape and volume also play an important role in the adsorption process. The SEM images of the adsorbent BTACC and CAC before and after adsorption of Pb (II) ions are given in Fig. 2a–d. This image shows the presence of isolated pores and uneven cavities before and after adsorption. The irregular and porous surface

activated carbon could be observed. On the basis of this fact, it can be concluded that BTACC presents an adequate morphology for Pb (II) adsorption (Deepak et al., 2013).

Energy dispersive X-ray analysis (EDX)

Energy dispersive X-ray analysis (EDX) technique is used for elemental analysis or chemical characterization of a sample in conjunction with scanning electron microscopy (SEM). To resolve the elemental content, the electron-beam strikes the surface of a conducting sample (SEM). The EDX spectra of the Pb (II) ion adsorbed on BTACC and CAC show peaks for metal ions in addition to the other cations, confirming the adsorption of Pb (II) ions on the surface of the adsorbent (Fig. 3 a–d).

Effect of agitation time

The effect of agitation time on the removal of Pb (II) by BTACC and CAC is shown in Fig. 4. Percentage removal increases with time and attains equilibrium at 3 h for BTACC and 4 h for CAC for an initial concentration of $10\text{ mg}\cdot\text{L}^{-1}$ of Pb (II). However, 4 h

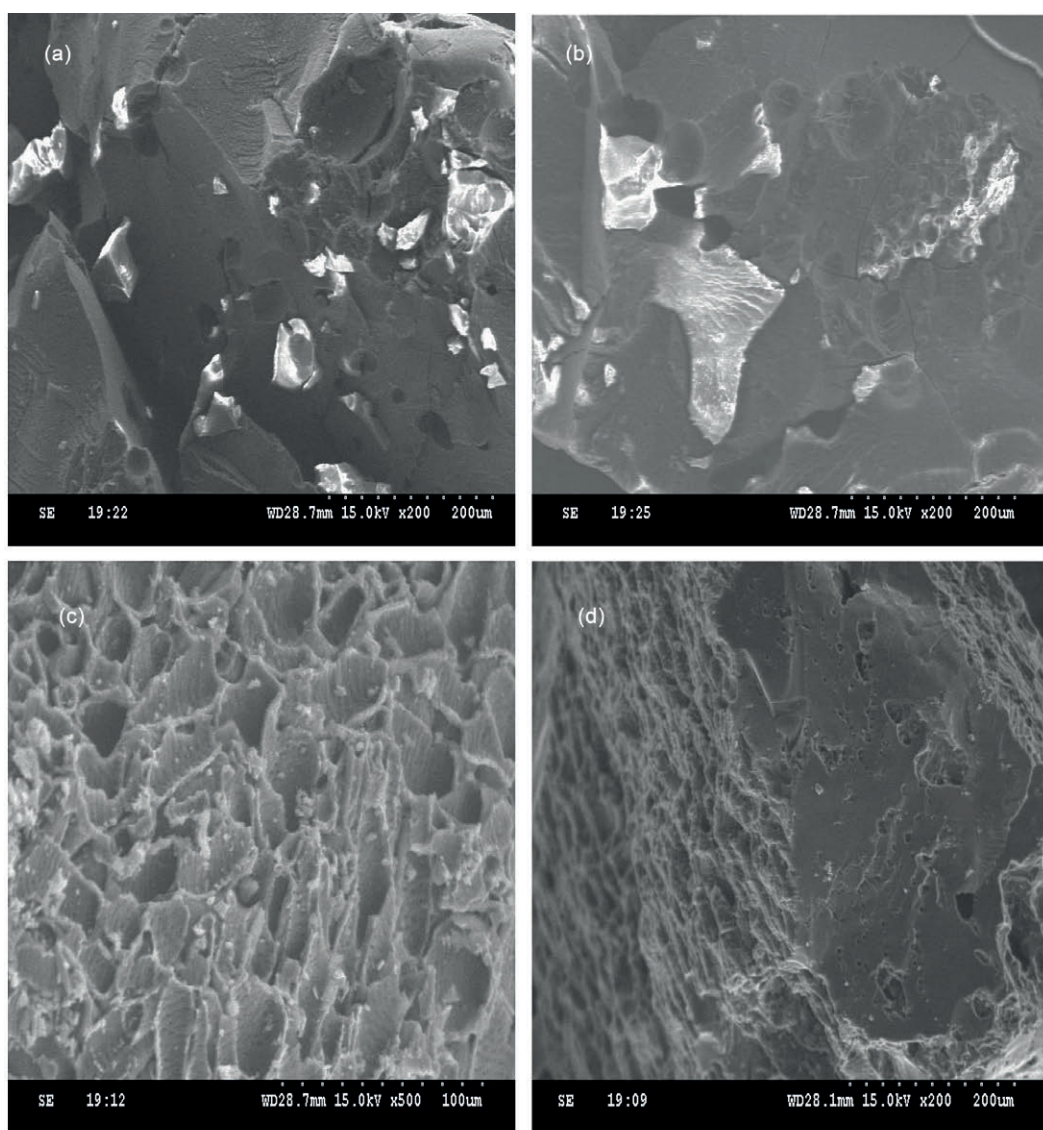


Figure 2. (a, b) SEM images before and after adsorption of Pb (II) ions on BTACC; (c, d) SEM images before and after adsorption of Pb (II) ions on CAC

and 5 h shaking time was maintained for BTACC and CAC, respectively, for further studies. It was found that BTACC showed 98% lead (II) removal whereas CAC showed 52% removal for a carbon dosage of 100 mg·100 mL⁻¹ of Pb (II) ion solution of initial concentration of 10 mg·L⁻¹.

Effect of carbon dosage

Figure 5 represents the removal of Pb (II) as a function of carbon dosage for both BTACC and CAC. It was shown that a minimum carbon dosage of 100 mg of BTACC was required for 98% removal of Pb (II) from a 10 mg·L⁻¹ lead solution. However, a maximum removal of 52% was observed for CAC with a carbon dosage of 100 mg for the same concentration of Pb (II). This shows that BTACC was nearly 2 times more efficient than CAC for Pb (II) removal.

Effect of pH

Figure 6 shows the effect of initial pH on the removal of Pb (II) by BTACC and CAC. It is clear that BTACC is effective for the quantitative removal of Pb (II) over the pH range of 4.0–10.0. However, for CAC the quantitative removal was only at a narrow pH of 5. The influence of pH on lead removal can likely be explained as follows. At lower pH, a higher concentration of hydrogen ions present in the mixture competes with the positively charged metal ion for the adsorption sites, resulting in the reduced uptake of metal ions. As the pH increases, the concentration of Pb (II) ions remains constant and therefore the uptake of metal ions can be explained as H⁺-Pb²⁺ exchange

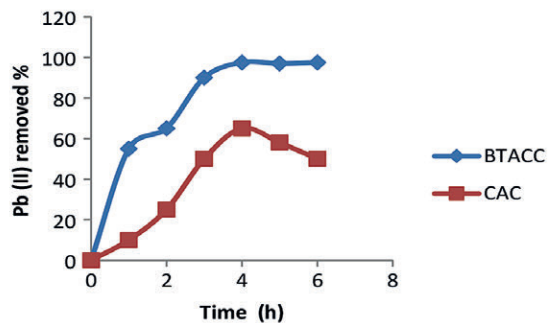


Figure 4. Effect of contact time for the removal of Pb (II) by BTACC and CAC (pH = 5.0)

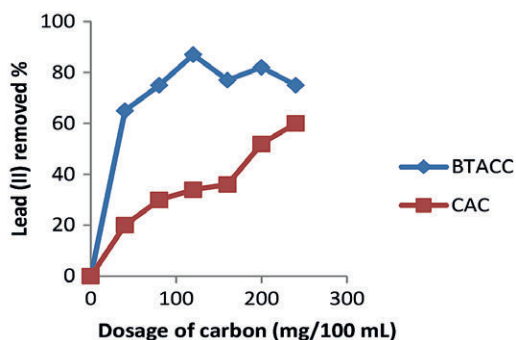


Figure 5. Effect of carbon dosage on the removal of Pb (II) by BTACC and CAC (pH = 5.0), equilibrium time = 3 h, concentration of Pb (II) = 10 mg·L⁻¹ for CAC

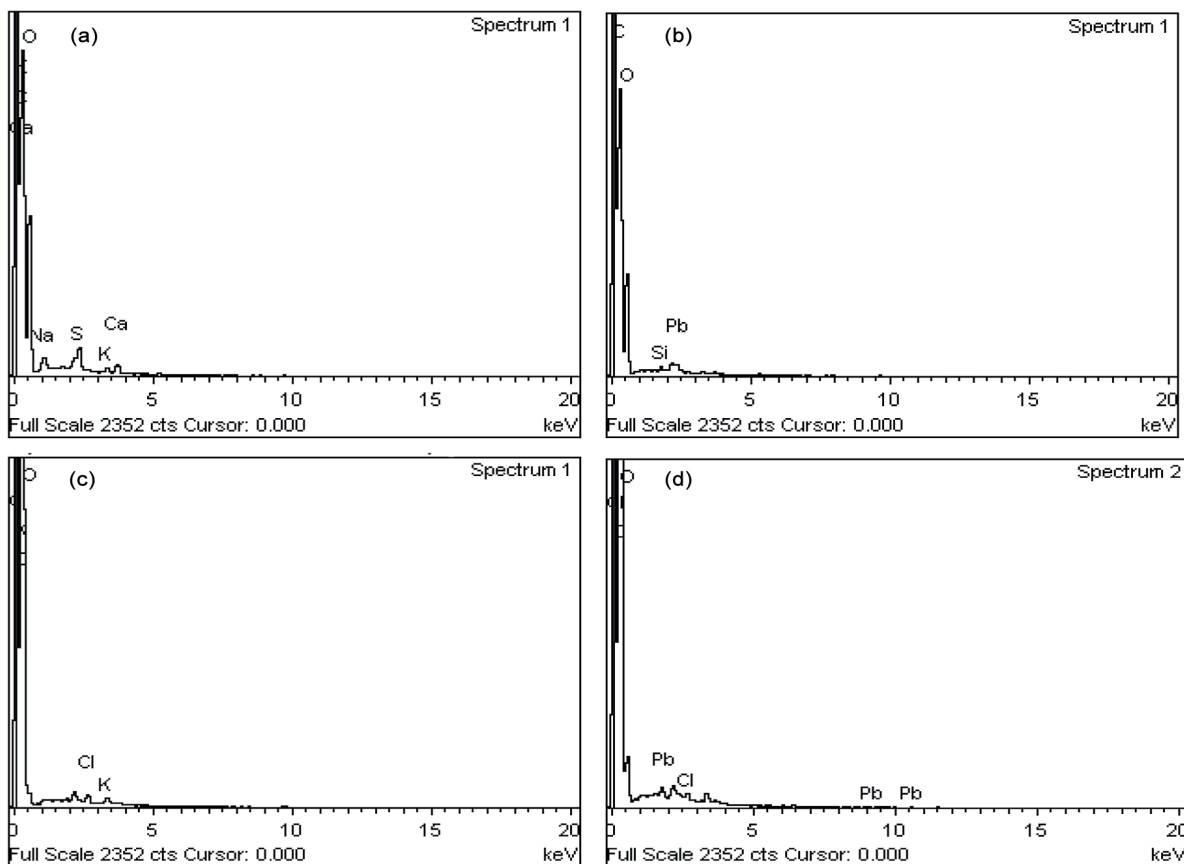


Figure 3. (a, b) Edx pattern of before and after adsorption of Pb (II) ions on BTACC; (c, d) Edx pattern of before and after adsorption of Pb (II) ions on CAC

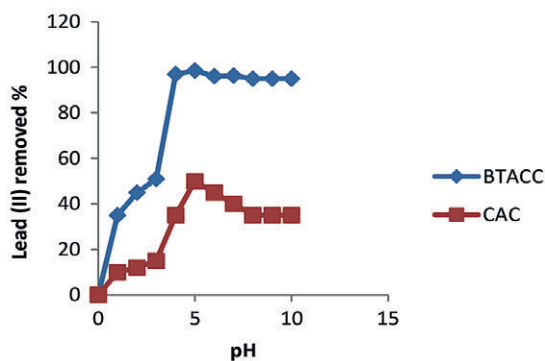
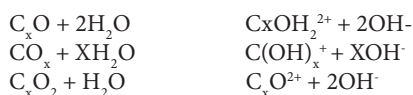


Figure 6. Effect of pH on the removal of Pb (II) by BTACC and CAC

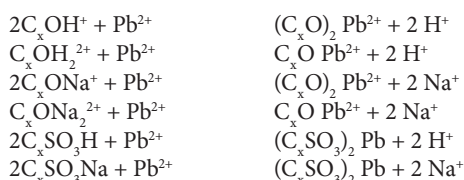
reaction. The major mechanism of adsorption of Pb^{2+} seems to be ion exchange. Furthermore, a pure carbon surface is considered to be non-polar, but in actual practice some carbon-oxygen complexes, such as C_xO , CO_x and C_xO_2 are usually present which render the surface slightly polar. The interaction of the above groups with aqueous phase may lead to the following hydrolytic reactions:



Since BTACC was prepared upon treatment with sulphuric acid and sodium bicarbonate, groups such as C_xONa^{2+} , C_xONa^+ , C_xSO_3H and C_xSO_3Na are also assumed to be present. Na^+ ions in the above groups are also exchanged with H^+ in the medium as follows:



Under these conditions lead ions are expected to exchange as follows:



Adsorption isotherms

Results for application of various isotherm models for adsorption by BTACC and CAC are given in Table 3. Reported adsorption capacities of various materials for Pb (II) are given in Table 4.

The Langmuir equation was applied for adsorption equilibrium for both BTACC and CAC:

$$C_e/q_e = 1/Q_0b + C_e/Q_0 \quad (1)$$

where C_e is equilibrium concentration ($mg \cdot L^{-1}$), q_e is amount adsorbed at equilibrium ($mg \cdot g^{-1}$) and Q_0 and b are Langmuir constants related to adsorption capacity and energy of adsorption, respectively (Hall et al., 1966; Webber et al., 1974). The linear plots of C_e/q_e vs. C_e showed that the experimental data fitted well and adsorption obeys the Langmuir model for both BTACC and CAC (Fig. 7).

The Freundlich isotherm is represented by Eq. 2:

$$\log_{10}(x/m) = \log K + 1/n \log_{10} C_e \quad (2)$$

where C_e is the equilibrium concentration ($mg \cdot L^{-1}$) and x/m is the amount adsorbed per unit weight of BTACC and CAC ($mg \cdot g^{-1}$). Plots of $\log(x/m)$ vs. $\log C_e$ are linear for both BTACC and CAC (Fig. 8). The straight-line nature of the plots indicates that the process followed was of the Freundlich adsorption type; K and n values for both carbon sources were calculated from the intercepts and slopes, respectively, and are shown in Table 2. The values of $1 < n < 10$ show favourable adsorption of Pb (II) on both BTACC and CAC (Guadalupe et al., 2008). The K values for BTACC and CAC are very much

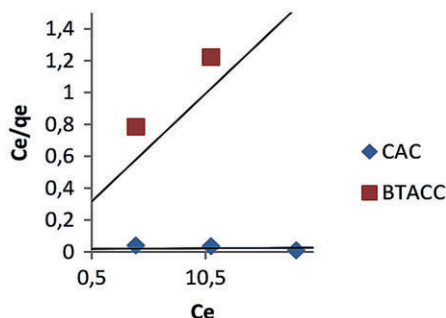


Figure 7. Langmuir adsorption isotherm for Pb (II) with BTACC and CAC system

Table 2. Comparison of K and n values in distilled water and tap water (Freundlich adsorption isotherm)

Carbon	Distilled water		Tap water	
	K	n	K	n
BTACC	31.62	1.15	72.44	0.7518
CAC	0.3565	2.0	19.05	2.82

Table 3. Isothermal parameters for Pb (II) removal from aqueous solutions

Isotherm model	Parameter	BTACC	CAC
		K ($mg \cdot g^{-1}$)	31.62
Freundlich	n ($mg \cdot L^{-1}$)	1.15	2.0
	R^2	0.071	0.899
Langmuir	Q_0 ($mg \cdot g^{-1}$)	80	18.18
	b ($mg \cdot L^{-1}$)	3.91	0.12
Temkin	R^2	0.681	0.712
	A ($L \cdot g^{-1}$)	43.63	1.69
	B	14.26	5.278
	R^2	0.818	0.493

Table 4. Adsorption capacities reported in the literature for Pb (II) adsorption

Agricultural waste	Adsorption model	Adsorption capacity	References
Cotton waste	Langmuir	44.67 $mg \cdot g^{-1}$	Riaz et al., 2009
Maize leaf	Freundlich	4.356 $mg \cdot g^{-1}$	Babarinde et al., 2006
Barley straws	Langmuir	23.20 $mg \cdot g^{-1}$	Pehlivan et al., 2008
Tamarind nut	Langmuir	33.3 $mg \cdot g^{-1}$	Ramadevi et al., 2005
<i>Acacia catechu</i>	Langmuir	80 $mg \cdot g^{-1}$	Present study
	Freundlich	31.62 $mg \cdot g^{-1}$	

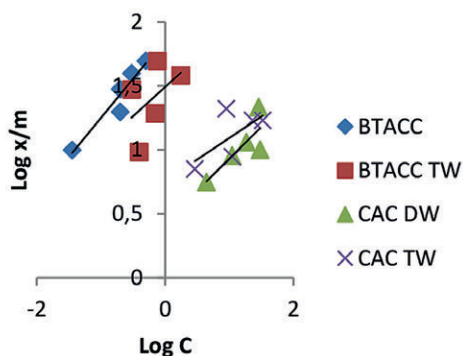


Figure 8. Freundlich adsorption isotherm for Pb (II) in distilled water (DW) and tap water (TW) for BTACC and CAC system

lower in distilled water than tap water because calcium and magnesium ions which are present in tap water compete for adsorption with lead.

Temkin isotherm

Temkin considered the effects of some implied adsorbent-adsorbate interactions on the adsorption isotherms, suggesting that these interactions cause the heat of adsorption for all of the molecules in the layer to decrease linearly with the coverage. The Temkin isotherm has been applied in the following form (Mousavi et al., 2011):

$$q_e = B_T \ln K_T + B_T \ln C_e \quad (3)$$

where $B_T = RT/b_T$, T (K) is the absolute temperature, R is the universal gas constant ($8.314 \text{ J}\cdot\text{mol}^{-1} \text{ K}^{-1}$), K_T ($\text{L}\cdot\text{mg}^{-1}$) is the equilibrium binding constant that corresponds to the maximum binding energy, b_T is the variation in the adsorption energy ($\text{kJ}\cdot\text{mol}^{-1}$), and B_T is a Temkin constant related to the heat of adsorption ($\text{kJ}\cdot\text{mol}^{-1}$). The Temkin constants can be derived by plotting q_e versus $\ln C_e$. Ho et al. (1996) indicated that the typical range for the binding energy during ion exchange mechanism is $8\text{--}16 \text{ kJ}\cdot\text{mol}^{-1}$. The heat of adsorption B_T value for BTACC and CAC is 14.26 and $5.278 \text{ kJ}\cdot\text{mol}^{-1}$, respectively. It is confirmed that BTACC has the binding energy within the limit $8\text{--}16 \text{ kJ}\cdot\text{mol}^{-1}$ and it was found that the ion exchange mechanism took place while it further confirmed the ion exchange capacity value of BTACC. CAC, however, does not fall within the limit of binding energy ($8\text{--}16 \text{ kJ}\cdot\text{mol}^{-1}$).

Adsorption studies

In order to explain the adsorption mechanism and rate-controlling steps, the kinetic adsorption data were modelled using pseudo-first order and pseudo-second order kinetic equations.

Pseudo-first order kinetics

The pseudo-first order reaction equations of Lagergren (1898) have been widely used for the adsorption of liquid/solid systems on the basis of solid capacity. The linear form is generally expressed as:

$$\log (q_e - q_t) = \log q_e - K_1 t / 2.303 \quad (4)$$

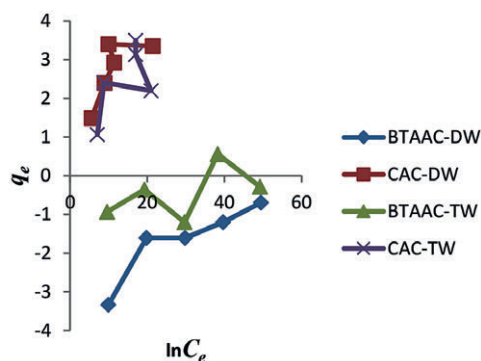


Figure 9. Temkin adsorption isotherm for Pb (II) in distilled water (DW) and tap water (TW) for BTACC and CAC system

where q_e ($\text{mg}\cdot\text{g}^{-1}$) and q_t ($\text{mg}\cdot\text{g}^{-1}$) are the amount of adsorption at equilibrium at time t (min); k_1 (min^{-1}) is the rate constant of the pseudo-first-order adsorption process. The values of k_1 , R^2 and q_e at different concentrations were obtained by plotting $\log (q_e - q_t)$ versus t . The constants were determined experimentally by plotting of $\log (q_e - q_t)$ vs. t (Fig 10a, b).

Pseudo-second-order kinetics

The pseudo-second-order model obtained from Ho et al. (1999) is based on the assumption that the adsorption follows the second-order rate equation. The linear form can be written as:

$$t / qt = 1 / K_2 q_e^2 + t / qt \quad (5)$$

where K_2 is the rate constant of adsorption.

To evaluate the applicability of kinetic data relative deviation (P %) was calculated using the equation:

$$P \% = 1/N \times 100 (q_e \text{ exp} - q_e \text{ cal} / q_e \text{ cal}) \quad (6)$$

where $q_e \text{ (exp)}$ and $q_e \text{ (cal)}$ are the experimental and calculated value of Pb (II) adsorbed on the adsorbents, N is the number of measurements made. The lower the percentage deviation (P%), the better is the fit. When P (%) is less than 5, the fit is considered to be excellent (Ugurlu et al., 2009). The correlation coefficient (R^2) for the pseudo-second-order model is much closer to unity. The calculated q_e value was much closer to the experimental q_e value. All kinetic parameters and correlation coefficients are listed in Table 5. Also the percentage relative deviation (P%) was found to be less than 5% in the case of pseudo-second-order. These values predict that the adsorption kinetics of Pb (II) ions onto the BTACC and CAC is mainly based on the pseudo-second-order equation. So the overall rate of Pb (II) ion adsorption may be controlled by the chemical process.

Adsorption mechanism

The pseudo-first-order and pseudo-second-order kinetic models were not able to explain the diffusion mechanism and also the rate-determining step for the adsorption of Pb (II) ions onto the BTACC and CAC. This is explained by the intra-particle diffusion model. For a solid-liquid sorption process, the solute transfer is usually characterized by either external mass transfer or intra-particle diffusion or both. During the

adsorption of Pb (II) ions onto BTACC and CAC, the following three consecutive steps were involved (Acharya et al., 2009):

- The movement of adsorbate molecules from the bulk solution to the external surface of the adsorbent (film diffusion)
- Adsorbate molecules move to the interior part of the adsorbent particles (intra-particle diffusion)
- Sorption of the solute on the interior surface of the pores and capillary spaces of adsorbent (sorption)

The effects of contact time data were analysed by the Weber et al. (1963) intra-particle diffusion model to elucidate the diffusion mechanism; the model is expressed as:

$$q_t = k_d t^{1/2} + I \quad (6)$$

where I ($\text{mg}\cdot\text{g}^{-1}$) is the intercept and k_d is the intra-particle diffusion rate constant ($\text{mg}\cdot\text{g}^{-1}\cdot\text{min}^{1/2}$), which can be evaluated from the slope of the linear plot of q_t versus $t^{1/2}$ as shown in Figs 10e and 10f. The dual nature of the curve was obtained due to the varying extent of adsorption in the initial and final stages of the adsorption

experiment. This can be attributed to the fact that in the initial stages the sorption was due to the boundary layer diffusion effect whereas in the later stages (linear portion of the curve) it was due to the intra-particle diffusion effect. The intercept of the plot shows the boundary layer effect. The larger the intercept the greater will be the contribution of the surface sorption in the rate-controlling step. If the regression in the plot of q_t versus $t^{1/2}$ is linear and passes through the origin, then intra-particle diffusion is the sole rate-limiting step. However, in the present study, the plots were not linear and did not pass through the origin. The results show that intra-particle diffusion was not only the rate-limiting step, but also the rate-controlling step of adsorption.

Film and pore diffusion coefficient

For the adsorption of heavy metals on carbon surfaces, Michaelson et al., (1975) proposed film diffusion to be the rate-determining process, and the value of film diffusion coefficient (D_f) to be between 10^{-6} and 10^{-8} $\text{cm}^2\cdot\text{s}^{-1}$. If the pore diffusion were

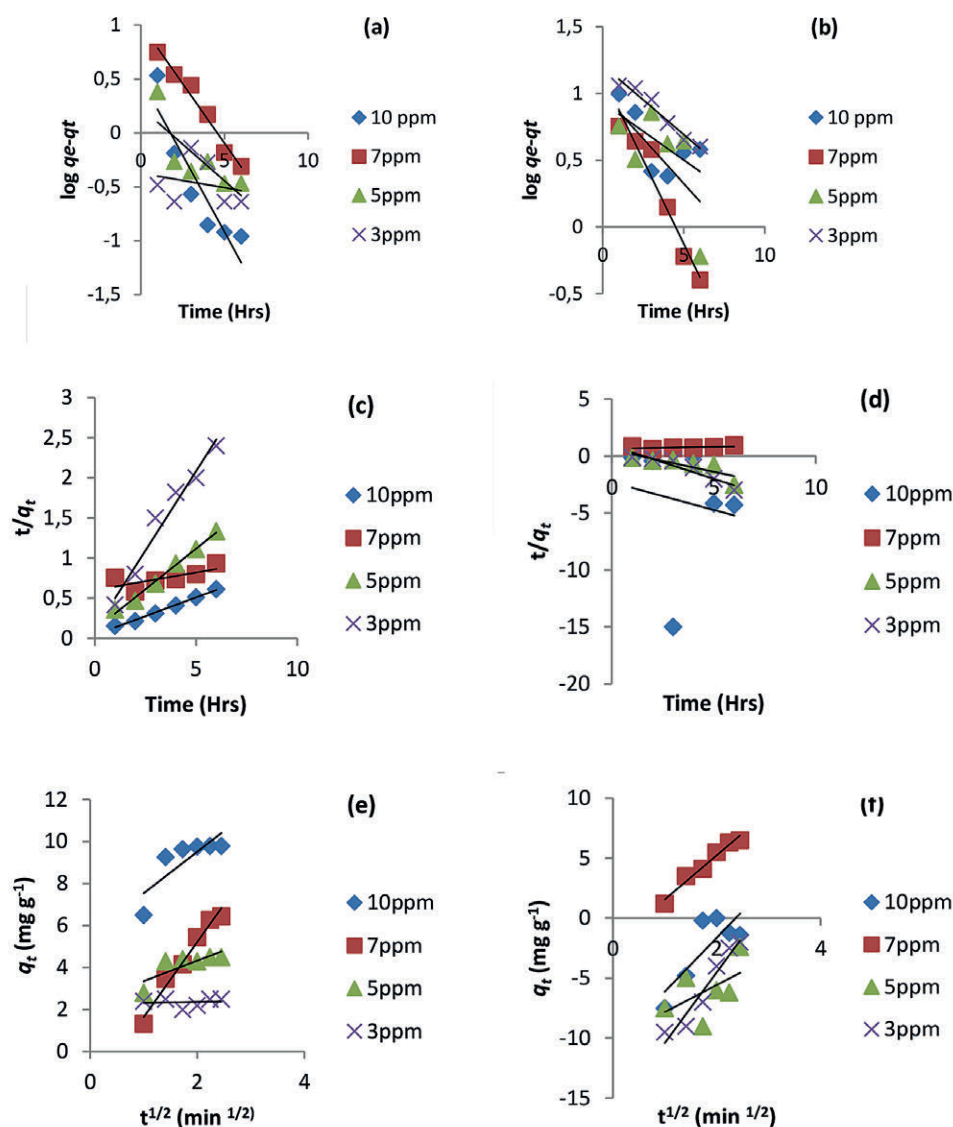


Figure 10. Adsorption studies for the Pb (II) ions adsorbed on BTACC and CAC: pseudo-first-order (a–b); pseudo-second-order (c–d); intra-diffusion model (e–f)

Table 5. Pseudo-first-order, pseudo-second-order and intra-particle diffusion model parameters for the adsorption of Pb (II) ions onto BTACC and CAC at different initial concentrations

Kinetic model	Parameters	Concentrations (mg·L ⁻¹)							
		BTACC				CAC			
		10	7	5	3	10	7	5	3
Pseudo-first order	q_e exp (mg·g ⁻¹)	9.91	6.93	4.84	2.73	2.4	6.9	1.8	2.0
	k_1 (min ⁻¹)	0.284	0.112	0.135	0.026	0.086	0.251	0.134	0.103
	q_e cal (mg·g ⁻¹)	3.17	4.49	1.71	2.37	8.55	13.49	9.93	16.26
	R^2	0.839	0.686	0.641	0.652	0.530	0.843	0.624	0.856
Pseudo-second order	q_e cal (mg·g ⁻¹)	10.75	22.73	4.98	2.53	2.06	32.26	2.62	1.72
	k_2 (g·mg ⁻¹ ·min ⁻¹)	2.27	0.07	1.89	0.01	0.21	0.05	0.72	0.64
	R^2	0.99	0.95	0.99	0.97	0.925	0.93	0.93	0.94
Intraparticle diffusion	K_d (mg·g ⁻¹ ·min ^{-1/2})	1.976	3.580	0.974	0.041	4.489	3.681	2.252	5.898
	I (mg·g ⁻¹)	5.559	1.946	2.374	2.276	10.62	2.129	10.08	16.31
	R^2	0.668	0.975	0.635	0.011	0.654	0.970	0.291	0.939
Film diffusion	D_f (cm ² ·s ⁻¹)	8.08×10^{-6}	4.55×10^{-6}	1.44×10^{-6}	3.45×10^{-6}	1.03×10^{-6}	3.47×10^{-6}	7.44×10^{-6}	6.29×10^{-6}
Pore diffusion	D_p (cm ² ·s ⁻¹)	2.79×10^{-6}	1.75×10^{-5}	1.79×10^{-5}	1.08×10^{-5}	1.23×10^{-5}	1.89×10^{-5}	1.07×10^{-5}	1.18×10^{-5}

to be the rate-limiting factor, the pore diffusion co-efficient (D_p) should be in the range of 10^{-11} – 10^{-13} cm²·s⁻¹ (Senthil Kumar et al., 2009). Assuming a spherical geometry for the sorbents, the theoretical rate constant of the process can be correlated to the pore diffusion co-efficient in accordance with the expression:

$$t_{1/2} = 0.03 r_0^2 / D_p \quad (7)$$

$$t_{1/2} = 0.23 r_0^3 c^* / c D_f \quad (8)$$

where r_0 is the radius of the sorbent expressed in cm, D_p is the pore diffusion coefficient expressed in cm²·s⁻¹, D_f the film diffusion co-efficient expressed in cm²·s⁻¹, c^*/c is the equilibrium loading factor of the sorbent and $t_{1/2}$ is the half-life period expressed in seconds. The results for BTACC and CAC are given in Table 5.

CONCLUSION

The present investigation showed that carbon prepared from a novel waste material, BTACC, is able to remove of 98% Pb (II) ion from aqueous solution, whereas CAC removed only 52%. The presence of hydroxyl, carboxylic and sulphonic acid groups in BTACC was confirmed by FT-IR spectroscopy, and give ion exchange properties to the BTACC. The ion exchange value of BTACC was found to be 0.6024 molar equiv·g⁻¹. Compared with CAC, which has a very narrow range of pH 5, BTACC as a wider applicable pH range from 4–10 where the carbon is applied in wastewater treatment. The equilibrium data agreed well with the Langmuir adsorption value. The sorption kinetic pseudo-second-order model and intraparticle diffusion is not the sole rate-controlling factor. The value of the film diffusion co-efficient obtained for BTACC and CAC lies between 10^{-6} to 10^{-8} cm²·s⁻¹ for various initial concentrations of Pb (II) ions. Hence it is evident that Pb (II) removal follows film diffusion which is the rate-limiting factor. Therefore, it can be concluded that carbon derived from BTACC is very effective for the removal of Pb (II) from aqueous solution.

ACKNOWLEDGEMENTS

The authors are very thankful to Dr A Elango ME, Ph.D, Principal, Alagappa Chettiar College of Engineering and

Technology, Karaikudi for providing necessary facilities and encouragement to complete this work. We also thank the management of CMS College of Engineering and Technology, Coimbatore for their constant support.

REFERENCES

- ABBASZADEH S, RAFIDAH WAN ALWI S, WEBB C, GHASEMI N and MUHAMAD II (2016) Treatment of lead-contaminated water using activated carbon adsorbent from locally available papaya peel biowaste. *J. Clean. Prod.* **118** 210–222. <https://doi.org/10.1016/j.jclepro.2016.01.054>
- ACHARYA J, SAHU JN, MOHANTY CR and MEIKAP BC (2009) Removal of lead (II) from wastewater by activated carbon developed from tamarind wood by zinc chloride activation. *Chem. Eng. J.* **149** 249–262. <https://doi.org/10.1016/j.ccej.2008.10.029>
- AHAMED M, VERMA S, KUMAR A and SIDDIQUI MKJ (2005) Environmental exposure to lead and its correlation with biochemical indices in children. *Sci. Total Environ.* **346** 48–55. <https://doi.org/10.1016/j.scitotenv.2004.12.019>
- APHA, AWWA (1973) *Standard Methods for the Examination of Water and Wastewater*. American Public Health Association, Washington D.C.
- BABARINDE NAA, BABALOLA JO and SANNNI RA (2006) Biosorption of lead ions from aqueous solution by maize leaf. *Int. J. Phys. Sci.* **1** 23–26.
- DEMIRBAS A (2008) Heavy metal adsorption onto agro-based waste materials: A review. *J. Hazardous Mater.* **157** 220–229. <https://doi.org/10.1016/j.jhazmat.2008.01.024>
- GOEL J, KADIRVELU K, RAJAGOPAL C and KUMAR GARG V (2005) Removal of lead (II) by adsorption using treated granular activated carbon, batch and column studies. *J. Hazardous Mater.* **125** 211–220. <https://doi.org/10.1016/j.jhazmat.2005.05.032>
- GROFFMAN A, PETERSON S and BROOKINS D (1992) Removing lead from wastewater using zeolite. *Water Environ. Technol.* **4** 54–59.
- GUADALUPE R, REYNEL-AVILA HE, BONILLA-PETRICIOLET A, CANO-RODRÍGUEZ I, VELASCO-SANTOS C and MARTÍNEZ-HERNÁNDEZ AL (2008) Recycling poultry feathers for Pb removal from wastewater: kinetic and equilibrium studies. *Int. J. of Chem. Molec. Eng.* **2** (11) 338–346.
- HALL KR, EAGLETON LC, ACRIVOS A and VERMEULEN TH (1966) Pore- and solid-diffusion kinetics in fixed-bed adsorption under constant-pattern conditions. *Fund. Ind. Eng. Chem.* **5** 212–223. <https://doi.org/10.1021/i160018a011>
- HO YS and MCKAY G (1999) Pseudo-second order model for sorption processes. *Process Biochem.* **34** 451–465. [https://doi.org/10.1016/S0032-9592\(98\)00112-5](https://doi.org/10.1016/S0032-9592(98)00112-5)
- HO YS, WASE DAJ and FORSTER CF (1996) Removal of lead ions

- from aqueous solution using sphagnum moss peat as adsorbent. *Water SA* **22** 219–224.
- HUA M, ZHANG S, PAN B, ZHANG W, LU LV and ZHANG Q (2012) Heavy metal removal from water/wastewater by nanosized metal oxides: A review. *J. Hazardous Mater.* **211** 317–331. <https://doi.org/10.1016/j.jhazmat.2011.10.016>
- INGLEZAKIS VJ, LOIZIDOU MD and GRIGOROPOULOU HP (2003) Ion exchange of Pb²⁺, Cu²⁺, Fe³⁺ and Cr³⁺ on natural clinoptilolite: selectivity determination and influence of acidity on metal uptake. *J. Colloid Interf. Sci.* **261** 49–54. [https://doi.org/10.1016/s0021-9797\(02\)00244-8](https://doi.org/10.1016/s0021-9797(02)00244-8)
- JYOTIKUSUM A, SAHU JN, MOHANTY CR and MEIKAP BC (2009) Removal of lead (II) from wastewater by activated carbon developed from tamarind wood by zinc chloride activation. *Chem. Eng. J.* **149** 249–262. <https://doi.org/10.1016/j.cej.2008.10.029>
- KADIRVELU K, THAMARAISELVI K and NAMASIVAYAM C (2001) Removal of heavy metals from industrial wastewaters by adsorption onto activated carbon prepared from an agricultural solid waste. *Bioresour. Technol.* **76** 63–65. [https://doi.org/10.1016/s0960-8524\(00\)00072-9](https://doi.org/10.1016/s0960-8524(00)00072-9)
- LAGERGREN S (1898) About the theory of so-called adsorption of solute substances. *Kungliga Svenska Vetenskapsakad Handlingar* **24** 1–39.
- MICHELSON LD, GIDEON PG, PACE EG and KUTAL LH (1975) US Department of Industry, Office of Water Research and Technology, Bulletin 74.
- MOUSAVI HZ and SEYEDI SR (2011) Nettle ash as a low cost adsorbent for the removal of nickel and cadmium from wastewater. *Int. J. Environ. Sci. Technol.* **8** (1) 195–202. <https://doi.org/10.1007/bf03326209>
- PATHANIA D, SHARMA S and SINGH P (2017) Removal of methylene blue by adsorption onto activated carbon developed from *Ficus carica* bast. *Arab. J. Chem.* **10** 1445–1451. <https://doi.org/10.1016/j.arabjc.2013.04.021>
- PEHLIVAN E, ALTUN T and PARLAYICI S (2008) Utilization of barley straws as biosorbents for Cu²⁺ and Pb²⁺ ions. *J. Hazardous Mater.* <http://dx.doi.org/10.1016/j.jhazmat.08.115>.
- RAUNIO S and TAHATI H (2001) Glutamate and calcium uptake in astrocytes after acute lead exposure. *Chemosphere* **44** 355–359. [https://doi.org/10.1016/s0045-6535\(00\)00305-2](https://doi.org/10.1016/s0045-6535(00)00305-2)
- RIAZA M, NADEEMA R, HANIFA MA, ANSARIC TM and REHMANA K (2009) Pb(II) biosorption from hazardous aqueous streams using *Gossypium hirsutum* (cotton) waste biomass. *J. Hazardous Mater.* **161** 88–94.
- SENTHIL KUMAR P and GAYATHRI R (2009) Adsorption of Pb²⁺ ions from aqueous solutions onto bael tree leaf powder. *J. Eng. Sci. Technol.* **4** 381–399.
- SRINIVASAN K and RAMADEVI A (2005) Removal of lead in aqueous medium by tamarind nut carbon. *Ind. J. Environ. Protect.* **25** 420–428.
- SURESH JEYAKUMAR RP and CHANDRASEKARAN V (2014) Adsorption of lead (II) ions by activated carbons prepared from marine green algae. *Int. J. Ind. Chem.* **5** (2) 1–10. <https://doi.org/10.1007/s40090-014-0010-z>
- THARANITHARAN V and SRINIVASAN K (2009) Studies on the adsorption of Ni (II) on to modified amberlite XAD-7HP resin. *Ind. J. Environ. Protect.* **29** (4) 294.
- UGURLU M, KULA I, HAMDI KARAOGLU M and ARSLAN Y (2009) Removal of Ni(II) ions from aqueous solutions using activated carbon prepared from olive stone by ZnCl₂ activation. *Environ. Progress Sustainable Energ.* **28** 547–557. <https://doi.org/10.1002/ep.10358>
- WEBBER TN and CHAKRAVARTI RK (1974) Pore and solid diffusion models for fixed bed adsorbents. *Am. Inst. Chem. Eng.* **20** 228–238. <https://doi.org/10.1002/aic.690200204>
- WEBER WJ and MORRIS JC (1963) Kinetics of adsorption on carbon from solution. *J. Sanit. Eng. Div.* **89** 31–60.
- WHO (World Health Organization) (1977) *Environmental Health Criteria, 3, Lead*. WHO, Geneva.

Smallholder willingness to pay and preferences in the way irrigation water should be managed: a choice experiment application in KwaZulu-Natal, South Africa

U Chipfupa¹ and E Wale¹

¹School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa

ABSTRACT

Efficient and sustainable utilization of irrigation water is the key to realizing the objective of enhancing agricultural productivity and commercializing smallholder irrigation farming. Valuing and recognizing the scarcity of irrigation water is essential for its sustainable use. Using cross-sectional data from 328 smallholders in and around Makhathini and Ndumo-B irrigation schemes in KwaZulu-Natal Province, the study aimed to assess smallholder farmers' preferences for the way irrigation water resources should be managed and their willingness to pay for irrigation water. This was done employing a choice experiment method. The results suggest the need for irrigation water pricing to reflect irrigation intensity. They also show that improving agricultural production and productivity, with market access can enhance farmers' willingness and ability to pay for irrigation water. The need to consider multiple uses of irrigation water for sustainable utilization of water resources is evident, while supporting women smallholders will have a positive impact on their willingness to pay for irrigation water. The paper recommends a shift towards on-farm volumetric water pricing in the irrigation schemes. The schemes should also have clearly defined boundaries and enforceable rules on collective use of water. The design of irrigation infrastructure should integrate other uses of irrigation water such as domestic and livestock purposes. Consequently, there is a need for further research to ensure that irrigation water prices reflect the marginal value of irrigation water use. Policies should address factors that inherently result in gender differences in terms of access to productive resources which negatively affect sustainable water utilization.

Keywords: water valuation, farmers' willingness to pay for irrigation water, water pricing, farmers' preference on irrigation water management, multiple uses of irrigation water

INTRODUCTION

Efficient utilization of water should underpin efforts towards commercializing agricultural production in smallholder irrigation. Globally, irrigated agriculture uses nearly 70% of freshwater withdrawn from rivers and aquifers, and the figure is even higher for sub-Saharan Africa (87%) (FAO, 2011). The improved performance of irrigation systems compared to rainfed agriculture has triggered the expansion of irrigation farming. Consequently, this has increased the demand for water, adding to the growing concerns of water scarcity, amid other competing water uses. South Africa is one of the water-scarce countries in Africa. The average annual rainfall in the country is approximately 500 mm compared to a global average of 860 mm (Speelman et al., 2011; Schreiner, 2015). Kruger and Nxumalo (2017) also showed some large variations in the rainfall distribution in the country, with observed increases in the south and decreases in the northern and north-eastern parts over the period 1921–2015. There are concerns that by 2030 available water in the country's catchments will not be enough to meet the national water requirements (Schur, 2000). Hence, the need to promote sustainable and efficient utilization of the available water resources, failure of which could have dire consequences for smallholder livelihoods, rural employment and poverty reduction.

Schur (2000) identifies economic incentives as vital tools to improve the allocation of water resources. If a resource, such as irrigation water, is scarce, its scarcity must be reflected in the market thereby inducing the incentive to use it more

efficiently and sustainably (Ray, 2011). Increasingly, research has been focusing on finding ways and means of improving efficiency in the irrigated agriculture sector to increase water supply for other uses (Reinders et al., 2013). The challenge is that the current water management arrangements in most smallholder irrigation schemes in South Africa provide no incentive for sustainable utilization of water, maintenance of irrigation infrastructure and collective management of the schemes (Muchara et al., 2014). The scheme-level institutional failures, the poverty of irrigation infrastructure and the non-availability of agricultural water markets means irrigation water is often considered as a common pool resource (non-excludable, but rival in consumption) (Barton and Bergland, 2010; Muchara et al., 2016). This makes it difficult to monitor water consumption in the irrigation schemes or even charge volumetrically. Coupled with poor record-keeping and lack of water measurement devices in most schemes, precision in irrigation water valuation is close to impossible (Lange and Hassan, 2007; Young and Loomis, 2014; Muchara et al., 2016). In the absence of credible water value estimates, there would be little or no incentive for changing irrigation practices and efficient utilization of water. This study argues that until the commitments to gradually phase in the full cost of providing irrigation water are met, the current irrigation water charges will have minimal impact on smallholder irrigation practices.

Several approaches have been employed in the past to determine the economic value of irrigation water and inform water pricing and management policies. These are categorised as direct (stated preference) and indirect (revealed preference) methods. Direct techniques, such as contingent valuation (CV) and choice experiment modelling (CEM), obtain preferences directly through questioning individuals on their

*Corresponding author, email: uchipfupa@gmail.com

Received 11 October 2017; accepted in revised form 13 June 2019

willingness to pay (WTP) for a good or a service. Indirect techniques depend on observed market behaviour and data to infer an economic value of water (e.g., residual valuation, hedonic pricing, production function, and demand function approaches) (Young and Loomis, 2014). The commonly applied approach to irrigation water valuation in South Africa is the residual valuation method (e.g. Speelman et al., 2011; Muchara et al., 2016; Njoko and Mudhara, 2017). However, the use of this or any other revealed preference method is problematic because markets for some key inputs in smallholder agriculture (such as land) are non-existent. In such cases, it is recommended to use the stated preference approaches, i.e., the CV or CEM (e.g. Kunimitsu, 2009; and Bhaduri and Kloos, 2013). However, comparing these two, the CEM is more suitable to valuation of irrigation water in the context of smallholder farmers for the following reasons: The approach can model heterogeneous preferences in irrigation water services among smallholder farmers which results in higher water use efficiency (Abu-Zeid, 2001). The CEM also often requires a small sample of data to achieve similar accuracy in water valuation estimates (Barton and Taron, 2010). This is especially important given the high cost of data collection in Africa. The CEM is also not affected by some common biases associated with the 'warm glow' effect (deriving moral satisfaction from doing something deemed socially right) and strategic positioning by respondents often encountered in CV (Birol et al., 2006).

The application of the CEM is relatively new to irrigation water valuation with just a few studies conducted so far (e.g. Kunimitsu, 2009; Barton and Bergland, 2010; Bhaduri and Kloos, 2013). Most of the studies are from Asia and there is a gap in the literature regarding Africa's smallholder irrigation sector. To the best of the authors' knowledge CEM has not been applied to irrigation water valuation in the context of South Africa. However, the method has been widely used in other sectors in the country to determine non-market values for different environmental goods (e.g. Jaeck and Lifran, 2009), municipal water (e.g. Kanyoka et al., 2008) and technology adoption (e.g. Asrat et al., 2010), among others. Therefore, this study aimed to use the CEM to assess farmers' preferences on how irrigation water should be managed and determine their willingness to pay (WTP) for the resource. The farmers' preferences are assessed from three angles, i.e., water management, multiple uses of water, and multiple cropping. These represent the institutional arrangements in irrigation water management, other possible uses of irrigation water and the demand for irrigation water, respectively. The study contributes to the literature on water pricing in smallholder irrigation. The findings can be used to inform irrigation water management and water pricing policies for resource-poor farmers, contributing to sustainable and efficient utilisation of irrigation water resources.

RESEARCH METHODOLOGY

The South African context

Agricultural water in most smallholder irrigation schemes in South Africa is provided as a free commodity, subsidized exclusively by the government (Muchara et al., 2014). Schreiner (2015) reports that the country is currently providing a subsidy of over US\$30 million per year to the irrigation sector. The 'Draft Pricing Strategy for Water Use Charges' drawn in terms of the National Water Act of 1998 (RSA, 1998) gives provisions for subsidized water pricing rates,

including operations and maintenance charges, for irrigation schemes benefiting resource-poor communities (DWS, 2015). The policy states that farmers in such communities incur no charge for the initial 5 years, and after that the water charges are phased in for the next 5 years at a rate of 20% per annum. The pricing strategy attempts to balance economic efficiency and the social equity side of irrigation water provision. However, this has created perceptions that water is a free good and situations where smallholder irrigation schemes are dependent on the government for operation and maintenance costs (Backeberg, 2006).

Study area

The study was conducted in two areas in Jozini, a local municipality in uMkhanyakhude District in the northern part of KwaZulu-Natal Province, South Africa (Fig. 1). Jozini covers 3 057km² of land and borders Mozambique to the north, eSwatini to the west and four other local municipalities to the east and south. It is predominately rural but has 4 semi-formalized towns that act as tertiary centres. The municipality has a population size of 186 502 which is largely youthful (72% below 29 years) and mostly female (54%). Education levels are low (13.5% have no schooling, 2% have a post-Grade 12 qualification) and this is reflected in the high poverty levels (43% of households reporting no income in the last census) (Jozini Local Municipality, 2015).

According to the Jozini Municipality, agriculture forms one of the major important sectors within the municipality (Jozini Local Municipality, 2015). The general livelihoods resemble a mixed farming system with farmers engaged in either crop farming (rainfed and irrigated) or livestock production or both. Due to persistent droughts, irrigation farming has assumed greater importance. Irrigation is conducted both in

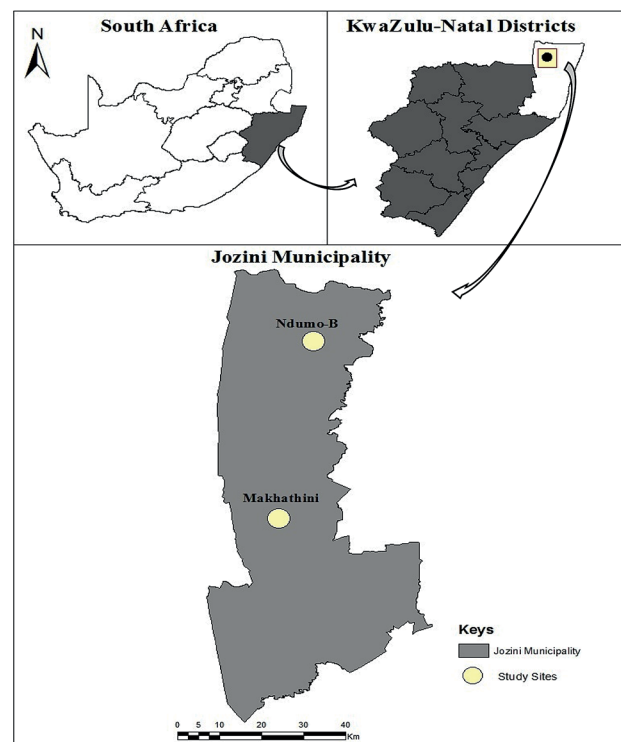


Figure 1. Location of the study area

schemes and outside. Land (in scheme or outside) is held on a permission to occupy (only use rights) basis granted by the traditional authorities. The land holding of the study sample ranges from as little as 0.2 ha to 10 ha of land. Those who irrigate outside the schemes include independent irrigators, homestead gardeners and community gardeners.

Scheme irrigation is mainly through two major schemes: i.e., Makhathini and Ndumo-B. These are approximately 80 km apart. Makhathini irrigation scheme covers an estimated 4 500 ha of irrigated land and has a total of 1 167 smallholders farming as individuals or part of cooperatives. It is managed by Mjindi Farming Private Limited, a state-owned entity. The scheme is serviced by a 34 km canal which carries water from the Jozini Dam, drawn by 6 pump stations. Ndumo-B irrigation scheme is relatively small and covers 500 ha of land. It is managed and operated by two cooperatives. At the time of the survey, only a part of the scheme with 21 members was operational (200 ha). Water is drawn from the Pongola River using an electric pump and brought to the plots using pipes. In both schemes, there is currently no volumetric water charging system at the farmer level.

The design of the choice experiment

Three critical steps are followed when designing a choice experiment, namely, the establishment of attributes of interest, assigning levels and, finally, the design of the choice sets (Mangham et al., 2009; Johnson et al., 2013). How each step is conducted has implications for the validity and credibility of the results. In this study, complementary processes were followed to identify and assign levels to attributes of irrigation water in the target communities. The processes include literature review (including policy documents), in-depth discussions with farmers, field observations and key informant interviews with relevant stakeholders. For selection, an attribute had to be relevant to the agricultural policy direction in South Africa, hold significant value to the smallholders in relation to the payment of water and have literature which supports its importance. This process resulted in 4 attributes (Table 1).

The attribute 'membership to an organization governing water use' represents institutional arrangements in the

irrigation schemes and knowledge on collective water management. Water governance and management of irrigation schemes are key aspects of sustainable management and success of smallholder irrigation in South Africa (Muchara et al., 2014). Currently, water management is implemented through cooperatives or a third-party institution managing the irrigation schemes on behalf of farmers. According to the National Water Act of 1998 (RSA, 1998), the water allocation rights are obtainable by any individual or organization drawing water from a surface or groundwater resource. However, resource-poor farmers are encouraged to form cooperatives, to assist them not only in water management but also in access to information, finance and high-value markets.

The attribute 'multiple uses of irrigation water' represents the possibility of using irrigation water for uses other than irrigation. The lack of consideration of the different uses of irrigation water results in undervaluation and inefficient allocation of the resource (Meinzen-Dick and Van Der Hoek, 2001). The current water pricing policy (see DWS, 2015) does not consider these other different dimensions in irrigation water valuation. Currently, 20% of the sampled smallholders use irrigation water solely for irrigation purposes while the remaining majority also use it for other purposes (watering of livestock and/or domestic use). Though not desirable, since access to water is at least in theory considered a human right, authorities indicate that it is possible to fence off the canal or use a pipe system that prevents access outside of the schemes.

The attribute 'number of crops per season' characterizes the demand for irrigation water by each farmer, i.e., irrigation intensity. Farmers growing more crops (multiple crops or more quantities of the same crop) are more likely to use more water per season, yet they pay the same amount of annual water fees per hectare. Currently, some scheme irrigators have voiced their concerns with the non-volumetric charging system and believe that even in the absence of water meters, those growing more crops should pay more for water.

The attribute 'annual payment for irrigation water per ha (water fees)' includes both raw water fees and water service charges (electricity and maintenance) paid by smallholders to access irrigation water. Raw irrigation water is subsidized, with farmers paying minimal fees or nothing at all. However, to some extent, they contribute to the maintenance of water

Table 1. Attributes for the choice experiment*

Attribute	Level	Expected impact on choice
Membership to an organization governing water use	Yes No (<i>status quo</i>)	+
Multiple uses of irrigation water	Irrigation only (<i>status quo</i>) Irrigation and domestic use Irrigation and livestock Irrigation, domestic and livestock use	+
Number of crops per season	1 crop per season (<i>status quo</i>) 2 crops per season 3 crops per season 4 and above crops per season	+
Annual payment for irrigation water (water fees) (ZAR-ha ⁻¹)	2 500 (<i>status quo</i>) 3 000 5 000 7 000	-

*A combination of the attribute levels with *status quo* in parentheses represents the current scenario used in the study. It is assumed that in the absence of institutional challenges farmers in both schemes will face the same water charges. In the absence of any other information, the same *status quo* is also used for out-of-scheme farmers.

infrastructure and pumping charges. Those from Makhathini irrigation scheme currently pay a subsidized charge of approximately 2 500 ZAR·ha⁻¹·yr⁻¹. The charge includes raw water and other related services (electricity and water infrastructure maintenance). Smallholders from Ndumo-B irrigation scheme pay almost 3 times more (7 200 ZAR·ha⁻¹·yr⁻¹) since they cater for the full cost of water provision. Though they have no charge for raw water, their electricity bill translates to an average of approximately 600 ZAR·ha⁻¹·yr⁻¹.

In designing the choice sets, the study aimed to achieve a balance between statistical efficiency of the design and response efficiency. Statistical efficiency refers to 'minimizing the confidence intervals around parameter estimates in a choice model' while response efficiency deals with the 'measurement error resulting from respondents' inattention to the choice questions or other unobserved, contextual influences' (Johnson et al., 2013 p. 6). Response efficiency could also be the result of respondent's failure to comprehend all the choice sets. Statistically efficient designs are orthogonal (levels of each attribute are statistically independent of each other), balanced (each attribute level appears in equal proportion across choices) and minimize overlap (repeating of the attribute level with a choice set) (Ryan et al., 2012). Due to the practical impossibility of presenting the full set of choices (128 (2 × 4³)), a fractional factorial design was adopted in the study (Kuhfeld, 2010). The orthogonal design option in SPSS v 24 was used to generate a choice set of 16 alternatives. Pairwise correlation coefficients of the attributes showed that the choice set met the orthogonality criterion while the level of balance was achieved. To increase response efficiency, a compromise was made on the minimum overlap condition. Johnson et al. (2013) state that overlap improves response efficiency by reducing the cognitive burden of evaluating huge attribute differences in a choice set. However, this was kept to a minimum to limit the negative impact on the design efficiency. Of the 16 alternatives, one was similar to the status quo scenario and hence was dropped from the list. The remaining list of 15 alternatives was divided into 5 choice sets of 4 alternatives, including the opt-out choice. Pretesting results showed that smallholders could respond to these with minimum difficulties in understanding or before

boredom sets in. Literature suggests a practical limit of 18 choice sets of 2 options that an individual can respond to with no difficulties (Mangham et al., 2009).



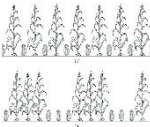

Sampling and data collection

The study targeted smallholders in and around the two irrigation schemes. This was done because the schemes also benefit other farmers outside the schemes. Moreover, it is important to align the study with the South African Government's objective of irrigation expansion. A total of 328 smallholder farmers were interviewed. The sample was stratified to include scheme irrigators (*n* = 109), non-scheme irrigators (*n* = 174) and dryland farmers (*n* = 45). Dryland farmers are those currently practicing rainfed farming. The study targeted at least 10% of scheme farmers actively engaged in farming at the time of conducting the study. Owing to their homogeneity, simple random sampling was used to identify the scheme irrigators using a list obtained from the scheme management and cooperatives. The other types of farmers were identified through snowball sampling since there was no available information regarding their populations. However, this was done proportionally considering the critical mass in each area. The choice experiment was conducted in a once-off survey in April 2016 through a questionnaire administered by 6 well-trained enumerators over a 7-day period. Besides the choice experiment questions, the questionnaire also covered information on the demographic and socioeconomic status of the farm households. To check the cognitive ability of the farmers in comprehending and responding to the proposed choice sets, a pre-testing exercise was conducted. Issues from the pre-test were addressed and changes made to the questionnaire before the actual survey.

Empirical approach

The theoretical foundation for choice modelling, the random utility model, is used to analyse the farmer's utility maximization problem (McFadden, 1973). It is founded on Lancaster's characteristics theory which indicates that it is not

Table 2. Example of a choice set in the study

	Attributes	Option 1	Option 2	Option 3	Status quo
	Membership to a water organization	No	No	Yes	No
	Multiple uses of water	Irrigation only	Irrigation, domestic and livestock	Irrigation and domestic use	Irrigation only
	Number of crops per season	Three crops per season	Two crops per season	At least 4 crops per season	One crop per season
	Annual payment of water (Water fees) (ZAR·ha ⁻¹)	7 000	3 000	2 500	2 500

Please tick only one

the good but the attributes it possesses that determines its value to a consumer (Lancaster, 1966). The decision maker is the only one with knowledge of this utility. What the researcher observes are the different levels of the attributes and not the utility of the decision maker (Train, 2009). The study assumes that smallholders are rational in their decisions, and their WTP for irrigation water is determined by the utility they derive from the use of that water. The utility depends on their preferences for the various factors which impact on irrigation water use. At any one time, given a set of alternatives, rational farmers choose an alternative that gives them the highest utility.

To illustrate this, if a farmer's utility depends on a choice made from a given choice set (J) of irrigation water use options, the utility function for the farmer is given by:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad j = 1, 2, \dots, J \quad (1)$$

where, for any farmer i , a given level of utility U is associated with alternative choice j . The utility function for each farmer has 2 parts, i.e., an observable part (V) as well and an unobservable part (ε). V is assumed to be a linear function of the attributes and any socio-economic characteristics of the farmer such as income and resource endowment. The exact estimation of the model depends on the assumptions made about the probability distribution of ε_{ij} . If ε_{ij} is independent and identically distributed with extreme value distribution, one should estimate the conditional logit model (Greene, 2012). In this model, the probability of individual farmer i choosing alternative j can then be expressed as:

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{j=1}^J \exp(V_{ij})} = \frac{\exp(\beta' X_{ij})}{\sum_{j=1}^J \exp(\beta' X_{ij})} \quad (2)$$

where X_{ij} are all the observed factors and β' represents parameters estimated from the model. If there are m attributes, V_{ij} is expressed as:

$$V_{ij} = \beta_0 ASC + \beta_1 X_{1j} + \beta_2 X_{2j} + \beta_m X_{mj} \quad (3)$$

where β_m is the coefficient of attribute X_m . The status quo or current situation is represented by ASC which is a dummy variable with 1 = choice of current status and 0 = any other alternative. The inclusion of the status quo provides an opt-out choice for those farmers not interested in any of the suggested alternatives.

However, if the error terms are correlated and not identically distributed, the independence from irrelevant alternatives (IIA) assumption of the conditional logit model is violated (Hausman and McFadden, 1984). The likelihood of this happening is high in the presence of heterogeneity in farmer preferences and socioeconomic factors. In such situations, estimating the conditional logit would result in biased estimates. The recommendation is to use the mixed logit, a less restrictive model that allows random taste variation and correlation in the error terms (Train, 2009; Greene, 2012).

In the mixed logit model, the probability P of individual farmer i choosing alternative j then becomes:

$$P_{ij} = \int \left(\frac{\exp(\beta' X_{ij})}{\sum_{j=1}^J \exp(\beta' X_{ij})} \right) f(\beta) d\beta \quad (4)$$

where $f(\beta)$ is the distribution function for β and X_{ij} is a vector of observed variables. Three interaction terms were introduced

in the model to test several assumptions regarding WTP for irrigation water, i.e., the effect of gender and crop income on the cost of irrigation water and the effect of livestock ownership on the multiple uses of irrigation water. In the estimation of the mixed logit, the non-price attributes were randomized while the cost attribute was treated as non-random (Layton, 2000; Lee et al., 2014), a preferred option because it allowed the distribution of the WTP to be the same as that of the attribute (Scarpa et al., 2008), making it easier to compute WTP estimates.

Willingness to pay estimation

Since the cost attribute (water fees) is taken as non-random, the WTP distribution takes the same form as that of the non-price attributes. According to Scarpa et al. (2008), the mean and standard deviation of the WTP can thus be given by the mean and standard deviation of the attribute scaled by the inverse of the price coefficient. For a given attribute, the ratio of the attribute to the price coefficient also represents the marginal WTP for a change in the attribute values (Lee et al., 2014). Following Bech and Gyrd-Hansen (2005), the coefficient of the dummy attributes' in the equation, e.g. 'multiple uses of irrigation water', is multiplied by two. The equation is slightly adjusted to incorporate the interaction effects associated with the price or non-price attributes (Giergiczny et al., 2012; Bhaduri and Kloos, 2013). For example, computing the WTP for the attribute 'multiple uses of irrigation water' should include 2 terms in the numerator, i.e., 'multiple uses' and 'multiple uses \times no_cattle'. The denominator, which will be the same for all attributes, should include 3 terms, i.e., 'water fees', 'water fees \times gender' and 'water fees \times crop_income'.

$$WTP_{multiple_uses} = - \frac{2\beta_{multiple_uses} + \beta_{multiple_uses \times no_cattle}}{\beta_{fee} + \beta_{fees \times gender} + \beta_{fees \times crop_income}} \quad (5)$$

RESULTS AND DISCUSSION

The comparative descriptive results

Table 3 shows the characteristics of the respondents in respect of the demographics and other variables related to the attributes used in the choice experiment. Comparison by farmer category shows statistically significant differences in the number of crops grown per season, cattle ownership, multiple uses of irrigation water and interest in collective water management. Non-scheme irrigators grow more crops per season while rainfed farmers own approximately 3 times the number of cattle compared to the other farmers. Evidence of multiple uses of irrigation water is higher among farmers outside compared to those in the schemes. Interest in collective water management is also higher among farmers outside of the schemes compared to those inside. This is because smallholder farmers in the schemes have negative experiences with collective water management. Non-compliance by some members result in consequences that affect even those who are compliant. For example, the failure by some to pay for water use charges often leads to the disconnection of electricity or water which affects everyone.

Regarding differences across the study areas (Makhathini and Ndumo-B), statistically significant differences are observed in gender, the number of crops grown per season, crop income, membership to cooperatives and multiple uses of water. Makhathini had more female respondents and a higher

proportion of farmers in cooperative membership compared to Ndumo-B. However, Ndumo-B farmers grow, on average, more crops per season and obtain approximately 4 times the crop income of Makhathini farmers. Furthermore, Ndumo-B has a higher proportion of farmers who use irrigation water for other purposes.

The results of the choice experiment

The study estimates the empirical models using a dataset of 6 450 (327 x 5 x 4) observations. Each farmer had 5 choices from choice sets containing 4 options. Of the 328 questionnaires completed one had incomplete information and hence was dropped. The estimation was conducted in STATA 13. To reduce simulation errors in parameter estimates, 100 Halton draws were used in the mixed logit estimation. For ease of analysis and interpretation of results the attribute 'multiple uses of water' was transformed into a dummy variable with 1 representing multiples uses of irrigation water and 0 otherwise.

After estimating the conditional logit model, a test for the IIA assumption using the Hausman-McFadden test was conducted. It compared the parameter estimates of the full model, which are consistent and efficient, to those of the restricted model that includes some of the outcomes (consistent and inefficient) (Cheng and Long, 2007). The test was estimated for each outcome of the dependent variable (alternative choices), which meant 4 tests were conducted. The significant test results in Table 4 meant that the IIA assumption did not hold and hence the mixed logit model was estimated. The log-likelihood, Bayesian information criterion (BIC), and Akaike information criterion (AIC) values also confirmed that the mixed logit, which allows heterogeneous preferences, is the better model compared to the conditional logit.

Table 5 presents the results of the mixed logit models estimated with and without interaction terms. As indicated earlier, the dependent variable is the 'farmer's choice of irrigation

water use options'. The results of the two are similar and the signs of the coefficients are as expected except for the 'membership to a water governing institution' attribute. The results suggest that the coefficient of the membership attribute does not statistically significantly affect choices. The other three attributes have a significant impact on choices. The negative coefficient of the attribute 'water fees' shows that higher fees reduce the probability of a farmer selecting an expensive option. Both the 'number of crops per season' and 'multiple uses of irrigation water' positively influence the choice of an option. This means all farmer categories prefer the option of growing more crops and are willing to pay more for increased irrigation intensity. The results confirm findings from other studies that identified the importance of recognizing multiple uses of irrigation water (e.g., Meinen-Dick and Van Der Hoek, 2001; Boelee et al., 2007) in water valuation. The negative sign of the ASC coefficient shows that farmers prefer the alternatives that offer different combinations of water services compared to the status quo. Only 20% of the farmers prefer the status quo situation.

The study tests the effect of gender differences on the WTP through an interaction term of water fees and gender. The results show a negative coefficient for the 'water fees x gender' interaction term, suggesting that being male negatively affects choices resulting in a lower WTP for irrigation water compared to females. Women farmers in the study community generally place a higher significance on crop farming compared to men, most of whom prefer the culturally valued livestock production. The SOFA Team and Doss (2011) drew similar conclusions for women in sub-Saharan Africa, and this could explain their higher WTP for irrigation water.

Table 4. Test for the IIA assumption

Hausman-McFadden test	Chi-square	p-value
Exclude Option 1	43.7	0.000
Exclude Option 2	168.8	0.000
Exclude Option 3	222.3	0.000
Exclude Option 4	31.0	0.000

Table 3. Demographic and other characteristics of the sample farmers

	Scheme irrigators (n = 109)	Non-scheme irrigators (n = 174)	Dryland farmers (n = 45)	P-value	Makhathini (n = 216)	Ndumo (n = 112)	P-value	Total (N = 328)
Gender (% female)	62.4	64.6	71.1	0.587	75.5	46.4	0.000	64.9
Age of farmer	47.6 (1.2)	49.2 (0.9)	50.2 (1.8)	0.357	49.1 (12.6)	48.4 (10.7)	0.607	48.8 (0.66)
Number of years in formal school	4.8 (0.4)	4.1 (0.4)	3.8 (0.7)	0.303	43 (4.5)	4.3 (4.6)	0.988	4.3 (0.3)
Number of crops	1.3 (0.1)	1.8 (1.0)	1.1 (0.3)	0.000	1.4 (0.7)	1.7 (1.0)	0.001	1.5 (0.1)
Number of cattle	5.3 (1.2)	4.6 (0.7)	14.8 (5.5)	0.001	6.3 (19.1)	6.0 (11.1)	0.875	6.24 (0.9)
Income from crops (R'000)	15.3 (5.2)	7.5 (1.3)	5.4 (1.4)	0.118	4.9 (0.6)	19.2 (5.3)	0.000	9.8 (1.9)
Membership in a cooperative (% members)	64.2	67.8	73.3	0.539	72.9	56.8	0.003	67.4
Interested in being part of an institution governing water (% interested)	59.0	68.2	66.7	0.069	64.6	65.7	0.293	65.0
Other uses of irrigation water:								
Livestock watering (%)	55.6	80.6	70.5	0.000	62.8	86.9	0.000	70.8
Domestic use (%)	59.3	81.2	70.5	0.001	65.6	86.9	0.000	72.4
Construction (%)	52.8	78.2	68.2	0.000	61.4	82.2	0.000	68.3

Note: Values in parentheses are standard errors

Table 5. Estimation results for all farmers ($n = 327$)

Attributes	MXL simple		MXL with interactions	
	Coef.	Std. Err.	Coef.	Std. Err.
ASC	-0.586 ^a	0.190	-0.554 ^a	0.190
Membership to water organization	-0.053	0.083	-0.046	0.083
Number of crops	0.354 ^a	0.062	0.358 ^a	0.063
Multiple uses	1.098 ^a	0.216	0.959 ^a	0.224
Water fees	-4.81 x 10 ^{4a}	2.48 x 10 ⁵	-4.34 x 10 ^{4a}	2.97 x 10 ⁵
Multiple uses x no. of cattle			0.035 ^b	0.015
Water fees x gender			-1.94 x 10 ^{4a}	5.17 x 10 ⁵
Water fees x crop income			1.52 x 10 ⁹	9.54 x 10 ¹⁰
SD				
Membership to water organization	0.170	0.219	0.126	0.242
Number of crops	0.712 ^a	0.074	0.733 ^a	0.075
Multiple uses	2.763 ^a	0.210	2.686 ^a	0.207
Number of observations	6 540		6 540	
LR chi ² (4)	541.7		529.7	
Prob > chi ²	0.000		0.000	
Log likelihood	-1 777.8		-1 763.9	
AIC	3 571.6		3 549.8	
BIC	3 625.9		3 624.5	

Note: a, b, c significant at the 1%, 5% and 10%, respectively; SD – standard deviations

The study also tests the hypothesis that higher crop income increases farmers' WTP for water using an interaction term between 'water fees' and 'crop income'. The results indicate that the potential for higher income earnings from crop production increases farmers' WTP. Thus, improving the productivity of agricultural enterprises and ensuring profitable markets for the marketable surplus will positively impact farmers' effective demand and hence their WTP for irrigation water. Similar results were also obtained in a study conducted in China where income had a positive and significant effect on WTP for irrigation water (Tang et al., 2013). The coefficient of the interaction term between cattle ownership and the 'multiple uses of irrigation water' attribute is statistically significant and positive. This shows that farmers with larger stocks of cattle have a higher probability of choosing the multiple uses attribute and are willing to pay more for water. These farmers are typically benefiting from the complementarity of crop-livestock integration. Other studies confirm these findings and show that on-farm irrigation water has value in livestock watering and feed production (Hewitt, 2013; FAO, 2018).

The standard deviations of the 'number of crops per season' and 'multiple uses of irrigation water' attributes are statistically significant ($p < 0.01$), showing heterogeneity in farmers' preferences for these attributes. The magnitudes of the mean and standard deviation show further information on the proportion of smallholders with a negative or positive preference of an attribute. Following Hole (2007), the proportions are given by $100 \times \Phi(-b_x/s_x)$, where b_x and s_x are the mean and standard deviation of the x^{th} coefficient while Φ is the cumulative standard normal distribution function. The results show that 69% of the smallholders prefer to use irrigation water for many more purposes and 65% prefer to grow more than one crop. Adding interaction effects to the model has no significant effect on the proportion of farmers with such preferences.

Preferences in managing irrigation water across farmer categories

The results in Table 6 show heterogeneity in preferences between the different categories of farmers and the study areas. The coefficient of the attribute 'multiple uses of irrigation water' is positive but significant only for the scheme and non-scheme irrigators. However, the coefficient of the interaction term between the attribute 'multiple uses of irrigation water' and the 'number of cattle owned' is statistically significant only for dryland farmers. This means, due to the value that dryland farmers place on their livestock, multiple uses of irrigation water are important only as they relate to the livestock enterprise. Other studies have shown that livestock are an integral component of smallholder mixed cropping systems and the integration enhances income diversification, land productivity and water efficiency (Liniger et al., 2011; Mekuria and Mekonnen, 2018). The coefficient of the interaction term 'water fees x crop income' has a significant positive influence on choices and the WTP for non-scheme irrigators only. This suggests that higher income from crops will enhance the ability to pay for irrigation water among farmers irrigating outside of the schemes.

Regarding spatial differences, the cattle ownership effect on 'multiple uses of irrigation water', and hence choices is statistically significant for Makhathini and not Ndumo-B area. This suggests that integrating livestock with crop production will enhance smallholder ability to pay for irrigation water in Makhathini. The impact of gender and crop income on 'water fees' and hence the ability to pay for irrigation water is statistically significant for Ndumo-B and not Makhathini. This means female smallholders in Ndumo-B are more price sensitive compared to men and have a lower WTP. Despite women valuing smallholder agriculture more than men, other factors make those in Ndumo-B more economically vulnerable and hence could face challenges paying for irrigation water. Njoko and Mudhara (2017) found similar

results that showed that in most parts of rural KwaZulu-Natal, men have a higher ability to pay for water compared to women. Sharaunga et al. (2016) report that women in South Africa are disproportionately economically disempowered compared to their male counterparts. The significance of the coefficient of 'water fees × crop income' interaction term for Ndumo-B shows that increased productivity and profitable markets for the marketable surplus will have more impact on smallholder WTP in this area and not Makhathini.

Farmers' WTP for different attributes

Table 7 presents the mean WTP for the different water-related services, estimated from the model with interaction effects. The negative WTP values show the lack of willingness to pay for that attribute. The bigger the number, the more unwilling are the farmers to pay for the attribute. The results suggest that farmers value the additional benefits derived from the use of irrigation water for other purposes more than the other attributes. 'Membership to a water governing institution' is the less valued of the three, for the reasons discussed earlier. The heterogeneity in preferences is observed through different WTP estimates for the attributes. Non-scheme irrigators are willing to pay ZAR1 213 more than what scheme irrigators are willing to pay for additional uses of irrigation water. This is because a higher proportion of non-scheme irrigators use irrigation water for other purposes compared to scheme irrigators (see Table 3). Similarly, for the same reasons, farmers in Ndumo-B are willing to pay more compared to Makhathini for the same attribute.

The results also suggest that farmers growing more crops (multiple and/or more of the same) are willing to pay extra for the use of more water, and the WTP is higher for scheme irrigators and Ndumo-B compared to other farmers. This finding suggests that irrigation water pricing should reflect irrigation intensity leading to efficient water allocation outcomes, an argument also put forward by Giraldo et al.

(2014). Although the initial cost might end up being prohibitive to resource-poor farmers (Abu-Zeid, 2001), volumetric water pricing remains the best option for improving efficient utilization of water in smallholder irrigation. However, in the absence of volumetric charging, enhancing irrigation design, such as having clearly defined and enforceable water-sharing arrangements can result in efficient allocation of water resources (Ostrom et al., 1994). Despite facing frequent crop failures, the low dryland farmers' WTP values across all attributes is an indication of their negative valuation and perceptions of irrigation water payment compared to the other smallholders. Payment for water is a new phenomenon to such farmers, most of whom have never paid for water use before. Their attitudes to irrigation water payment are thus bound to be different from the rest of the farmers.

CONCLUSION AND RECOMMENDATIONS

Water valuation is an important step to address market failure in irrigation water and induce efficient utilization of the resource in the smallholder irrigation sector. Water scarcity threatens agricultural production and productivity and, therefore, endangers food security, employment and rural livelihoods. The study set to assess farmers' preferences in managing irrigation water resources and determine their WTP for irrigation water using the CEM approach. It contributes to the debate on irrigation water pricing for resource-poor farmers and improving efficient utilization of water resources. The farmers' WTP to produce more valuable crops implies that irrigation water use charges should reflect irrigation intensity. This will also contribute to curbing over-irrigation which will continue to happen if water is under-priced. It is recommended that the water pricing policy shifts from the current average charge per hectare to a volumetric charging system at the farmer level. The government should fund the initial meter installation costs with a cost recovery

Table 6. Mixed logit estimation results for different farmer categories and study areas

Attributes	Scheme irrigators		Non-scheme irrigators		Dryland farmers		Makhathini		Ndumo-B	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Membership to water organization	-0.216	0.152	0.088	0.116	-0.121	0.275	-0.104	0.102	0.028	0.155
Number of crops	0.365 ^a	0.121	0.349 ^a	0.080	0.321 ^c	0.195	0.165 ^b	0.081	0.741 ^a	0.104
Multiple uses	0.657 ^c	0.404	1.339 ^a	0.317	0.684	0.660	0.662 ^b	0.292	1.693 ^a	0.401
Water fees	-3.21 x 10 ^{4a}	5.10 x 10 ⁵	-0.001 ^a	4.33 x 10 ⁵	-3.80 x 10 ^{4a}	7.86 x 10 ⁵	-3.72 x 10 ^{4a}	3.60 x 10 ⁵	-0.001 ^a	6.53 x 10 ⁵
Multiple uses × no. of cattle	0.037	0.027	0.041	0.026	0.037 ^c	0.020	0.027 ^b	0.013	0.035	0.023
Water fees × gender	-1.85 x 10 ^{4b}	8.91 x 10 ⁵	-1.61 x 10 ^{4b}	7.11 x 10 ⁵	-0.001 ^a	2.13 x 10 ⁴	-3.91 x 10 ⁵	6.20 x 10 ⁵	-2.86 x 10 ^{4a}	9.29 x 10 ⁵
Water fees × crop income	9.44 x 10 ¹⁰	1.23 x 10 ⁹	5.27 x 10 ^{9a}	1.81 x 10 ⁹	-5.92 x 10 ⁸	4.26 x 10 ⁸	-1.95 x 10 ⁹	3.06 x 10 ⁹	2.45 x 10 ^{9b}	1.17 x 10 ⁹
ASC	-0.410	0.312	-0.565 ^b	0.268	-0.695	0.554	-0.802 ^a	0.232	0.181	0.354
SD										
Membership to water organization	-0.460	0.353	-0.040	0.244	-0.557	0.527	-0.234	0.384	-0.049	0.269
Number of crops	0.898 ^a	0.149	0.596 ^a	0.092	0.854 ^a	0.240	0.789 ^a	0.098	0.560 ^a	0.120
Multiple uses	2.673 ^a	0.364	2.598 ^a	0.267	-3.119 ^a	0.640	-3.201 ^a	0.312	2.092 ^a	0.304
Number of observations	2180		3460		900		4300		2 240	
LR chi ² (4)	213.0		225.5		88.0		447.5		80.6	
Prob > chi ²	0.000		0.000		0.000		0.000		0.000	
Log likelihood	-600.3		-916.8		-219.9		-1 172.1		-553.0	

Note: ^{a, b, c} significant at the 1%, 5% and 10% levels, respectively; SD- standard deviations

Table 7. Mean WTP for water estimates

Attribute	Mean WTP					
	Pooled sample	Scheme irrigators	Non-scheme irrigators	Dryland farmers	Makhathini	Ndumo-B
Membership to water organization	-145.3 (199.8)	-854.7 (910.0)	250 (56.8)	-217.3 (499.1)	-504.5 (569.9)	62.5 (55.9)
Number of crops per season	569.8 (1 165.7)	722.1 (1 775.1)	497.8 (851.3)	287.2 (764.6)	401.6 (1 919.7)	836.9 (631.9)
Multiple uses of water	3 108.1 (4 274.3)	2 671.0 (5 286.3)	3 884.0 (3 710.5)	1 259.7 (2 793.7)	3 288.1 (7 791.2)	3 863.3 (2 362.5)

Note: Estimates in South African Rand; figures in parenthesis () are standard deviations of mean WTP

strategy so that farmers can contribute towards this cost over time. To control the use of irrigation water, the design of farmer-managed irrigation systems should include clearly defined collective use arrangements with enforceable rules and mechanisms for conflict resolution. More WTP for additional uses of irrigation water confirms why it is important to consider multiple uses of irrigation water for efficient allocation and improved water management. Ignoring this value results in the undervaluation and unsustainable utilization of the resource. More importantly, the design and management of irrigation water infrastructure (such as canals) should integrate these other uses like livestock watering. This will reduce the damage to the irrigation infrastructure and avoid unnecessary water losses.

The study demonstrates how improving agricultural production and productivity with market access will improve farmers' willingness and ability to pay for irrigation water. Thus, there is a need to support policies and programmes (such as Agri-parks) that enhance the profitability of smallholder irrigation. This will positively impact farmers' contribution towards operation and maintenance costs for the schemes. The study also reiterates the importance of smallholder agriculture to women. Supporting women farmers has positive implications for efficient and sustainable utilization of irrigation water. Policies should seek to address existing gender gaps in accessing resources. Regarding dryland farmers, more awareness creation on water scarcity and the importance of efficient and sustainable utilization of water is required.

The study only focused on two irrigation schemes and hence studies in the future should expand this research nationally and regionally. A more difficult question for future research regarding irrigation water pricing concerns the possibility of integrating other irrigation water uses into the water pricing system. Proper cost-benefit analysis on the volumetric water charging at farmer level in the schemes should also be conducted. This will assist policy makers in planning for a smooth shift in the water pricing strategy.

ACKNOWLEDGEMENTS

The study was undertaken as part of a project (K5/2278/4) funded by the Water Research Commission (WRC). The final project report can be accessed via www.wrc.org.za as 'Appropriate entrepreneurial development paths for homestead food gardening and smallholder irrigation crop farming in KwaZulu-Natal Province (WRC Report No. 2278/1/18)'. The first author would like to acknowledge the WRC for financial support of his PhD study through this project. Special mention goes to the enumerators who worked hard during the data collection.

REFERENCES

- ABU-ZEID M (2001) Water pricing in irrigated agriculture. *Int. J. Water Resour. Dev.* **17** (4) 527–538.
- ASRAT S, YESUF M, CARLSSON F and WALE E (2010) Farmers' preferences for crop variety traits: Lessons for on-farm conservation and technology adoption. *Ecol. Econ.* **69** (12) 2394–2401. <https://doi.org/10.1016/j.ecolecon.2010.07.006>
- BACKEBERG GR (2006) Reform of user charges, market pricing and management of water: problem or opportunity for irrigated agriculture? *Irrig. Drain.* **55** (1) 1–12. <https://doi.org/10.1002/ird.221>
- BARTON DN and BERGLAND O (2010) Valuing irrigation water using a choice experiment: an 'individual status quo' modelling of farm specific water scarcity. *Environ. Dev. Econ.* **15** (3) 321–340. <https://doi.org/10.1017/s1355770x10000045>
- BARTON DN and TARON A (2010) Valuing irrigation water using survey-based methods in the Tungbhadra River Basin, India. *Irrig. Drain. Syst.* **24** (3–4) 265–277. <https://doi.org/10.1007/s10795-010-9098-2>
- BECH M and GYRD-HANSEN D (2005) Effects coding in discrete choice experiments. *Health Econ.* **14** (10) 1079–1083. <https://doi.org/10.1002/hec.984>
- BHADURI A and KLOOS J (2013) Getting the water prices right using an incentive-based approach: an application of a choice experiment in Khorezm, Uzbekistan. *Eur. J. Dev. Res.* **25** (5) 680–694. <https://doi.org/10.1057/ejdr.2013.30>
- BIROL E, KAROUSAKIS K and KOUNDOURI P (2006) Using economic valuation techniques to inform water resources management: a survey and critical appraisal of available techniques and an application. *Sci. Total Environ.* **365** 105–122. <https://doi.org/10.1016/j.scitotenv.2006.02.032>
- BOELÉE E, LAAMRANI H and VAN DER HOEK W (2007) Multiple use of irrigation water for improved health in dry regions of Africa and South Asia. *Irrig. Drain.* **56** (1) 43–51. <https://doi.org/10.1002/ird.287>
- CHENG S AND LONG JS (2007) Testing for IIA in the multinomial logit model. *Sociol. Meth. Res.* **35** (4) 583–600.
- DWS (Department of Water and Sanitation, South Africa) (2015) Draft pricing strategy for water use charges. Department of Water and Sanitation, Pretoria.
- FAO (2011) The state of the world's land and water resources for food and agriculture – managing systems at risk.
- FAO (2018) Water use of livestock production systems and supply chains – Guidelines for assessment (Draft for public review). Livestock Environmental Assessment and Performance (LEAP) Partnership. FAO, Rome.
- GIERGICZNY M, VALASIUK S, CZAJKOWSKI M, DE SALVO M and SIGNORELLO G (2012) Including cost income ratio into utility function as a way of dealing with 'exploding' implicit prices in mixed logit models. *J. For. Econ.* **18** (4) 370–380. <https://doi.org/10.1016/j.jfe.2012.07.002>
- GIRALDO L, CORTIGNANI R and DONO G (2014) Simulating volumetric pricing for irrigation water operational cost recovery under complete and perfect information. *Water* **6** (5) 1204–1220. <https://doi.org/10.3390/w6051204>

- GREENE WH (2012) *Econometric Analysis*. (7th edn). Pearson Education Limited, London.
- HAUSMAN J and MCFADDEN D (1984) Specification tests for the multinomial logit model. *Econometrica* 1219–1240. <https://doi.org/10.2307/1910997>
- HEWITT K (2013) Water efficiency usage in agriculture: Evaluating the use of allocated water in food production. Undergraduate Honours thesis, University of Colorado.
- HOLE AR (2007) Estimating mixed logit models using maximum simulated likelihood. *Stata J.* 7 (3) 388–401. <https://doi.org/10.1177/1536867x0700700306>
- JAECK M AND LIFRAN R (2009) Preferences, norms and constraints in farmers' agro-ecological choices: case study using a choice experiments survey in the Rhone River Delta, France. Paper presented at: *AARES 2009 Conference*, Cairns, 11–13 February 2009.
- JOHNSON FR, LANCSAR E, MARSHALL D, KILAMBI V, MÜHLBACHER A, REGIER DA, BRESNAHAN BW, KANNINEN B and BRIDGES JF (2013) Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. *Value in Health* 16 (1) 3–13. <https://doi.org/10.1016/j.jval.2012.08.2223>
- JOZINI LOCAL MUNICIPALITY (2015) Jozini integrated development plan 2015/16. uMkhanyakhude District, KwaZulu-Natal.
- KANYOKA P, FAROLFI S and MORARDET S (2008) Households' preferences and willingness to pay for multiple use water services in rural areas of South Africa: an analysis based on choice modelling. *Water SA* 34 (6) 715–723.
- KRUGER AC and NXUMALO MP (2017) Historical rainfall trends in South Africa: 1921–2015. *Water SA* 43 (2) 285–297. <https://doi.org/10.4314/wsa.v43i2.12>
- KUHFELD WF (2010) *Marketing Research Methods in SAS*. SAS 9.2. SAS Institute, NC, USA.
- KUNIMITSU Y (2009) Measuring the implicit value of paddy irrigation water: application of RPML model to the contingent choice experiment data in Japan. *Paddy Water Environ.* 7 (3) 177–185. <https://doi.org/10.1007/s10333-009-0159-9>
- LANCASTER KJ (1966) A new approach to consumer theory. *J. Polit. Econ.* 74 (2) 132–157.
- LANGE G and HASSAN R (2007) *The Economics of Water Management in Southern Africa: An Environmental Accounting Approach*. Edward Elgar, Cheltenham, UK.
- LAYTON DF (2000) Random coefficient models for stated preference surveys. *J. Environ. Econ. Manag.* 40 (1) 21–36.
- LEE D, HOSKING S and DU PREEZ M (2014) A choice experiment application to estimate willingness to pay for controlling excessive recreational fishing demand at the Sundays River Estuary, South Africa. *Water SA* 40 (1) 39–40. <https://doi.org/10.4314/wsa.v40i1.5>
- LINIGER H, MEK DASCHI SR, HAUERT C AND GURTNER M (2011) Sustainable land management in practice: guidelines and best practices for Sub-Saharan Africa: field application. Rome: FAO.
- MANGHAM LJ, HANSON K and MCPAKE B (2009) How to do (or not to do). Designing a discrete choice experiment for application in a low-income country. *Health Polic. Plann.* 24 (2) 151–158. <https://doi.org/10.1093/heapol/czn047>
- MCFADDEN D (1973) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed.) *Frontiers in Economics*. Academic Press, New York.
- MEINZEN-DICK RS and VAN DER HOEK W (2001) Multiple uses of water in irrigated areas. *Irrig. Drain. Syst.* 15 (2) 93–98.
- MEKURIA W and MEKONNEN K (2018) Determinants of crop–livestock diversification in the mixed farming systems: evidence from central highlands of Ethiopia. *Agric. Food Security* 7 (1) 1–15. <https://doi.org/10.1186/s40066-018-0212-2>
- MUCHARA B, ORTMANN G, MUDHARA M and WALE E (2016) Irrigation water value for potato farmers in the Mooi River irrigation scheme of KwaZulu-Natal, South Africa: a residual value approach. *Agric. Water Manage.* 164 243–252. <https://doi.org/10.1016/j.agwat.2015.10.022>
- MUCHARA B, WALE E, ORTMANN GF and MUDHARA M (2014) Collective action and participation in irrigation water management: a case study of Mooi River irrigation scheme in KwaZulu-Natal Province, South Africa. *Water SA* 40 (4) 699–708. <https://doi.org/10.4314/wsa.v40i4.15>
- NJOKO S AND MUDHARA M (2017) Determinant of farmers' ability to pay for improved irrigation water supply in rural KwaZulu-Natal, South Africa. *Water SA* 43 (2) 229–237. <https://doi.org/10.4314/wsa.v43i2.07>
- OSTROM E, LAM WF and LEE M (1994) The performance of self-governing irrigation systems in Nepal. *Hum. Syst. Manage.* 13 (3) 197–207.
- RAY I (2011) Farm-level incentives for irrigation efficiency: some lessons from an Indian canal. *J. Contemp. Water Res. Educ.* 121 (1) 64–71.
- REINDERS FB, STOEP I and BACKEBERG GR (2013) Improved efficiency of irrigation water use: a South African framework. *Irrig. Drain.* 62 (3) 262–272. <https://doi.org/10.1002/ird.1742>
- RSA (Republic of South Africa) (1998) National Water Act. Act No. 36 of 1998. *Government Gazette* 19182. Government Printer, Cape Town.
- RYAN M, KOLSTAD J, ROCKERS P and DOLEA C (2012) How to conduct a discrete choice experiment for health workforce recruitment and retention in remote and rural areas: a user guide with case studies. World Health Organization & Capacity Plus: World Bank. <https://doi.org/10.2471/blt.09.068494>
- SCARPA R, THIENE M and TRAIN K (2008) Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps. *Am. J. Agric. Econ.* 90 (4) 994–1010. <https://doi.org/10.1111/j.1467-8276.2008.01155.x>
- SCHREINER B (2015) Water pricing: the case of South Africa. In: Dinar A, Pochat V and Albiac-Murillo J (eds.) *Water Pricing Experiences and Innovations*. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-16465-6_1
- SCHUR M (2000) Pricing of irrigation water in South Africa, country report. World Bank, Washington DC.
- SHARAUNGA S, MUDHARA M and BOGALE A (2016) Effects of 'women empowerment' on household food security in rural KwaZulu-Natal province. *Dev. Polic. Rev.* 34 (2) 223–252. <https://doi.org/10.1111/dpr.12151>
- SOFA TEAM and DOSS C (2011) The role of women in agriculture. FAO, Rome.
- SPEELMAN S, FRIJA A, PERRET S, D'HAESE M, FAROLFI S and D'HAESE L (2011) Variability in smallholders' irrigation water values: study in North-West Province, South Africa. *Irrig. Drain.* 60 (1) 11–19. <https://doi.org/10.1002/ird.539>
- TANG Z, NAN Z and LIU J (2013) The willingness to pay for irrigation water: a case study in Northwest China. *Glob. Nest J.* 15 (1) 76–84.
- TRAIN K (2009) *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, New York.
- YOUNG RA and LOOMIS JB (2014) *Determining the Economic Value of Water: Concepts and Methods*. (2nd edn). RFF Press, Routledge, New York, USA.

The role of water conservation strategies and benchmark ecotopes for increasing yields in South Africa's semi-arid croplands

Malcolm Hensley¹, Pieter AL le Roux¹, J Jacobus Botha^{2*} and Leon D van Rensburg¹

¹Department of Soil, Crop and Climate Sciences, University of the Free State, PO Box 339, Bloemfontein 9300, South Africa

²Agricultural Research Council - Institute for Soil, Climate and Water, Private Bag X01, Glen, 9360, South Africa

ABSTRACT

Recently published results regarding South Africa's cropping potential show that about one third of the arable land is of low potential, located mainly in semi-arid areas, with the main problem being water shortage. This is therefore an appropriate time to review priorities and procedures, for selecting benchmark ecotopes to represent marginal areas, and for research needs with regard to water conservation strategies to mitigate the problems of low yields. Relevant international principles encapsulated in the words agro-ecology, sustainability and socio-economic conditions, are discussed. Relevant new technologies are described, namely: digital soil mapping that will facilitate the identification of benchmark ecotopes; a stochastic procedure to predict rainfall intensity data from daily rainfall that will facilitate runoff predictions; a crop yield cumulative probability procedure that enables sustainability to be described quantitatively. As a case study, results from a successful field experiment using the infield rainwater harvesting production technique on benchmark ecotopes in a semi-arid area, inhabited by subsistence farmers, are presented. The objectives of the study, procedures used and the method of expressing the results are recommended as guidelines for contributing towards mitigating the problem of low crop productivity across a large portion of the arable area in South Africa.

Keywords: water conservation, benchmark ecotopes, semi-arid areas, subsistence farmers, infield rainwater harvesting

INTRODUCTION

The recently published results regarding the cropping potential of South Africa (Le Roux et al., 2016) show that at least one third of the arable land is of low potential. This land is located mainly in semi-arid areas where the main cropping problem is a water shortage due to a low and erratic rainfall pattern, high evaporation rates and often with soils having some unsatisfactory characteristics. The available estimates also indicate that in the case of a serious drought, such as in 1983/84, the production of maize, the staple food of a large fraction of the population, could easily be less than required. With regard to land evaluation in the general sense, for example, compared to grazing or forestry potential, these results dictate that land evaluation for crop production needs to have the highest priority. Le Roux et al. (2016 p. 83) also recommend that 'detailed assessment of the 15 m ha occupied by subsistence farmers needs to receive the highest priority'. This priority is also intimated in the following statement by Paterson et al. (2015 p. 6): 'The challenges begin with the need for recognition of the future central role to be played by the Natural Agricultural Resource Information System towards food security, environmental sustainability and ultimately the nation's social well-being.'

Considering the current land evaluation needs in South Africa leads to the logical conclusion that the most important current objective needs to be the identification and evaluation of ecotopes that are representative of the cropping areas of low potential, and especially those in areas occupied by subsistence farmers. These ecotopes will be described as benchmark ecotopes, equivalent to the term 'benchmark sites' used by Uehara and Tsuji (1990) for their international project titled 'International Benchmark Sites Network for Agro-technology Transfer (IBSNAT)', and as used in the following Water

Research Commission reports that appear in the reference list: Bennie et al. (1998); Hensley et al. (1997); Hensley et al. (2000).

Considerable advances have been made internationally in recent years regarding the principles of land evaluation. The following, accentuated by Bouma et al. (2012 p. 34.1–34.2), are relevant, especially the title of their paper, i.e. land evaluation for landscape units:

- 'In all cases land use is the core issue to be studied, while socio-economic considerations often play a crucial role as well'
- The focus needs to be on 'agro-ecology as related to sustainable development'
- Defining land qualities in terms of land characteristics is unsatisfactory; details in this regard are presented in Le Roux et al. (2016)
- Ongoing consultation with all stakeholders is necessary

Agro-ecology is described by Altieri (1989 p. 37–38) as a 'new research and development paradigm for world agriculture' and also states that, 'solving the sustainability problem of agriculture is the primary aim of agro-ecology'. Three aspects are emphasized: '(i) the need for a systems framework analysis; (ii) the need to focus on both biophysical and socio-economic constraints on production; (iii) the need for a suitable land unit for the analysis'. Regarding the last need, the ecotope concept of MacVicar et al. (1974) is appropriate.

The recent FAO publication on land evaluation (FAO, 2007) is a revised version of its predecessor (FAO, 1976). In FAO (2007), among the eight proposed principles of land evaluation, the three following ones are particularly relevant to the current prevailing conditions in South Africa, and accentuate similar principles to those of Bouma et al. (2012) that:

- Land evaluation requires a multi-disciplinary and cross-sectoral approach.
- Suitability refers to use or services on a sustained basis; sustainability should incorporate productivity, social equity and environmental concerns.
- Land evaluation needs to consider all stakeholders.

*Corresponding author, email: BothaC@arc.agric.za

Received 29 June 2018; accepted in revised form 19 June 2019

Addressing sustainability of agricultural production, FAO (2007) accentuates the need for information regarding risk assessment and the environmental impact of production strategies on different classes of land. Addressing social equity, it is also stated (p. 1) that ‘most current rural development is directed at areas where people face economic and social problems, in particular hunger and poverty’. The importance of reliable land evaluation procedures is accentuated, especially in developing countries due to the increasing pressure on land resources largely due to rapid increases in population.

Altieri and Nicholls (2005 p. 99) present a convincing motivation for international attention to the needs of subsistence farmers with these words: ‘Throughout the developing world, resource-poor farmers (about 1.4 billion people) located in risk-prone, marginal environments, remain untouched by modern agricultural technology’. Equivalent estimates for South Africa in 1985 were reported by Fényes (1985 p. 15) as follows: ‘The predominantly subsistence-orientated African agricultural sector in South Africa presently occupies 15.1 million ha of land, 14% of which is cultivated and employs 1.1 million (1970) people’. Van der Merwe (1985) also reported a similar area of land occupied by subsistence farmers, and that it was, at that stage, relatively underutilized in terms of its potential. Cousins (2015) estimates that the number of market-orientated smallholder farmers in communal areas in South Africa and in land reform context amount to between 200 000 and 250 000; and, in addition, that subsistence-orientated smallholder farmers growing food for themselves, and selling occasionally, amount to between 2 and 2.5 million. Data about the arable area of land used by these two groups of farmers is not currently available.

NEW LAND EVALUATION TECHNOLOGIES

Identifying benchmark crop ecotopes and improving the procedure for estimating areas

This is a subject of ongoing importance in South Africa because of the lack of permanent food security for all its people. Reliable information about the area of marginal crop ecotopes, and what can be done to mitigate the problems, is needed by decision makers. The provisional procedure for identifying ecotopes using land type survey (LTS) data in the form of climate-soil-slope units, described by Schoeman and MacVicar (1978), was a valuable first approximation that served a useful purpose. Because it did not involve actual measurements, it was subjective in nature and could therefore have resulted in considerable errors. Van Zijl (2013) has developed a procedure using modern technology in the form of digital soil mapping (DSM) that could greatly facilitate and improve the efficiency of this step. An example is presented by Van Zijl et al. (2013) of their procedure, and termed ‘disaggregation of land types using terrain analysis, expert knowledge and GIS methodology’. It is recommended that this procedure, developed further with careful testing, be used for the urgently needed identification of important benchmark ecotopes.

Runoff prediction models

Although the semi-arid areas of South Africa have the disadvantage of low rainfall, they nevertheless have the advantage that a significant portion of the rain is in the form of storms of reasonably high intensity (Pi). Therefore, by harvesting as much as possible of the runoff water from these

events, i.e., those with $P_i >$ the final infiltration rate of the soil (Pf), maximum use can be made of the limited rainfall. In a Water Research Commission (WRC) project aimed at contributing to infield rainwater harvesting (IRWH) studies, Walker and Tsubo (2003) describe a stochastic procedure developed in the USA that enables the prediction of Pi data for rainfall stations with long-term daily rainfall (Pd) data and relatively short-term Pi data. They used measured rainfall intensity data for 30 years for Bloemfontein (20 km from Glen) to develop and successfully calibrate the model. Both Zere et al. (2005) and Anderson (2007) used the model, together with carefully selected input parameters for the Morin and Cluff (1980) mechanistic runoff model, to simulate the results of a long-term (18 years) runoff experiment by Du Plessis and Mostert (1965) at Glen on a soil of Tukulu form. Their objective was to test with what degree of reliability the combination of the two models could predict the measured runoff results obtained by Du Plessis and Mostert (1965). Willmot (1982) statistics were used to determine the reliability of the predictions and found by both researchers to be satisfactory. These results were only obtained after the field experiments on the two benchmark ecotopes at Glen (Botha, 2006) had been completed. Anderson (2007) showed, however, that the original runoff procedure used in the crop model titled ‘crop yield prediction for semi-arid areas’ CYP-SA (Botha, 2006) gave similar yield prediction results to those that could be obtained using the Morin and Cluff model together with predicted Pi data. Useful results were also obtained using Pi data and the Morin and Cluff (1980) runoff model for three benchmark ecotopes in Ethiopia by Welderufael (2006) and Welderufael et al. (2008; 2009).

Considering the future need to evaluate the productivity of many marginal benchmark ecotopes in semi-arid areas, if the use of the combination of the two models described above proves to be consistently reliable, this strategy offers a useful means of estimating the yield benefits obtainable with the in-field rainwater harvesting (IRWH) production technique before conducting expensive field experiments on an unnecessary number of benchmark ecotopes.

Appropriate procedures for expressing crop productivity

The importance of sustainability is relevant, as stressed by Bouma et al. (2012) and FAO (2007). For marginal ecotopes in semi-arid areas this consideration is of particular importance because of the low and variable rainfall coupled with high evaporation rates. Reliable recent contributions in this regard are the use of appropriate kinds of crop models to predict long-term yields with different production techniques, coupled with results expressed in the form of cumulative probability functions (CPF's). The basic procedure is described in detail by Uehara and Tsuji (1990), and has been widely and effectively used, e.g., Monteith and Virmani (1990) in India; Muchow et al. (1990) in Australia; Bouma and Droogers (1998) in The Netherlands; Hensley et al. (2000) in South Africa; Botha (2006) in South Africa.

The choice of the type of crop model to use for the CPF procedure is important. Passioura (1996) differentiates between ‘scientific’ and ‘engineering’ crop models, and provides detailed evidence why he favours the latter for solving practical problems and providing sound management advice. He states furthermore (p. 690) that ‘the best engineering models are based on robust empirical relations between plant behaviour

and the main environmental variables'. In support of the views of Passioura (1996), Penning de Vries and Spitters (1990 p. 123) state that, 'Summary models and regression models derived from simulation results, are ultimately ideal tools for application and predictive purposes'. These authors also indicate that for water-limiting environments, adherence to these criteria is of additional importance. A local example in support of this contention is available in the improvement made by Botha (1997) with regard to the crop-stress algorithm used for maize in the PUTU crop model (De Jager, 1990). Although a specific algorithm was used in PUTU before 1997 for a specific crop, soil differences in different ecotopes were not taken into account. Careful and detailed measurements in a maize crop, growing in a Glen/Sepane ecotope as the soil dried out and stress increased, showed that an improved crop-stress algorithm was needed on this type of ecotope, i.e., high clay duplex soil with a semi-arid climate. With the improved algorithm the performance of the model improved significantly.

The crop model SWAMP (soil water management programme; Bennie et al., 1998) and CYP-SA (Botha 2006) are of the type described by Passioura as 'engineering models'.

BENCHMARK ECOTOPES

The ecotope concept

The ecotope concept is described in detail by MacVicar et al. (1974). The main part of the definition is described as: An ecotope is a class of land defined in terms of its macro-climate (including where necessary, aspect), soil and soil surface characteristics (mainly slope) such that, in terms of the farming enterprises that can be carried out on it, the potential yield class for each enterprise, or the production techniques needed for each enterprise, there is a significant difference between one ecotope and any other. The original agro-ecological nature of the ecotope concept is revealed by the second part of this definition, and also by the initial procedure used by Schoeman and MacVicar (1978) for identifying ecotopes using LTS data in the form of climate-soil-slope units and obtaining crop yield estimates for each unit by consulting with knowledgeable and experienced local people.

Currently the most important ecotopes in South Africa are the benchmark ones representing cropping areas of low potential, and especially those in areas occupied by subsistence farmers. The benchmark ecotope concept is not new in South African literature associated with land evaluation. It played a prominent role in the following publications: Hensley et al. (1997); Bennie et al. (1998); Hensley et al. (2000).

Selecting benchmark crop ecotopes

Important sources of information for this process are the supporting database of around 2 500 modal soil profiles described and analysed for the LTS (Paterson et al., 2015), together with the detailed soil knowledge gained by all those intensively involved over a long period with the survey. This knowledge is expressed in the numerous LTS memoirs and in the following publications by soil scientists involved: Verster (1974); Schoeman and MacVicar (1978); Eloff (1984); Ellis (1984); Turner (2000); Schoeman et al. (2002); Turner et al. (2014). The tacit knowledge of these people is also valuable as most of them are fortunately still available for consultation purposes.

Information from the LTS will indicate in which particular land types, or group of land types, such as the relatively

homogenous farming areas defined by Scheepers et al. (1984) for the Highveld Region, and the bio-resource units defined by Camp (1999) for KwaZulu-Natal, the sought-after benchmark ecotopes will be located. For these areas, it is important that diligent consultation with relevant expert knowledge and local farmers be carried out to obtain reliable information about potential yields and their sustainability. This procedure was employed by Turner et al. (2014) for the survey of Eastern Cape Province. It yielded valuable information with regard to proposed suitability class maize yields for subsistence farmers, considerably lower than those normally adopted for commercial farming. The importance of the fact that the socio-economic parameters for subsistence farming are very different from those for commercial farming, as explained by Kundhlande et al. (2004), is accentuated.

CASE STUDY

Land type Dc17, with an area of 239 080 ha (Eloff, 1984), is located east of Bloemfontein. It is a semi-arid area characterized by a low and variable rainfall, high evaporative demand, and predominantly duplex soils or soils with marginalitic A horizons, both being unfavourable for crop production with the prevailing climate. Eloff (1984) estimated the fraction of arable land in Dc17 as 10%, all with low cropping potential. As Dc17 is occupied mainly by subsistence farmers who often experience a lack of food security, it was considered a suitable area for identifying benchmark crop ecotopes and appropriate research to attempt to find solutions to the problem of low productivity. Two ecotopes similar to those in Dc17 were identified on the Glen Research Station, a suitable site for long-term field experiments. The soils of these ecotopes are (Soil Classification Working Group, 1991): Swartland form/Rouxville family (Glen/Sw); and Bonheim form/Onrus family (Glen/Bo). Detailed profile descriptions and analytical data for these soils are presented in Hensley et al. (2000).

The hypothesis for the proposed experiment was that with conventional tillage (CON), an unnecessary loss of water occurs by runoff (R) and evaporation (Es) from the rough soil surface. Therefore, if R could be reduced to zero and Es also significantly reduced, crop yield would be significantly increased (Hensley et al. 2000; van Rensburg et al. 2010). Regarding Es, detailed quantitative measurements by Schwartz et al. (2010), comparing losses with tilled and untilled surfaces in a semi-arid area, showed that the former were significantly greater than the latter throughout a season. The instrumentation used was time-domain reflectometry and a neutron water meter.

To test our hypothesis at Glen, the IRWH production technique depicted in Fig. 1 was proposed. The plan was that the only mechanical cultivation on the land would be that needed to establish the basins shown in Fig. 1. The cultivation consisted of one deep ploughing on the contour with soil deposited onto the downslope side. Cross walls were made in the furrows at 3 m intervals to form basins. Maize rows were located as shown in Fig. 1. No-till was maintained on the smooth crusted runoff strips and weeds were controlled with chemical sprays. CON started with ploughing, followed by disking for seedbed preparation, and about 1 month after germination a shallow cultivation was given for weed control. Plant populations, planting dates and fertilisation were always the same as on the IRWH plots.

The first IRWH vs. CON experiments with maize and sunflower on both ecotopes were conducted over 3 growing

seasons (1996–1999). They were followed by similar experiments on Glen/Bo over 4 seasons. Measured results for critical parameters needed for the experiments are presented in Fig. 2 and Table 1, with measured maize yields and other relevant data in Tables 2 and 3. Statistical results show that out of 10 comparisons made, IRWH gives significantly better yields than CON in 9 cases. The failure for the 1996/97 season on the Glen/Bo IRWH plots was logical because it had not been possible for the first year to complete the construction of basins

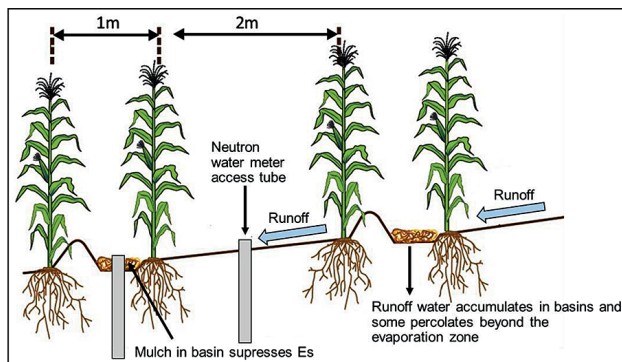


Figure 1. A diagrammatic representation of the in-field rainwater harvesting technique (modified from Botha, 2006)

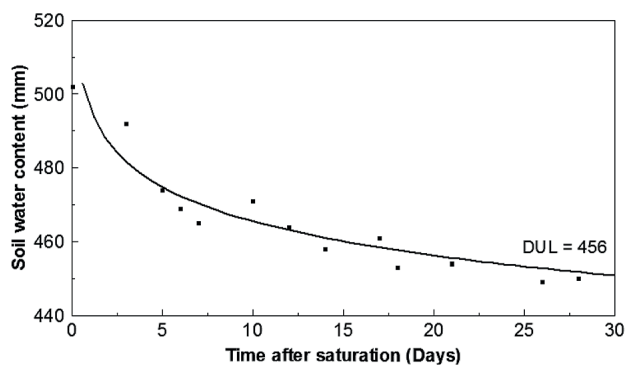


Figure 2. Drainage curve for Glen/Bo (data from Botha; 2006)

on that treatment before the start of the spring rains. The crop was therefore denied the benefit of the additional runoff water during the spring rains.

Soil water content measurements were by neutron water meter access tubes, replicated many times, as shown in Fig. 1. The measured effective root zone for maize and sunflower on both ecotopes was 1 200 mm.

Equation 1 provides a mathematical description of the curve and enables the drainage rate at any time after field saturation (F_{Sat}) to be calculated, and to estimate deep drainage after periods of heavy rain. The value given for F_{Sat} is the measured value at field saturation, and not the intercept on the vertical axis by the curve's equation.

$$\text{Glen/Bo } (F_{Sat} = 512 \text{ mm}): Y = 512 - 8.92 (\ln t) \quad r^2 = 0.91 \quad (1)$$

where:

Y = water content of the root zone (mm)

t = time (h) after drainage starts at a root zone water content of F_{Sat}

The drainage curve for the Glen/Sw has a similar shape to that of Glen/Bo, with the resultant curve described by the following equation (Van Staden, 2000):

$$\text{Glen/Sw } (F_{Sat} = 433 \text{ mm}): Y = 433 - 11.43 (\ln t) \quad r^2 = 0.80 \quad (2)$$

The data in Table 1 shows that the pedocutanic horizons in both ecotopes have not impaired root distribution to a significant extent. For maize on Glen/Sw the fraction of TESW from the 0–600 mm layer (A, B1 and B2 horizons) is 57%, and for the 600–1 200 mm layers (B3, B4 and C horizons) is 43%; and the equivalent results for the Glen/Bo are 68% and 32%, respectively. The considerably higher TESW values for Glen/Bo are due to the higher content of clay and silt compared to Glen/Sw. It is expected that the maintenance of a relatively higher soil water content in the IRWH plots, compared to CON plots, will have had a beneficial effect throughout each season on TESW, and therefore on yield. The wide variation in seasonal rainfall, a characteristic of semi-arid areas in South Africa, and hence its influence on yield, is clearly shown by the data in Tables 2 and 3. This factor accentuates the importance with regard to

Table 1. The soil water extraction properties of the Glen/Sw (data from Hensley et al., 2000) and Glen/Bo (data from Botha; 2006)

Ecotope	Profile detail						Maize		Sunflower		
	Diagnostic horizon ^{*1}	Colour	Clay (%)	Silt (%)	Bd ^{*2} (Mg m ⁻³)	Lower depth (mm)	DUL (mm)	LL (mm)	TESW ^{*3} (mm)	LL (mm)	TESW (mm)
Glen/Sw	ot	DkBr	38	7	1.50	300	82	33	49	23	59
	vp	DkBr	40	7	1.66	600	96	67	29	62	34
	vp	DkBr	44	9	1.51	900	96	62	34	62	34
	vp & so ^{*1}	DkBr & mottled ^{*4}	35	7	1.46	1 200	84	60	24	60	24
					Total	358	222	136	207	151	
Glen/Bo	ml	DkBr	45	10	1.30	300	122	39	83	45	77
	vp	DkBr	43	12	1.45	600	123	74	49	67	56
	vp	DkBr	40	14	1.45	900	106	74	32	67	39
	so	mottled ^{*4}	38	20	1.45	1 200	105	76	29	61	44
					Total	456	263	193	240	216	

^{*1} Abbreviations from Soil Classification Working Group (1991): ot = orthic; vp = pedocutanic so = saprolite; ml = melanic

^{*2} Bulk density

^{*3} Total extractable soil water, calculated as the measured drained upper limit of available water (DUL) minus the measured lower limit of plant available water (LL)

^{*4} Due to CaCO₃ concretions

the sustainability of any new production technique that is being tested in an attempt to ameliorate socio-economic conditions where food security is a problem. Hence, the value of a reliable crop model and the CPF procedure for expressing long-term yields (Fig. 3).

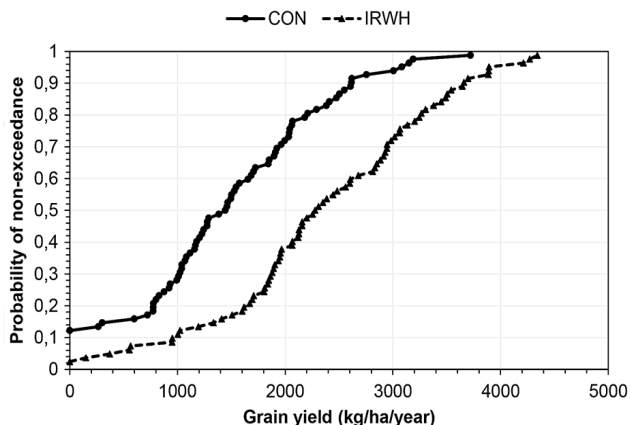


Figure 3. CPF graphs of long-term maize yields simulated with CYP-SA (80 seasons, 1922–2003), on the Glen/Bo ecotope with CON and IRWH (Botha, 2006)

The final condensed results of the experiment to test the hypothesis IRWH vs CON, are presented in the crop yield CPF in Fig. 3, and the yields and statistical results in Tables 2 and 3. The long-term superiority of IRWH is clearly demonstrated, indicating that there is generally a 50% probability on the Glen/Bo ecotope of achieving sustainable maize grain yields of approximately 1 300 kg·ha⁻¹ and 2 300 kg·ha⁻¹ with the CON and IRWH tillage treatments, respectively. For maize on the Glen/Sw the equivalent figures are very similar. The hypothesis for the ecotopes is therefore proved to be valid, and IRWH shown to be a suitable production technique for subsistence farmers in the Thaba Nchu area. Soils of the Bonheim, Sepane, Swartland and Valsrivier forms in this area with rooting depth at least 900 mm, should be considered as arable using IRWH where the aridity index (AI) value is at least 0.25, and the slope approximately 5%.

CONCLUSION AND RECOMMENDATIONS

Considering the current land evaluation needs in South Africa it is logical that the highest priority should now be given to the identification of the most important benchmark crop ecotopes that require research attention in order to seek possible remedial actions. Also important is the identification of arable regions in land types where crop ecotopes occur that are

Table 2. Maize grain yields on the Glen/Sw and Glen/Bo ecotopes for three seasons comparing results for the CON and IRWH treatments. Data from Hensley et al. (2000)

Season	Ecotope	Treatment ^{*1}	P ^{*2} (mm)	Maize grain yield (kg·ha ⁻¹)	Stat. test ^{*3}	R ^{*4} (mm)
96/97	G/Sw	CON	297	1 138	*	28
		IRWH		1 917		0
	G/Bo	CON	297	2 282	NS	14
		IRWH		2 274		0
97/98	G/Sw	CON	451	3 187	*	41
		IRWH		5 308		0
	G/Bo	CON	451	3 133	*	33
		IRWH		4 678		0
98/99	G/Sw	CON	208	41	*	0
		IRWH		234		0
	G/Bo	CON	208	0	*	0
		IRWH		132		0

^{*1} CON = conventional tillage; IRWH = infield rainwater harvesting

^{*2} Precipitation during the crop growing period

^{*3} Statistical significance at the 5% level indicated by *

^{*4} Runoff during the crop growing period

Table 3. Maize grain yields on the Glen/Bo ecotope for 4 seasons 1999–2003; comparing results for the CON and IRWH treatments. Data from Botha (2006).

Season	Treatment	P ^{*1} (mm)			Maize grain yield (kg·ha ⁻¹)	Stat. test ^{*2}	R ^{*3} (mm)
		F	G	T			
99/00	CON	157	228	385	3 093	*	31
	IRWH				3 455		0
00/01	CON	233	281	514	1 489	*	56
	IRWH				2 543		0
01/02	CON	360	247	607	1 521	*	64
	IRWH				3 281		0
02/03	CON	315	215	530	459	*	52
	IRWH				2 401		0

^{*1} Precipitation for: F = fallow period; G = crop growing period; T = overall production period

^{*2} Statistical significance at the 5% level indicated by *

^{*3} Runoff during the overall production period, i.e. crop growing season + preceding fallow period

similar to the selected benchmark crop ecotopes.

We propose that in order for a benchmark ecotope to become a registered 'habitat pigeon hole' in the South African crop potential database in which 'production yield data can be stored' (MacVicar et al., 1974; Paterson et al., 2015), the requirements listed below need to be met.

1. A name for the benchmark ecotope consisting of 4 words in the following order: a place name indicating the approximate geographical location; soil form; soil family; soil series
2. Detailed profile description as in the modal profiles of the LTS, and in accordance with the format proposed by Turner (1991); plus analytical data for each horizon of the solum, as in modal profiles for the LTS
3. Drainage curve of the root zone (as in Fig. 2)
4. Soil water extraction properties of the root zone for each main crop expected to be grown on the benchmark ecotope (as in Table 1)
5. Detailed long-term daily climate data, including as much rainfall intensity data as possible
6. Expected long-term yield potential with specific crop/crops, using specified production techniques. The yield potential should be expressed in terms of a CPF of long-term yields so that the risk is quantified. The latter is of particular importance for marginal crop ecotopes in semi-arid areas. To obtain the necessary data for this, a field experiment over at least three seasons with at least two of the main crops for the region, using the most appropriate production techniques available at the time, will be needed.
7. To enable long-term yield predictions obtained with a crop model to be determined, measured values for all the parameters required by the selected crop model need to be provided for each benchmark crop ecotope.

ACKNOWLEDGEMENTS

Appreciation is expressed for funding provided for many IRWH projects by the Water Research Commission (Sanewe and Backeberg, 2012) and for assistance from the Free State Department of Agriculture for providing suitable land at Glen for the experiments and for the loan of farm equipment for field experimentation at Glen over many years.

REFERENCES

- ALTIERI MA (1989) Agro-ecology: A new research paradigm for world agriculture. *Agric. Ecosyst. Environ.* **27** 37–41.
- ALTIERI MA and NICHOLLS CI (2005) *Agro-Ecology and the Search for a Truly Sustainable Agriculture*. Basic Textbooks for Environmental Training. University of California, Berkeley, USA.
- ANDERSON JJ (2007) Rainfall-runoff relationships and yield responses of maize and dry beans on the Glen/Bonheim ecotope using conventional tillage and in-field rainwater harvesting. PhD Agric. Thesis, University of the Free State, Bloemfontein.
- BENNIE ATP, STRYDOM MG and VREY HS (1998) The application of computer models for agricultural water management at ecotope level [in Afrikaans, with English abstract]. WRC Report No. 625/1/98. Water Research Commission, Pretoria.
- BOTHA JJ (1997) Quantifying of the stress curve for selected crop ecotopes [in Afrikaans]. MSc. Agric. thesis, University of the Free State, Bloemfontein.
- BOTHA JJ (2006) Evaluation of maize and sunflower in a semi-arid area using in-field rainwater harvesting. PhD Agric. thesis, University of the Free State, Bloemfontein.
- BOUMA J and DROOGERS P (1998). A procedure to derive land quality indicators for sustainable agricultural production. *Geoderma* **85** 103–110. [https://doi.org/10.1016/s0016-7061\(98\)00031-7](https://doi.org/10.1016/s0016-7061(98)00031-7)
- BOUMA J, STOOORVOGEL JJ and SONNEVELD MPW (2012) Land evaluation for landscape units. In: Huang PM, Li Y and Sumner ME (eds) *Handbook of Soil Sciences: Properties and Processes* (2nd edn). CRC Press, Boca Raton. 34-1–34-22.
- CAMP KGT (1999) A bio-resource classification for KwaZulu-Natal, South Africa. MSc. Agric. thesis, University of KwaZulu-Natal, Pietermaritzburg.
- COUSINS B (2015) Through a glass darkly: towards agrarian reform in South Africa. In: Ben Cousins and Cheryl Walker (eds.) *Land Divided, Land Restored, Land Reform in South Africa for the 21st Century*. Jacana, Auckland Park. 250–269.
- DE JAGER JM (1990) Maize production risk in the RSA [in Afrikaans]. Report to the Department of Agricultural Development. Dept. Agrometeorology, University of the Free State. 85 pp.
- DU PLESSIS MC and MOSTERT JWC (1965) Runoff and soil losses at the Agricultural Research Centre, Glen. [Afrikaans, with English summary]. *S. Afr. J. Agric. Sci.* **8** 1051–1061.
- ELLIS F (1984) The soils of the Karoo [in Afrikaans]. PhD thesis, University of Stellenbosch, Stellenbosch.
- ELOFF JF (1984) The soil resources of the Free State Province [in Afrikaans]. PhD thesis, University of Stellenbosch, Stellenbosch, South Africa.
- FAO (Food and Agriculture Organization of the United Nations) (1976) A framework for land evaluation. *Soils Bulletin* 32. FAO, Rome.
- FAO (Food and Agriculture Organization of the United Nations) (2007) Land evaluation. Towards a revised framework. Land and water discussion paper No. 6. FAO, Rome.
- FÉNYES TI (1985) An assessment of the present situation of subsistence farmers in southern Africa: implications for the future. In: *Proceedings of the 23rd Annual Conference of the Agricultural Economics Association of South Africa*, 30 April–1 May 1985, Bloemfontein, South Africa.
- HENSLEY M, ANDERSON JJ, BOTHA JJ, VAN STADEN PP, SINGELS A, PRINSLOO, M and DU TOIT A (1997) Modelling the water balance on benchmark ecotopes. WRC Report No. 508/1/97. Water Research Commission, Pretoria.
- HENSLEY M, BOTHA JJ, ANDERSON JJ, VAN STADEN PP and DU TOIT A (2000) Optimizing rainfall use efficiency for developing farmers with limited access to irrigation water. WRC Report No. 878/1/00. Water Research Commission, Pretoria.
- KUNDHLANDE G, GROENEWALD DG, BAIPHETHI MN, VILJOEN MF, BOTHA JJ, VAN RENSBURG LD and ANDERSON JJ (2004) Socio-economic study on water conservation techniques in semi-arid areas. WRC Report No. 1267/1/04. Water Research Commission, Pretoria.
- LE ROUX PAL, HENSLEY M, VAN RENSBURG LD and BOTHA JJ (2016) The cropping potential of South Africa: land evaluation results obtained during the last 50 years. *S. Afr. J. Plant Soil* **33** (2) 83–88. <https://doi.org/10.1080/02571862.2014.981878>
- MACVICAR CN, SCOTNEY DM, SKINNER TE, NIEHAUS HS and LOUBSER JH (1974) A classification of land (climate, terrain form, soil) primarily for rain fed agriculture. *S. Afr. J. Agric. Ext.* **3** 21–24.
- MONTEITH JL and VIRMANI SM (1990) Quantifying climatic risk in the semi-arid tropics: ICRISAT experience. In Muchow DC and Bellamy JA (eds.). *Proceedings of the International Symposium on Climatic Risk in Crop Production: Models and Management for the Semi-Arid Tropics and Subtropics*, Brisbane, Australia, July 1990.
- MORIN J and CLUFF CB (1980) Runoff calculations on semi-arid watersheds using a rotadisk rain simulator. *Water Resour. Res.* **16** 1085–1093. <https://doi.org/10.1029/wr016i006p01085>
- MUCHOW RC, HAMMER GL and CARBERRY PS (1990) Optimising crop and cultivar selection in response to climatic risk. In: Muchow DC and Bellamy JA (eds). *Proceedings of the International Symposium on Climatic Risk in Crop Production: Models and Management for the Semi-Arid Tropics and Subtropics*; Brisbane, Australia, July 1990.
- PASSIOURA JB (1996) Simulation models: science, snake oil, education or engineering? *Agron. J.* **88** 690–694. <https://doi.org/10.2134/agronj1996.00021962008800050002x>
- PATERSON DG, TURNER DP, WIESE LD, VAN ZIJL GM, CLARKE

- CE and VAN TOL JJ (2015) Spatial soil information in South Africa: Situational analysis, limitations and challenges. *S. Afr. J. Sci.* **111** (516) Art # 2014-0178, 7 pp. <http://dx.doi.org/10.17159/sajs.2015/20140178>
- PENNING DE VRIES FWT and SPITTERS CJT (1990) The potential for improvement in crop simulation. In: Muchow DC and Bellamy JA (eds) *Proceedings of the International Symposium on Climatic Risk in Crop Production: Models and Management for the Semi-Arid Tropics and Subtropics*; Brisbane, Australia, July 1990.
- SANEWE AJ and BACKEBERG GR (2012) Overview of research on rainwater harvesting and conservation by the Water Research Commission. In: Van Rensburg LD (ed.) *Rainwater Harvesting and Conservation techniques for improving household food security. Irrigation and Drainage*: Volume 61, Suppl. 2 October 2012. 1–6. <https://doi.org/10.1002/ird.1691>
- SCHEEPERS JJ, SMIT JA and LUDICK BP (1984) An evaluation of the agricultural potential of the Highveld region in terms of dryland cropping and livestock production. Technical Communication no. 185. Department of Agriculture, Pretoria.
- SCHOEMAN JL and MACVICAR CN (1978) A method for evaluating and presenting the agricultural potential of land at regional scales. *Agrochemophysica* **10** 25–37.
- SCHOEMAN JL, VAN DER WALT M, MONNIK KA, THACKRAH J, MALHERBE J and LE ROUX RE (2002) Development and application of a land capability classification system for South Africa. ARC-ISCW Report no. GW/A/2000/5T. ARC-Institute for Soil, Climate and Water, Pretoria.
- SCHWARTZ RC, BAUMHARD RL and EVETT SR (2010) Tillage effects on soil water distribution and bare soil evaporation throughout a season. *Soil & Till. Res.* **110** 221–229. <https://doi.org/10.1016/j.still.2010.07.015>
- SOIL CLASSIFICATION WORKING GROUP (1991) Soil classification – A taxonomic system for South Africa. *Mem. Agric. Natl. Resour. S. Afr. No. 15*. Department of Agricultural Development, Pretoria.
- TURNER DP (Ed) (1991) A procedure for describing soil profiles. Soil and Irrigation Research Institute Report No. GB/A/91/67. Soil and Irrigation Research Institute, Pretoria.
- TURNER DP (2000) Soils of KwaZulu-Natal and Mpumalanga. PhD thesis, University of Pretoria.
- TURNER DP, CHIRIMA JG, PATERSON DG, CARSTENS JP, SEABI FT and MUSHIA N (2014) Integrated food and nutrition security initiative. Mapping of crop producing areas in the Eastern Cape Province. ARC-ISCW Report No: GW/A/2014/23. ARC Institute for Soil, Climate and Water, Pretoria.
- UEHARA G and TSUJI GY (1990) Progress in crop modelling in the IBSNAT project. In: Muchow DC and Bellamy JA (eds). *Proceedings of the International Symposium on Climate Risk in Crop production: Models and management for Semi-Arid Tropics and Subtropics*, Brisbane, Australia, July 1990.
- VAN DER MERWE FJ (1985) Food production in Southern Africa and the limiting factors [in Afrikaans]. Priorities up to the year 2000. Unpublished report. South African Academy for Science and Arts, Pretoria.
- VAN RENSBURG LD, BOTHA JJ, ANDERSON JJ and HENSLEY M (2012) The nature and function of the in-field rainwater harvesting system to improve agronomic sustainability. In L.D. van Rensburg (Editor). *Rainwater Harvesting and Conservation techniques for improving household food security. Irrigation and Drainage*: Volume 61, Suppl. 2, October 2012. 34–40. <https://doi.org/10.1002/ird.1678>
- VAN STADEN PP (2000) The influence of runoff-water harvesting relationships on water use by maize on high-risk clay soils. [in Afrikaans]. MSc. Agric. thesis, University of the Free State, Bloemfontein.
- VAN ZIJL GM (2013) Developing a digital soil mapping protocol for Southern Africa using case studies. PhD thesis, University of the Free State, Bloemfontein.
- VAN ZIJL GM, LE ROUX PAL and TURNER DP (2013) Disaggregation of land types using terrain analysis, expert knowledge and GIS methods. *S. Afr. J. Plant Soil* **30** (3) 123–129. <https://doi.org/10.1080/02571862.2013.806679>
- VERSTER E (1974) The soils of the Transvaal region [in Afrikaans]. PhD thesis, University of the Free State, Bloemfontein.
- WALKER S and TSUBO M (2003) Estimation of rainfall intensity for potential crop production on clay soil with in-field water harvesting practices in a semi-arid area. WRC Report No. 1049/1/02. Water Research Commission, Pretoria.
- WELDERUFAEL WA (2006) Quantifying rainfall-runoff relationships on selected benchmark ecotopes in Ethiopia: a primary step in water harvesting research. PhD thesis, University of the Free State, Bloemfontein.
- WELDERUFAEL WA, LE ROUX PAL and HENSLEY M (2008) Quantifying rainfall-runoff relationships on the Dera Calcic Fluvic Regosol ecotope in Ethiopia. *Agric. Water Manage.* **95** (11) 1223–1232. <https://doi.org/10.1016/j.agwat.2008.04.007>
- WELDERUFAEL WA, LE ROUX PAL and HENSLEY M (2009) Quantifying rainfall-runoff relationships on the Melkassa Hypo Calcic Regosol ecotope in Ethiopia. *Water SA* **35** (5) 639–648. <https://doi.org/10.4314/wsa.v35i5.49189>
- WILLMOTT CJ (1982) Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.* **63** 1309–1313.
- ZERE TB, VAN HUYSTEEN CW and HENSLEY M (2005) Estimation of runoff at Glen in the Free State Province of South Africa. *Water SA* **31** 17–21. <https://doi.org/10.4314/wsa.v31i1.5116>

Developing a fuzzy logic model for predicting soil infiltration rate based on soil texture properties

Ahmed Z Dewidar^{1*}, Hussein Al-Ghobari¹ and Abed Alataway²

¹Agricultural Engineering Department, King Saud University, Riyadh 11451, Kingdom of Saudi Arabia

²PSIPW Chair, Prince Sultan Institute for Environmental, Water and Desert Research, King Saud University, Kingdom of Saudi Arabia

The prediction of the soil infiltration rate is advantageous in hydrological design, watershed management, irrigation, and other agricultural studies. Various techniques have been widely used for this with the aim of developing more accurate models; however, the improvement of the prediction accuracy is still an acute problem faced by decision makers in many areas. In this paper, an intelligent model based on a fuzzy logic system (FLS) was developed to obtain a more accurate predictive model for the soil infiltration rate than that generated by conventional methods. The input variables that were considered in the fuzzy model included the silt and clay contents. The developed fuzzy model was tested against both the observed data and multiple linear regression (MLR). The comparison of the developed fuzzy model and MLR model indicated that the fuzzy model can simulate the infiltration process quite well. The coefficient of determination, root mean square error, mean absolute error, model efficiency, and overall index of the fuzzy model were 0.953, 1.53, 1.28, 0.953, and 0.954, respectively. The corresponding MLR model values were 0.913, 2.37, 1.92, 0.913, and 0.914, respectively. The sensitivity results indicated that the clay content is the most influential factor when the FLS-based modelling approach is used for predicting the soil infiltration rate.

Keywords: FLS, infiltration rate, MLR, modelling, sensitivity analysis

INTRODUCTION

Soil and water are the main natural resources that can be used in a crop production system. The process of the percolation of rainwater or irrigation water into the soil is commonly known as infiltration. Increased infiltration control would aid in solving such wide-ranging problems as upland flooding, declining water tables, surface and groundwater pollution, inefficient irrigation of agricultural lands, and wastage of useful water (Hajiaghahi et al., 2014; Singh et al., 2018). Soil infiltration measurement is an important indicator of irrigation and drainage efficiency and is used for optimizing water availability for plant growth and metabolism, improving crop yield, and reducing erosion (Patle et al., 2018).

Infiltration rate (IR) is defined as the rate at which filtration occurs through the soil (Vand et al., 2018). Soil IR is the main factor that affects surface irrigation uniformity and efficiency owing to its mechanism of distributing water from the surface to the soil profile (Hajiaghahi et al., 2014). The IR is affected by a number of parameters, such as rainfall, vegetation cover, initial soil moisture, and fertilization and physical characteristics of the soil (Suryoputro et al., 2018). The physical characteristics of soil are influenced by the soil textures, which comprise mineral particles including sand, silt, and clay (Haghnazari et al., 2015). Clay particles are important because of their small size which allows them to fill the voids between the larger particles, and their charge orientation which makes them play a crucial role in binding the soil matrix into larger structures (Haghnazari et al., 2015). Several studies have focused on investigation of the physical characteristics of soil and their relationship with infiltration. Some of these have been focused on porosity (Bouma, 1982; Smettem, 1987), whereas others have concentrated on aggregation and structure (Helming et al., 1998), soil texture (Doerr et al., 2000; Ramos et al., 2003), and organic matter (Lado et al., 2004).

Field measurements of soil infiltration require considerable time and are costly. In contrast, infiltration models could significantly reduce the required time and cost (Mudiare and Adewumi, 2000). Many researchers have developed various conventional models to estimate the IR (Kostiakov, 1932; Philip, 1957; Sihag et al., 2017). Alternatively, the IR can be modelled using soft computing techniques such as the artificial neural network, adaptive neuro-fuzzy inference system, and fuzzy logic system (FLS) approaches (Singh et al., 2018; Singh et al., 2019). The conventional models are site specific and require model parameters, whereas the soft computing-based models are used generally for the study area. FLS is one of the methods that has been used as a dominant tool in solving and overcoming water resource problems (Azamathulla et al., 2016; Kisi et al., 2017; Kumar et al., 2018). The main advantage of fuzzy logic is its ability to deal with uncertain data in the form of reasonably continuous categories (Metternicht, 2001) and to accomplish more flexible knowledge-based modelling (Tran et al., 2002).

The fuzzy logic model (Zadeh, 1965) is a logical mathematical procedure based on the 'IF-THEN' rule system, which allows the human thought process to be reproduced in a mathematical form. According to Zadeh (1975), four basic units are required for the effective application of any fuzzy modelling strategy: the fuzzification unit (process that converts traditional inputs into fuzzy inputs), fuzzy rules unit (IF-THEN logical system that links the input to output variables), fuzzy inference unit (process that explains and combines rule outputs), and defuzzification unit (process that converts the fuzzy output to a traditional output). The methodologies proposed by Mamdani (1974) and Takagi and Sugeno (1985) are the most widespread in the development of fuzzy rule systems. The fuzzy model presented by Mamdani (1974) can represent a general class of systems that may include static or dynamic nonlinear systems.

Fuzzy logic applications have been used in estimating the daily reference evapotranspiration (Odhiambo et al., 2001), predicting multiple soil properties (Lee et al., 2003), assessing

*Corresponding author, email: dewidar@ksu.edu.sa

Received 25 August 2017; accepted in revised form 7 June 2019

the water quality in rivers (Ocampo-Duque et al., 2006), determining irrigation efficiency (Kangrang and Chaleeraktragoon, 2007), developing rainfall-runoff models to describe the non-linear relationship between the rainfall and runoff in a real system (Jacquin and Shamseldin, 2006), and predicting the suspended sediment content of a river (Demirci and Baltaci, 2013). Abdel and Adeb (2014) showed that the design, management, operation, and hydraulic evaluation associated with on-farm water applications depends heavily on the infiltration attributes of the soil. This is because the infiltration behaviour of the soil directly determines essential parameters such as inflow rate, run time, application time, and percolation depth in irrigation systems. Therefore, prediction of the soil IR is an important issue in agricultural engineering, water resource management, and soil desalination. The aims of this study were to: (i) develop a predictive model for soil IR using FLS, (ii) assess the performance of the developed fuzzy model using a statistical comparison between the results of the soil infiltration obtained from the developed fuzzy model and experimental findings, and (iii) compare the developed fuzzy model with the MLR in terms of their appropriateness for predicting the soil IR. Moreover, a sensitivity analysis (SA) was also performed and discussed to evaluate the effects of the input parameters on the infiltration modelling process.

MATERIALS AND METHODS

Data and site description

Data used in building the fuzzy model were obtained from an experiment published by Hajiaghahi et al. (2014). The published experiment was conducted at the agricultural fields of Karaj, Alborz Province, Iran (35° 59' N, 51° 6' E and altitude of 1 300 m amsl). The climate in the centre of Iran is classified as semi-arid (345 mm rainfall annually). The infiltration rates were estimated with the help of a cylindrical infiltrometer for all the selected treatments (Fig. 1). As shown in Fig. 1, the cylindrical infiltrometer has two concentric rings. The initial reading of the water level was taken once the two rings were filled with water. The level of water in the inner ring of the infiltrometer was recorded at a regular interval of 2.5, 5, 10, 15, 20, 25 and 30 minutes until the rate of infiltration became constant. The IR was then calculated from the observed

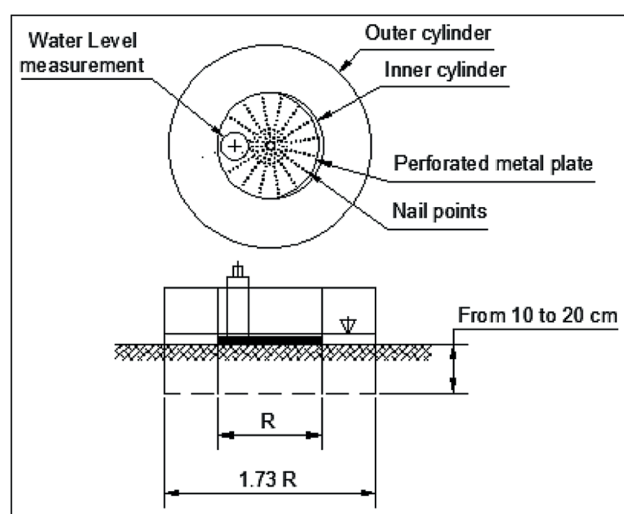


Figure 1. Schema of the double ring infiltrometer

cumulative infiltration data. The summary statistics of infiltration rates and soil texture properties which greatly affect the infiltration process (Fig. 2) are listed in Table 1.

Fuzzy logic system

The FLS is a theory in formal mathematics that enables a definitive solution to be obtained for problems that are complex, uncertain, and unstructured (Bojórquez-Tapia et al., 2002). A general fuzzy system (Zadeh, 1975) is composed of four primary elements: fuzzification, fuzzy rules, a fuzzy inference engine, and defuzzification (Fig. 3).

Fuzzification

Fuzzification is a process that transforms a numerical value into a fuzzy value. Fuzzification is usually created with fuzzy sets that are defined by a membership function. Fuzzy sets commonly assign a domain of interest to the interval [0, 1]. The fuzzy set $A \subset X$ is given by:

$$A = \left\{ \left(x_i, \mu_A(x) \right) \mid x \in X \right\}, \quad 0 \leq \mu_A(x) = 1 \quad (1)$$

where x_i are the elements of the universe of discourse (X), and $\mu_A(x)$ is the membership function of x in A .

Trapezoidal and triangular shapes are the most commonly used membership functions (Kosko and Toms, 1993).

Table 1. Summary statistics of infiltration rates and soil texture properties used in the developed fuzzy model

Statistics	Infiltration rate (mm/h)	Silt (%)	Clay (%)
Mean	10.3	34.62	31.57
Standard deviation	7.75	12.73	8.03
Kurtosis	0.17	-0.01	-0.3
Skewness	1.25	-1.12	-0.52
Minimum	1.7	5	11
Maximum	28.5	52	46

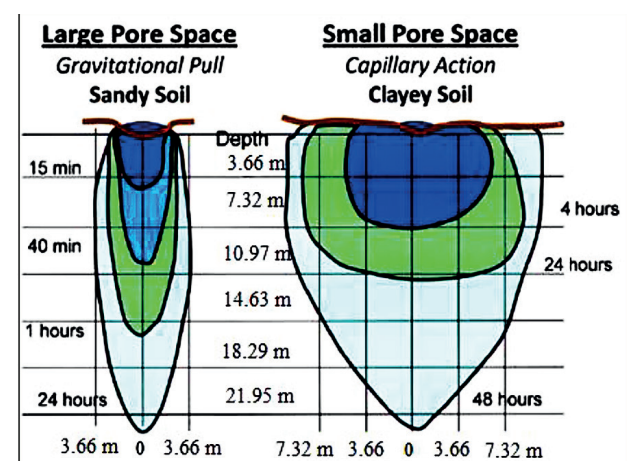


Figure 2. Comparative movement of water in sandy and clayey soils. In sandy soils, water moves downwards readily owing to the gravitational pull. In clay soils, water moves slowly in all directions by capillary action (Whiting et al., 2005).

Fuzzy rule base

The fuzzy rule base comprises fuzzy rules that represent all possible fuzzy relations between the input and output variables. These rules are expressed in IF-THEN statements (Eq. 2) (Ross, 2005). The number of rules is determined by the number of input parameters along with the membership functions.

$$\text{if } x_1 \text{ is } A_{i,1} \text{ and } x_2 \text{ is } A_{i,2} \text{ and } \dots, x_p \text{ is } A_{i,p} \text{ then } y \text{ is } B_{i,1} = 1, \dots, k \quad (2)$$

where y is the variable to be inferred, and x_1, x_2, \dots, x_p are the input variables. A_{ij} are the terms that linguistic variables can assume (fuzzy sets), and i is the index of the rule k . B_i is the term assumed by the output linguistic variable.

Fuzzy inference engine

The fuzzy inference engine is the strategic unit that takes into consideration all the possible fuzzy rules in the fuzzy rule base and learns how to transform a set of inputs into corresponding outputs. The fuzzy inference engine uses IF-THEN rules along with the connectors 'OR' or 'AND' for drawing essential decision rules (Zadeh, 1975). The methodologies introduced by Mamdani (1974) and Takagi and Sugeno (1985) are the most

commonly used for the fuzzy inference system. Figure 4 shows the Mamdani strategy represented in Eq. 2 for two simple rules.

Defuzzification

The defuzzification is the process by which the fuzzy results of the inference are transformed into a numerical output (Van-Leekwijck and Kerre, 1999). The methods frequently utilised in the defuzzification procedure are the mean of the maximums, smallest of the maximums, centre of gravity (COG) and the centroid of area. More detailed descriptions of fuzzy logic and fuzzy inference system models can be found in Ross (2005). An example of defuzzification using COG is presented in Fig. 5.

Building the fuzzy model

The fuzzy model was developed in the fuzzy logic toolbox of MATLAB software (Mathworks, 2016) using the Mamdani minimum–maximum inference engine. Flowchart of the FLS for modelling the soil IR is represented schematically in Fig. 6. The input and output variables that were considered in the fuzzy model were silt content (%), clay content (%), and IR (mm/h). The numerical inputs and output were transformed into fuzzy variables. The input variables were

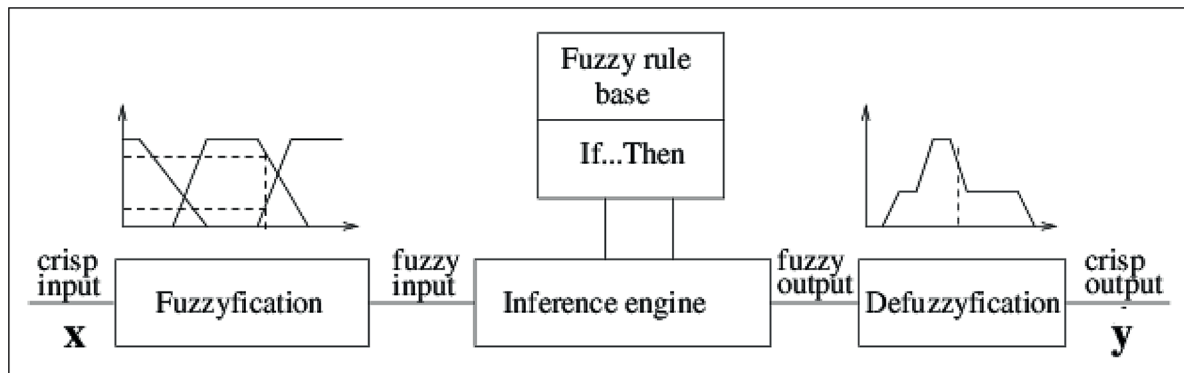


Figure 3. Basic structure of the fuzzy logic system

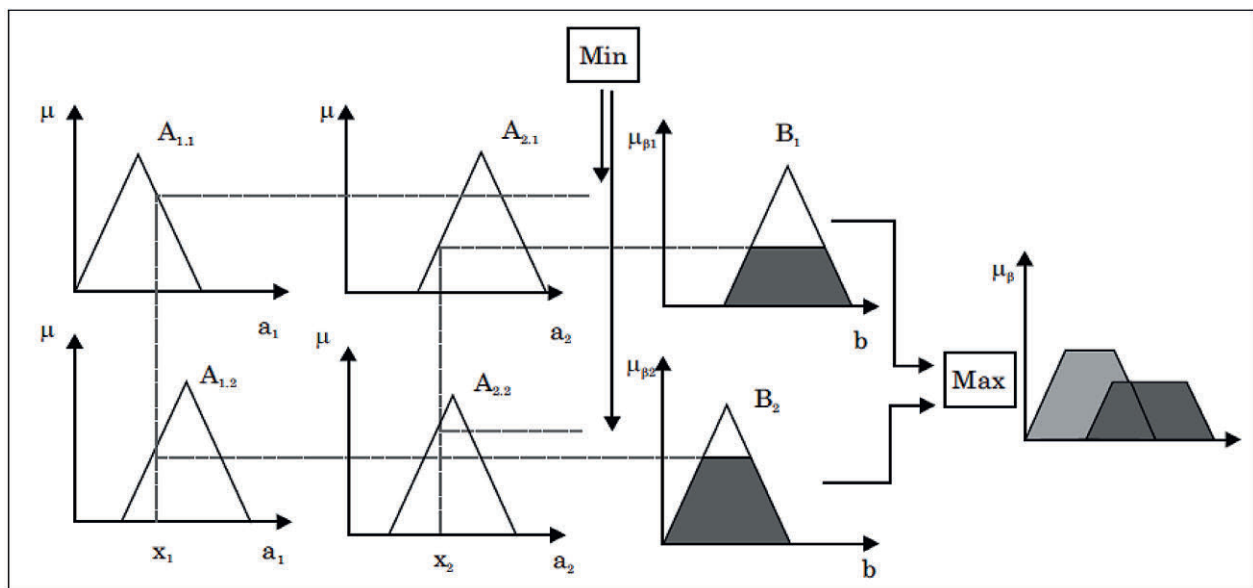


Figure 4. Mamdani strategy for fuzzy inference (Ross, 2005)

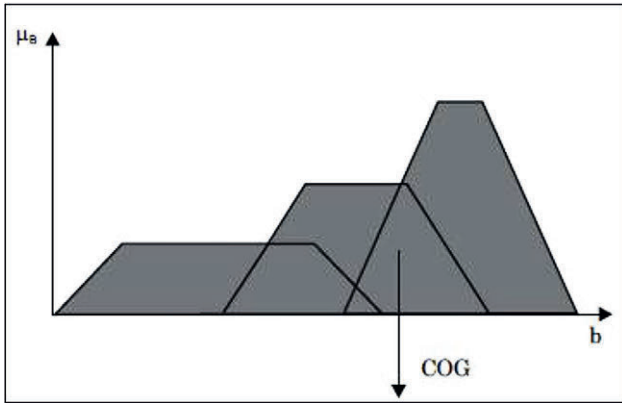


Figure 5. Centre of gravity (COG) method for defuzzification

then described using 5 linguistic terms: ‘very low’ (VL), ‘low’ (L), ‘medium’ (M), ‘high’ (H), and ‘very high’ (VH) which are used to describe all the possible fuzzy inputs. The output variable was also categorised into 5 classes: ‘very slow’ (VS), ‘slow’ (S), ‘medium’ (M), ‘rapid’ (R), and ‘very rapid’ (VR) to characterise all the possible fuzzy outputs. The fuzzy model was trained using the triangular membership functions for both the input and output parameters (Fig. 7). The choice of the number of membership functions and their initial values was based on knowledge of the system and experimental conditions (Eqs 3–5). The developed fuzzy model relied on 25 rules generated by the rule editor to describe all relationships between the input and output variables (Table 2). Finally, COG was used, which is the most popular defuzzification method for obtaining a real-value (numerical) output.

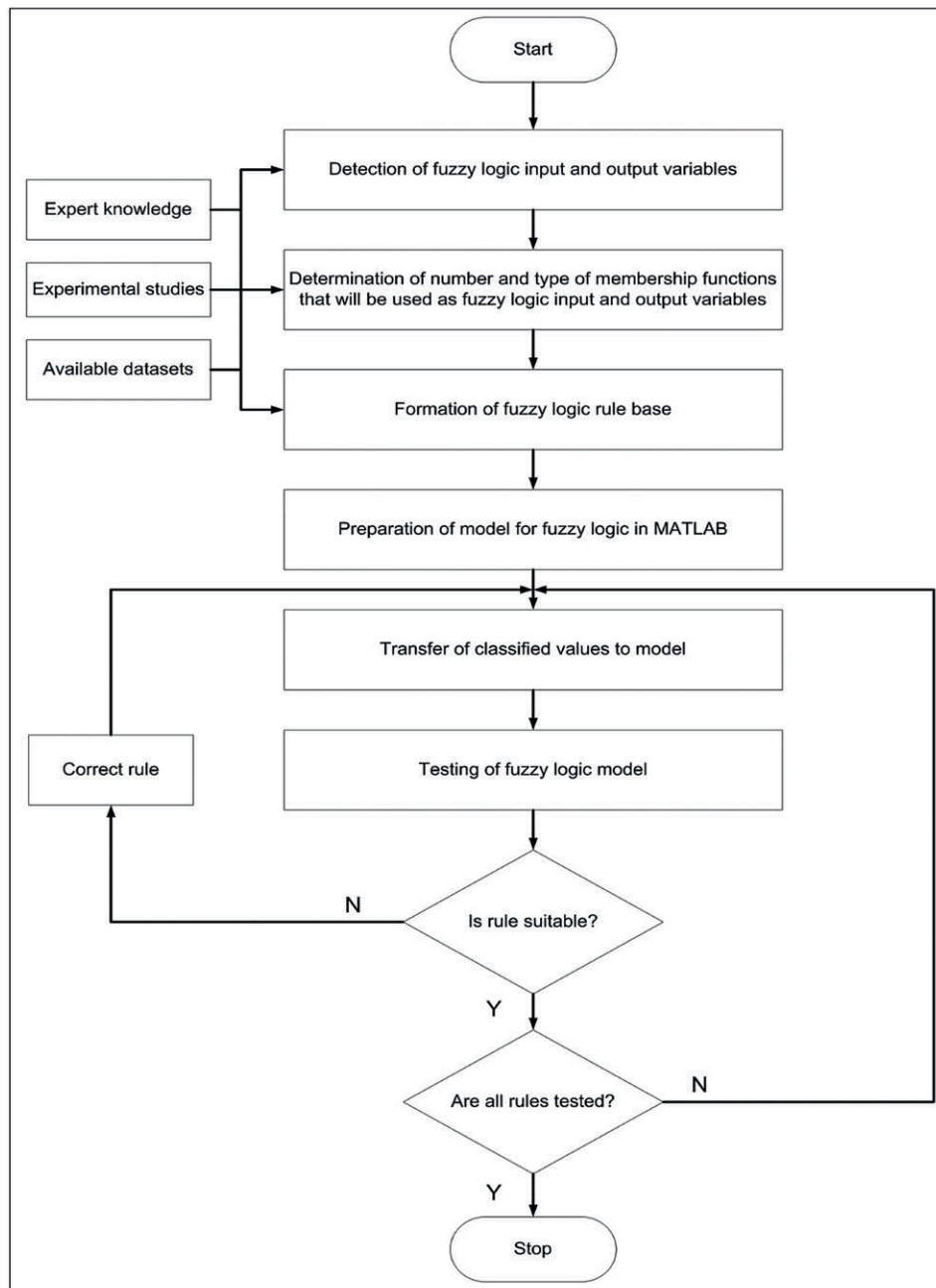


Figure 6. Flowchart of the fuzzy logic algorithm

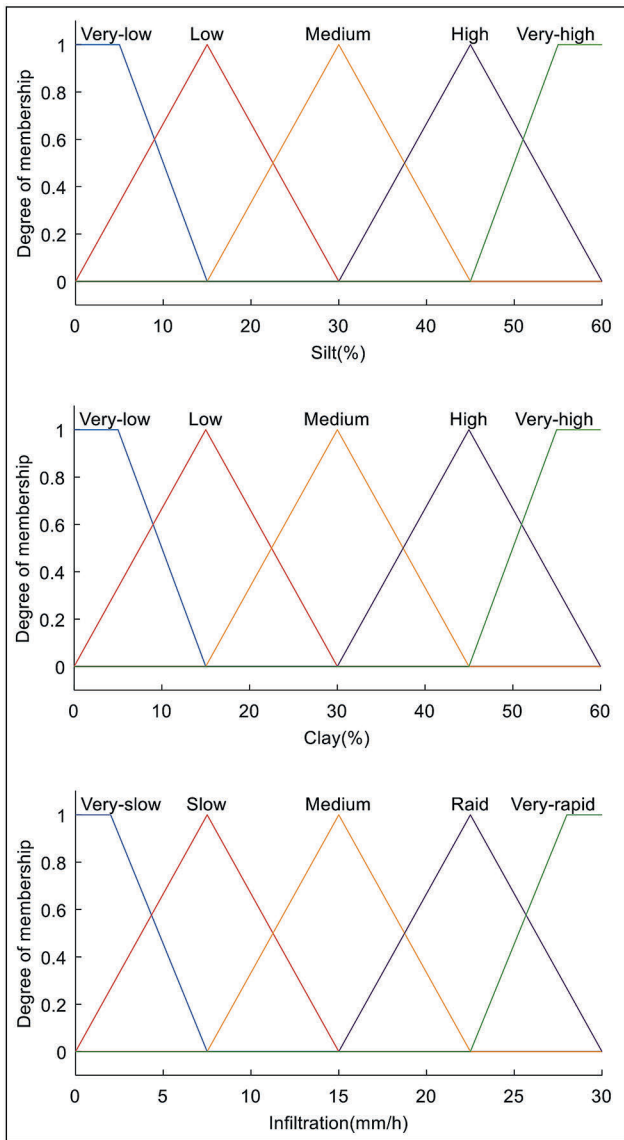


Figure 7. Membership functions defined for fuzzy input and output variables

$$\text{Silt content } (i_1) = \begin{cases} i_1 : 0 \leq i_1 \leq 60 \\ 0 : \text{otherwise} \end{cases} \quad (3)$$

$$\text{Clay content } (i_2) = \begin{cases} i_2 : 0 \leq i_2 \leq 60 \\ 0 : \text{otherwise} \end{cases} \quad (4)$$

$$\text{Infiltration rate } (j) = \begin{cases} o_j : 0 \leq o_j \leq 30 \\ 0 : \text{otherwise} \end{cases} \quad (5)$$

Multiple linear regression

MLR is a method that is used to model the linear relationship between a dependent variable (response) and one or more

Table 2. Rules generated for the infiltration model

Rule No.	IF	AND	THEN
1	(Silt is 'Very low')	(Clay is 'Very low')	(Infiltration is 'Very rapid')
2	(Silt is 'Very low')	(Clay is 'Low')	(Infiltration is 'Rapid')
3	(Silt is 'Very low')	(Clay is 'Medium')	(Infiltration is 'Medium')
4	(Silt is 'Very low')	(Clay is 'High')	(Infiltration is 'Slow')
5	(Silt is 'Very low')	(Clay is 'Very high')	(Infiltration is 'Very slow')
6	(Silt is 'Low')	(Clay is 'Very low')	(Infiltration is 'Rapid')
7	(Silt is 'Low')	(Clay is 'Low')	(Infiltration is 'Rapid')
8	(Silt is 'Low')	(Clay is 'Medium')	(Infiltration is 'Medium')
9	(Silt is 'Low')	(Clay is 'High')	(Infiltration is 'Slow')
10	(Silt is 'Low')	(Clay is 'Very high')	(Infiltration is 'Very slow')
11	(Silt is 'Medium')	(Clay is 'Very low')	(Infiltration is 'Rapid')
12	(Silt is 'Medium')	(Clay is 'Low')	(Infiltration is 'Medium')
13	(Silt is 'Medium')	(Clay is 'Medium')	(Infiltration is 'Medium')
14	(Silt is 'Medium')	(Clay is 'High')	(Infiltration is 'Slow')
15	(Silt is 'Medium')	(Clay is 'Very high')	(Infiltration is 'Very slow')
16	(Silt is 'High')	(Clay is 'Very low')	(Infiltration is 'Medium')
17	(Silt is 'High')	(Clay is 'Low')	(Infiltration is 'Slow')
18	(Silt is 'High')	(Clay is 'Medium')	(Infiltration is 'Slow')
19	(Silt is 'High')	(Clay is 'High')	(Infiltration is 'Very slow')
20	(Silt is 'High')	(Clay is 'Very high')	(Infiltration is 'Very slow')
21	(Silt is 'Very high')	(Clay is 'Very low')	(Infiltration is 'Medium')
22	(Silt is 'Very high')	(Clay is 'Low')	(Infiltration is 'Slow')
23	(Silt is 'Very high')	(Clay is 'Medium')	(Infiltration is 'Very slow')
24	(Silt is 'Very high')	(Clay is 'High')	(Infiltration is 'Very slow')
25	(Silt is 'Very high')	(Clay is 'Very high')	(Infiltration is 'Very slow')

independent variables (predictors). The corresponding general equation is as follows:

$$y = b_1 x_1 + b_2 x_2 + \dots + b_n x_n + c \quad (6)$$

where y is the dependent variable; b_1, b_2, \dots, b_n are the coefficients of regression; and x_1, x_2, \dots, x_n are the explanatory variables (predictors).

In this study, the MLR was built using the data used in the fuzzy model to create a mathematical relationship for predicting the IR as a function of the silt and clay contents (Eq. 7). The

standard error (SE), coefficient of correlation (CC), t statistic and probability value (*p*-value) of independent parameters were employed to evaluate the accuracy of the predictive MLR:

$$IR = \beta_0 + \beta_1 SI + \beta_2 CL \quad (7)$$

where IR is infiltration rate (mm/h); SI is silt content (%); and CL is clay content (%).

Criteria for evaluation

The performance of both the fuzzy and MLR models was evaluated by the coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), overall index (OI), and model efficiency (ME). The R^2 , MAE, RMSE, OI, and ME were calculated using Eqs 8–12, as presented in Table 3 (Rahman and Bala, 2010; Alazba et al., 2011; Zangeneh et al., 2012). The higher values of R^2 represent a greater similarity between the observed and predicted values. On the other hand, the lower values of MAE and RMSE show a high accuracy between the observed and predicted values. An ME and OI value of 1 implies that the observed and predicted results agree well (Alazba et al., 2011). In other words, the MAE and RMSE values have to be closer to zero, whereas the values of R^2 , OI, and ME should approach 1 as much as possible.

P_i and \bar{P} represent the predicted and average predicted values, respectively; Q_i and \bar{Q} represent the observed and average of the observed values, respectively; Q_{\max} is the maximum observed value; Q_{\min} is the minimum observed value; and N is the number of data points.

Sensitivity analysis

A sensitivity analysis (SA) is a useful tool for determining the contribution and relative importance of parameters in the modelling process. The SA assesses and describes how the model output values are affected by changes in the input values. The cosine amplitude method (CAM) was used in this study to distinguish the most sensitive factors affecting the soil IR. This is an effective method of performing an SA (Ross, 2005). The

degree of sensitivity of each input factor (silt and clay contents) was assigned by establishing the strength of the relationship (R_{ij}) between the IR and the input factors under consideration. A higher CAM value indicates a greater impact on the IR.

Let us assume that n data samples are gathered from a common data array X ; then the datasets employed to construct a data array X are defined as follows:

$$X = \{x_1, x_2, x_3, \dots, x_m\} \quad (13)$$

Each of the elements x_i in the data array X is a vector of length m , that is

$$x_i = \{x_{i1}, x_{i2}, x_{i3}, \dots, x_{im}\} \quad (14)$$

Thus, each of the data pairs can be thought of as a point in an m -dimensional space, wherein each point requires m coordinates for a complete description. The strength of the relationship between the data pairs x_i and x_j is estimated and demonstrated using the following equation:

$$R_{ij} = \frac{\sum_{k=1}^m x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^m x_{ik}^2 \sum_{k=1}^m x_{jk}^2}}, \quad 0 \leq R_{ij} \leq 1 \quad (15)$$

where $i, j = 1, 2, \dots, n$

RESULTS AND DISCUSSION

Fuzzy logic model

Figure 8 shows the graphical depiction of the 25 rules produced to map the inputs (silt and clay contents) into the output (IR). In this figure, we observe that each rule is represented in a row individually, whereas the variables are represented in individual columns. The first two columns of the plot describe the membership functions for the input factors, which is indicated through the antecedent (IF-part) of each rule. The third column shows the membership functions represented via the consequent (THEN-part) of each rule. The vertical red

Table 3. Statistical performance parameters

Parameter	Description	Equation
Coefficient of Determination	Measures the degree of correlation between the experimental and predicted values	$R^2 = \frac{\left[\sum_{i=1}^N (P_i - \bar{P})(Q_i - \bar{Q}) \right]^2}{\sum_{i=1}^N (P_i - \bar{P})^2 \sum_{i=1}^N (Q_i - \bar{Q})^2} \quad (8)$
Root mean square error	Measures the error in the same units as the variable between the experimental and predicted values	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - Q_i)^2} \quad (9)$
Mean absolute error	Measures the average magnitude of the errors in a set of forecasts without taking into consideration their direction	$MAE = \frac{1}{N} \sum_{i=1}^N P_i - Q_i \quad (10)$
Model efficiency	Measures the efficiency between the experimental and predicted values	$ME = 1 - \frac{\sum_{i=1}^N (Q_i - P_i)^2}{\sum_{i=1}^N (Q_i - \bar{Q})^2} \quad (11)$
Overall index	Measures the performance of the mathematical models and fitting between the experimental and predicted values	$OI = \frac{1}{2} \left(1 - \left(\frac{RMSE}{Q_{\max} - Q_{\min}} \right) + ME \right) \quad (12)$

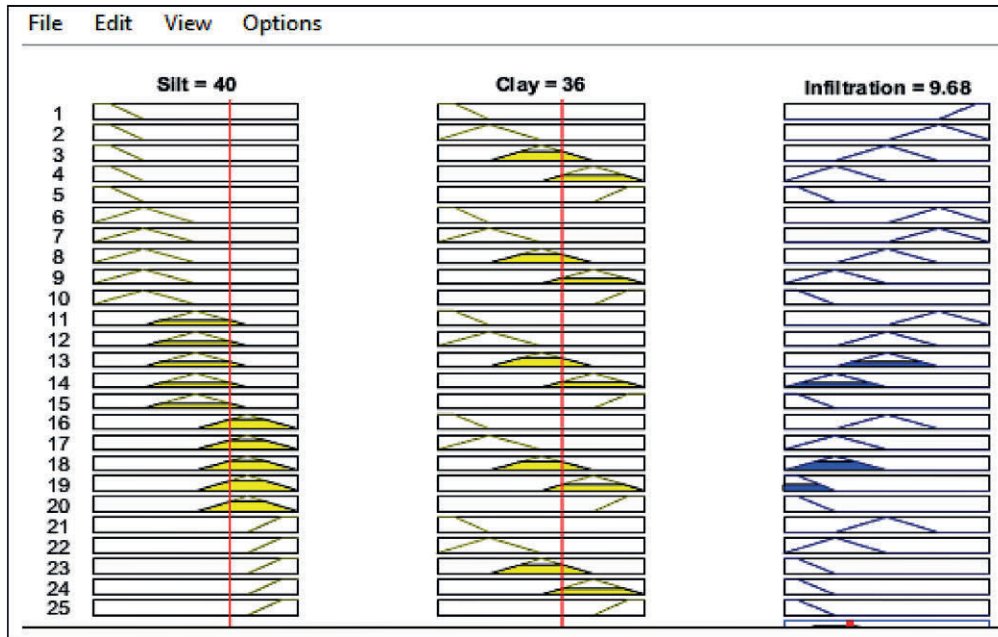


Figure 8. Rule viewer window of the developed fuzzy model

lines inside the first and second columns show the current data of the inputs. A yellow patch under the actual curve of the membership function is used to represent the value of the fuzzy membership. Thus, the properties of each variable in relation to the input index line are presented. The lowest plot inside the right column is the combination of each consequent. The value of the defuzzified output is indicated by a bold vertical line crossing the aggregate fuzzy set. For example, for the system inputs of 40% silt and 36% clay, the defuzzified IR is found to be 9.68 mm/h. Furthermore, the surface viewer model provides a three-dimensional curve that maps the silt and clay contents to the IR (Fig. 9). In Fig. 9, the IR values decrease with the increasing clay content of the soil. Similarly, the same trend was observed in the case of the silt content. However, the relative change was smaller.

Simulating the fuzzy model

The fuzzy model was simulated in the Simulink environment of the MATLAB software which is a graphical programming environment for modelling, simulating, and analysing multi-domain dynamic systems (Fig. 10). As shown in Fig. 10, two inputs of silt content (40%) and clay content (36%) are multiplexed and sent into the fuzzy logic controller, while the output of the IR (9.60 mm/h) is captured on a display box. It is apparent that there is a close agreement between the observed (9.60 mm/h) and predicted (9.68 mm/h) values of the IR. This high level of agreement between the observed and predicted values demonstrates the effectiveness of the developed fuzzy model for predicting the soil IR.

MLR model

The effect of the independent variables (SI and CL) and their interactions on the dependent variable (IR) were statistically analysed. The SE, *t*-statistic and *p*-value are presented in Table 4. On examining the *p*-value, a significant effect was observed between the independent variables and the IR at

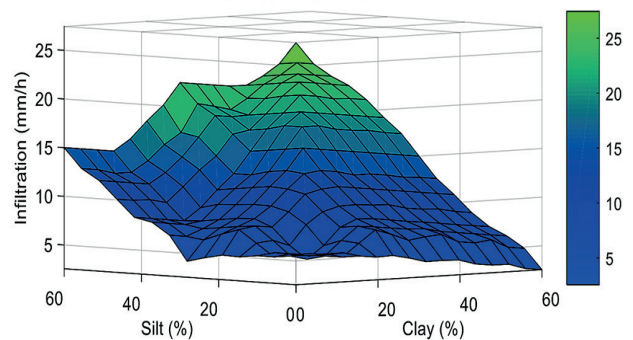


Figure 9. Surface graph showing the relationship of the IR with the silt and clay contents

an alpha level of 0.05. Both the SI and CL were found to be influential variables within the calculation of the IR, wherein the SE of those variables were ± 0.02 and ± 0.04 , respectively. Moreover, Fig. 11 shows the observed and predicted values of the IR using the MLR. From this figure, it can be observed that the values predicted by the MLR are in conformity with those obtained from the field. Also the SI and CL variables are inversely proportional to the IR. These results agree with the results of Hajiaghahi et al. (2014).

Table 4. Results of regression analysis of IR for the developed MLR

Variables	Coefficients	Value	SE	<i>t</i> -statistic	<i>P</i> -value	CC
Intercept	β_0	37.78	0.94	40.01	0.00	0.96
SI	β_1	-0.32	0.02	-13.64	0.00	
CL	β_2	-0.52	0.04	-14.28	0.00	

CC: coefficient of correlation; SE: standard error of regression coefficients; *p*-value: probability value; SI: silt content; CL: clay content

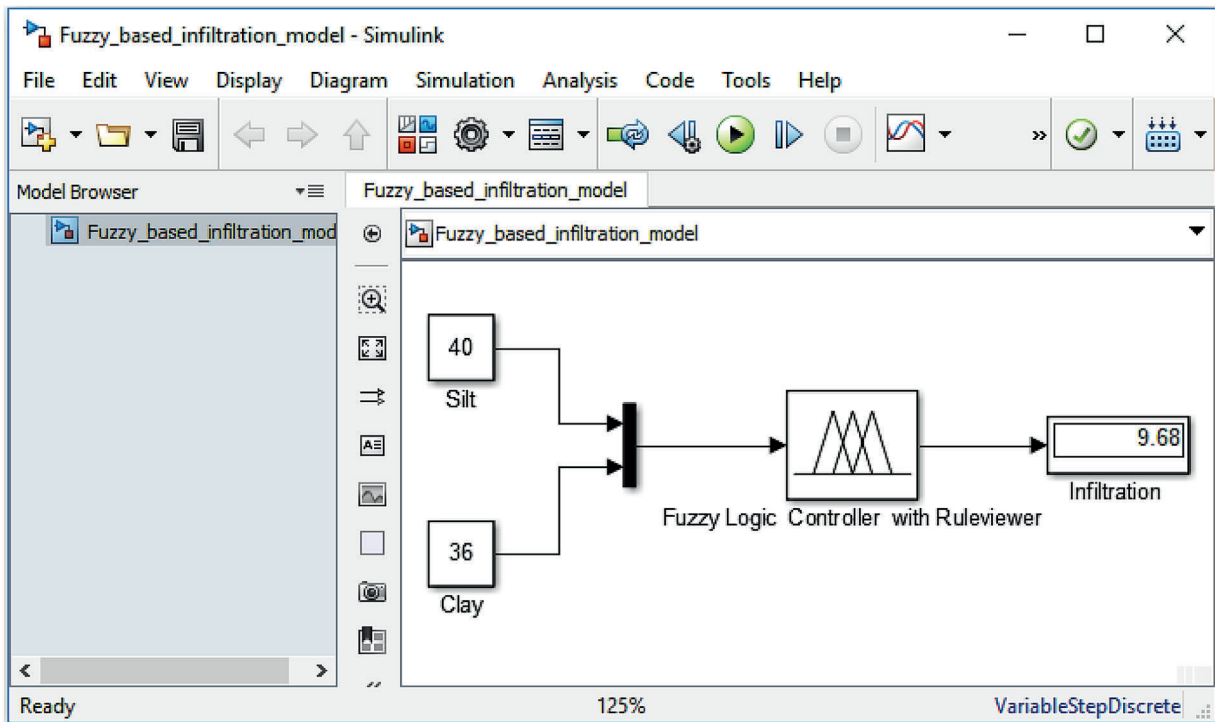


Figure 10. Simulation of soil IR prediction using fuzzy logic controller developed in Simulink environment

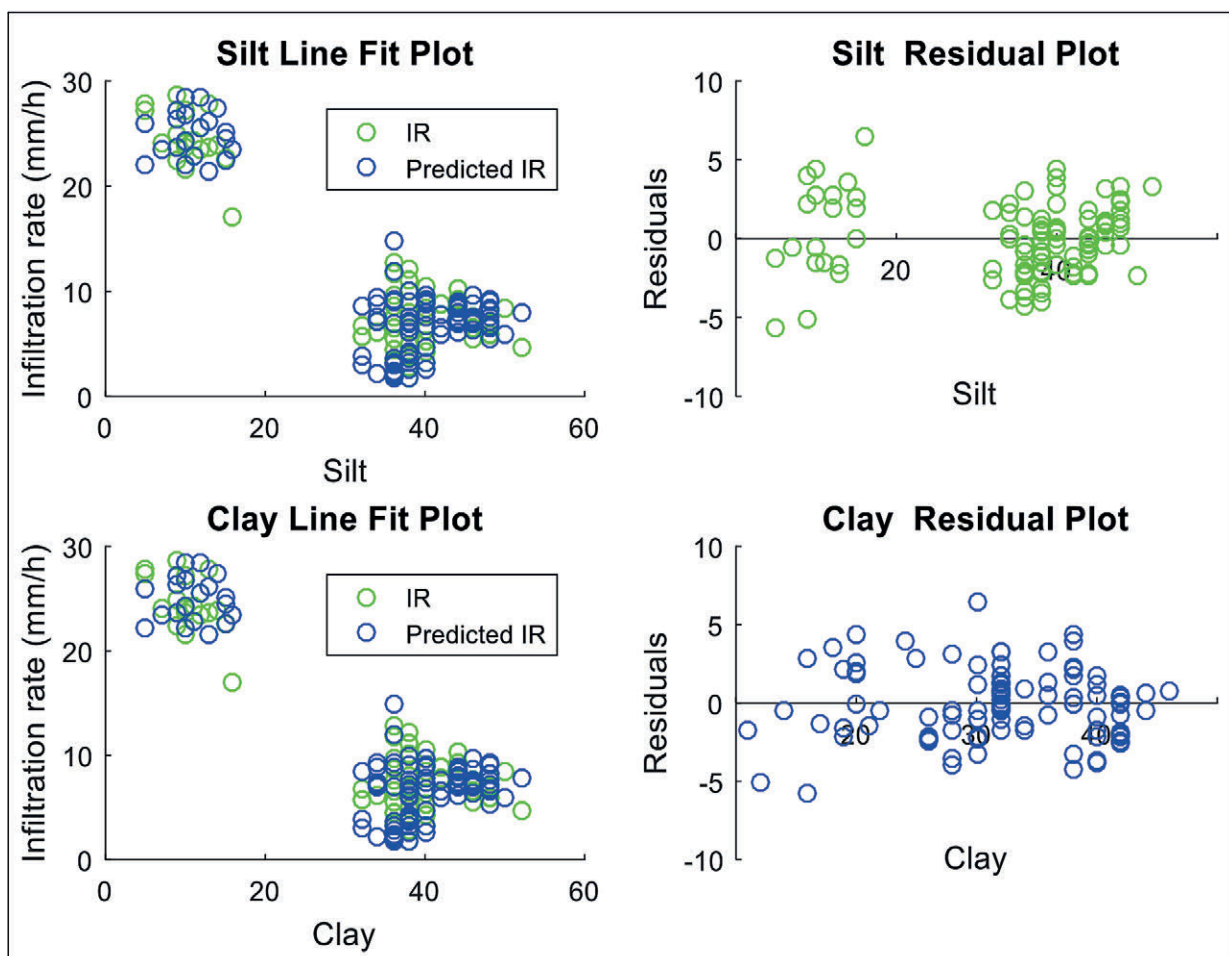


Figure 11. Comparison between observed and predicted values of IR using MLR

Comparing the fuzzy model with MLR

Figure 12 compares the experimental and predicted values of the soil IR using the fuzzy and MLR models. The dotted line represents the condition in which the outputs and targets are equal, whereas the circles represent data points. The solid line represents the best fit between the outputs and targets. It is apparent that the data points (circles) are clustered along the dotted line, which emphasizes that the IR values correlated by the developed fuzzy model and MLR provide a good match between the experimental and predicted results. However, results generated by the fuzzy model are more accurate than those calculated using the MLR. The higher level of agreement is also reflected in the values of the statistical parameters, such as correlation coefficient (R), which was 0.97654 and 0.95554 for the fuzzy model and MLR, respectively. Generally, a close agreement between the observed and predicted results demonstrates the consistency of the developed fuzzy model for predicting the soil IR.

Validating the developed fuzzy model

Table 5 shows the results of R^2 , MAE, RMSE, OI, and ME, which are used to evaluate the agreement between the measured and predicted results. The developed fuzzy model had R^2 , OI, and ME values that were approximately 4.3%, 94.5%, and 95.3%, respectively, more accurate than those from

MLR, which demonstrates a close agreement between the observed and predicted results by the FLS. Also, it should be noted (Table 5) the MLR had RMSE and MAE values that were approximately 1.5 times as high as those obtained using the developed fuzzy model. Furthermore, a graph of the relative errors for the fuzzy model and MLR using the soil IR dataset is shown in Fig. 13. The fuzzy model was found to have the ability to predict the soil IR with errors (93% of the values) fluctuating mostly from -0.5 to $+0.5$. In contrast, the values of relative error obtained from the MLR which fall within $\pm 0.5\%$ were less than 85% of the entire error values. In conclusion, the comparison between the developed fuzzy model and MLR demonstrated that the fuzzy model can predict the soil IR more effectively than MLR.

Table 5. Statistical parameters for evaluating the performance of the fuzzy model and MLR

Models	Statistical parameters				
	R^2	RMSE	MAE	ME	OI
Fuzzy	0.953	1.53	1.28	0.953	0.945
MLR	0.913	2.37	1.92	0.913	0.914

R^2 : coefficient of determination; RMSE: root mean square error; MAE: mean absolute error; ME: model efficiency; OI: overall index of model performance.

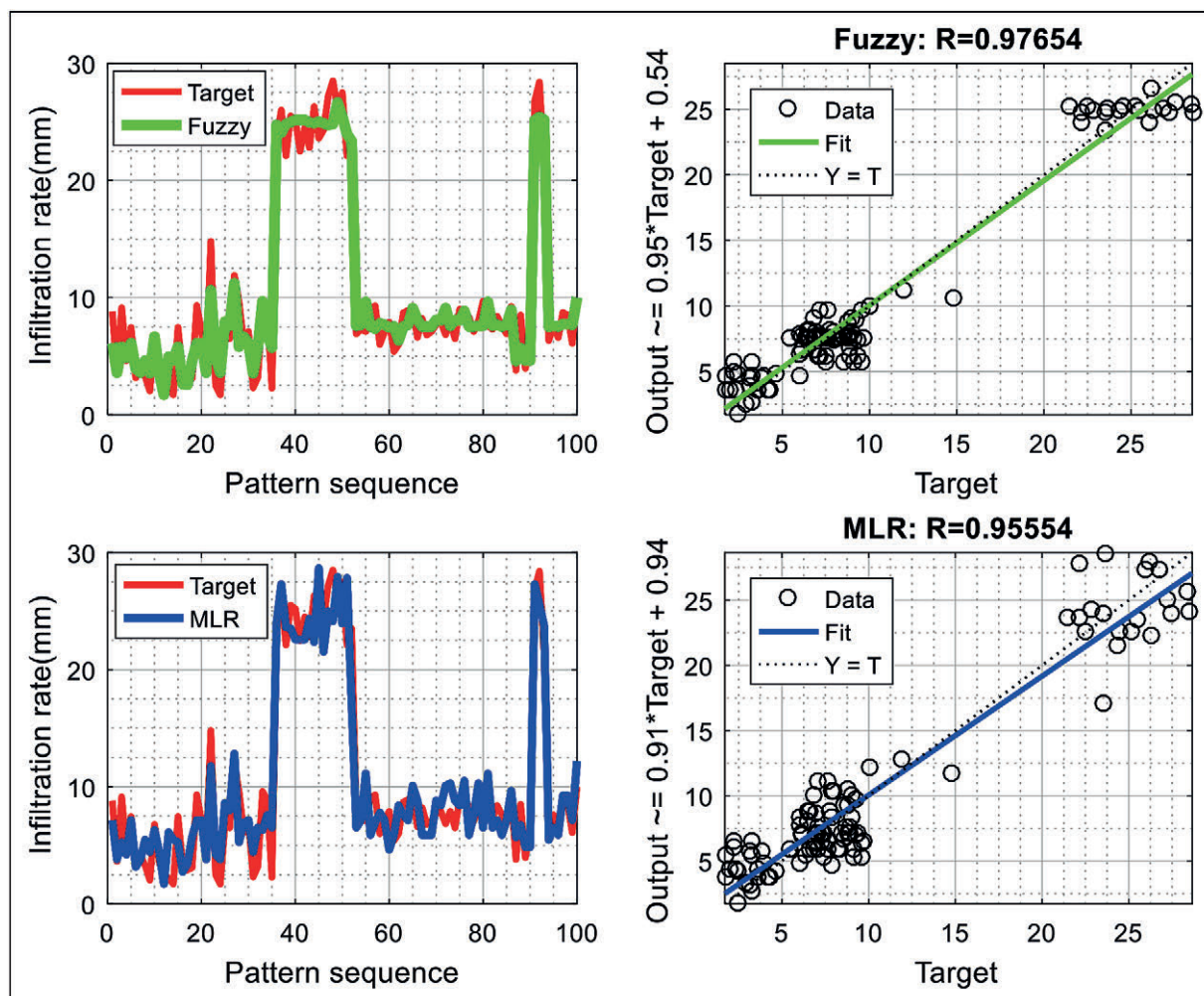


Figure 12. Comparing the experimental and predicted values of IR using the developed fuzzy model and MLR

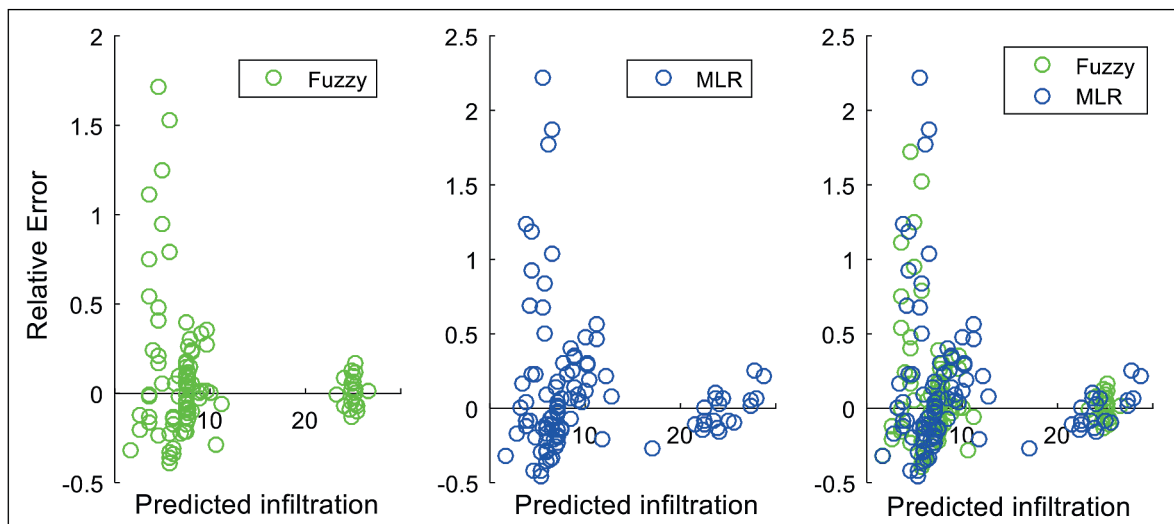


Figure 13. Relative error of the fuzzy model and MLR using the soil IR dataset

Sensitivity analysis

Using Eq. 14, a series of SAs was conducted on the input and output parameters. The values of R_{ij} between the IR values predicted by the fuzzy and MLR models and related input factors using the CAM method are shown in Fig. 14. As shown, the obtained R_{ij} values for all the inputs were not significantly different but close to 1, which indicates that all inputs strongly contributed to the developed fuzzy and MLR models. Furthermore, this indicates that all the inputs were important in predicting the soil IR and none should be neglected. The obtained results agree with the findings of Hajiaghaei et al. (2014).

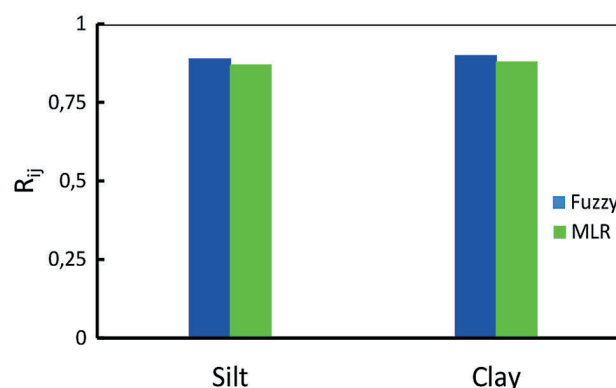


Figure 14. Sensitivity analysis of the soil IR and each input parameter for the fuzzy and MLR models

CONCLUSIONS

Knowledge of infiltration processes is imperative in irrigation management and other agricultural studies. In this study, an intelligent model based on the fuzzy logic system was developed as an alternative algorithm to estimate the infiltration rate of the soil. For this study, a dataset containing 100 observations in the field, obtained by double ring infiltrometer, was analysed. Based on the obtained results, the fuzzy model has a suitable capability to predict the infiltration rate of the soil. The fuzzy model also provides better performance than the MLR model that was used for evaluation purposes. Sensitivity analysis suggests that the clay content is the most effective parameter for the estimation of soil infiltration rate.

ACKNOWLEDGEMENTS

This project was supported by the Vice Deanship of Research Chairs at King Saud University.

REFERENCES

ABDEL DAM and ADEEB WAM (2014) Comparison between hydro-flume and open field head ditch irrigation systems at Kenana Sugar Scheme, Sudan. *Agric. Sci.* 5 (07) 588–603. <https://dx.doi.org/10.4236/as.2014.57062>

ALAZBA AA, MATTAR MA, ELNESR MN and AMIN MT (2011) Field assessment of friction head loss and friction correction factor equations. *J. Irrig. Drain. Eng.* 138 (2) 166–176. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000387](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000387)

AZAMATHULLA HM, HAGHIABI AH and PARSAIE A (2016) Prediction of side weir discharge coefficient by support vector machine technique. *Water Sci. Technol. Water Suppl.* 16 (4) 1002–1016. <https://doi.org/10.2166/ws.2016.014>

BOJÓRQUEZ-TAPIA LA, JUÁREZ L and CRUZ-BELLO G (2002) Integrating fuzzy logic, optimization, and GIS for ecological impact assessments. *Environ. Manage.* 30 (3) 418–433. <https://doi.org/10.1007/s00267-002-2655-1>

BOUMA J (1982) Measuring the hydraulic conductivity of soil horizons with continuous macropores 1. *Soil Sci. Soc. Am. J.* 46 (2) 438–441. <https://doi.org/10.2136/sssaj1982.03615995004600020047x>

DEMIRCI M and BALTACI A (2013) Prediction of suspended sediment in river using fuzzy logic and multilinear regression approaches. *Neural Comput. Appl.* 23 (1) 145–151. <https://doi.org/10.1007/s00521-012-1280-z>

DOERR SH, SHAKESBY RA and WALSH R (2000) Soil water repellency: Its causes, characteristics and hydro-geomorphological significance. *Earth-Sci. Rev.* 51 (1–4) 33–65. [https://doi.org/10.1016/S0012-8252\(00\)00011-8](https://doi.org/10.1016/S0012-8252(00)00011-8)

HAGHNAZARI F, SHAHGHOLI H and FEIZI M (2015) Factors affecting the infiltration of agricultural soils. *Int. J. Agron. Agric. Res.* 6 (5) 21–35.

HAJIAGHAEI A, RASHIDI M, SADEGHI MA, GHOLAMI M and JABERINASA B (2014) Prediction of soil infiltration rate based on silt and clay content of soil. *Am.-Euras. J. Agric. Environ. Sci.* 14 (8) 702–706.

HELMING K, RÖMKENS MJM and PRASAD S (1998) Surface roughness related processes of runoff and soil loss: A flume

- study. *Soil Sci. Soc. Am. J.* **62** (1) 243–250. <https://doi.org/10.2136/sssaj1998.03615995006200010031x>
- JACQUIN AP and SHAMSELDIN AY (2006) Development of rainfall–runoff models using Takagi–Sugeno fuzzy inference systems. *J. Hydrol.* **329** (1–2) 154–173. <https://doi.org/10.1016/j.jhydrol.2006.02.009>
- KANGRANG A and CHALEERAKTRAKOON C (2007) A fuzzy-GAS model for determining varied irrigation efficiency. *Am. J. Appl. Sci.* **4** (6) 339–345. <https://doi.org/10.3844/ajassp.2007.339.345>
- KISI O, KESHAVARZI A, SHIRI J, ZOUNEMAT-KERMANI M and OMRAN ESE (2017) Groundwater quality modeling using neuro-particle swarm optimization and neuro-differential evolution techniques. *Hydrol. Res.* **48** (6) 1508–1519. <https://doi.org/10.2166/nh.2017.206>
- KOSKO B and TOMS M (1993) *Fuzzy Thinking: The New Science of Fuzzy Logic*. Hyperion, New York. 350 pp.
- KOSTIAKOV AN (1932) On the dynamics of the coefficient of water percolation in soils and the necessity of studying it from the dynamic point of view for the purposes of amelioration. *Sixth Commission of the International Society of Soil Science* **1** 7–21.
- KUMAR M, RANJAN S, TIWARI NK and GUPTA R (2018) Plunging hollow jet aerators-oxygen transfer and modelling. *ISH J. Hydraul. Eng.* **24** (1) 61–67. <https://doi.org/10.1080/09715010.2017.1348264>
- LADO M, PAZ A and BEN-HUR M (2004) Organic matter and aggregate size interactions in infiltration, seal formation, and soil loss. *Soil Sci. Soc. Am. J.* **68** (3) 935–942. <https://doi.org/10.2136/sssaj2004.9350>
- LEE K, ZHANG N and DAS S (2003) Comparing adaptive neuro-fuzzy inference system (ANFIS) to partial least-squares (PLS) method for simultaneous prediction of multiple soil properties. Paper number 033144, 2003 ASAE Annual Meeting American Society of Agricultural and Biological Engineers. <https://doi.org/10.13031/2013.15017>
- MAMDANI EH (1974) Application of fuzzy algorithms for control of simple dynamic plant. In: *Proc. Inst. Elec. Eng.* **121** (12) 1585–1588. <http://dx.doi.org/10.1049/piee.1974.0328>
- MATHWORKS (2016) MATLAB and Fuzzy Logic Toolbox. MathWorks, Inc., Natick, Massachusetts, United States.
- METTERNICHT G (2001) Assessing temporal and spatial changes of salinity using fuzzy logic, remote sensing and GIS. Foundations of an expert system. *Ecol. Model.* **144** (2) 163–179. [https://doi.org/10.1016/S0304-3800\(01\)00371-4](https://doi.org/10.1016/S0304-3800(01)00371-4)
- MUDIARE OJ and ADEWUMI JK (2000) Estimation of infiltration from field-measured sorptivity values. *Nigerian J. Soil Res.* **1** 1–3.
- OCAMPO-DUQUE W, FERRE-HUGUET N, DOMINGO JL and SCHUHMACHER M (2006) Assessing water quality in rivers with fuzzy inference systems: A case study. *Environ. Int.* **32** (6) 733–742. <https://doi.org/10.1016/j.envint.2006.03.009>
- ODHIAMBO LO, YODER RE and YODER DC (2001) Estimation of reference crop evapotranspiration using fuzzy state models. *Trans. ASAE* **44** (3) 543–550.
- PATLE GT, SIKAR TT, RAWAT KS and SINGH SK (2018) Estimation of infiltration rate from soil properties using regression model for cultivated land. *Geol. Ecol. Landscapes* **3** (1) 1–13. <https://doi.org/10.1080/24749508.2018.1481633>
- PHILIP JR (1957) The theory of infiltration: 1. The infiltration equation and its solution. *Soil Sci.* **83** (5) 345–358.
- RAHMAN MM and BALA BK (2010) Modelling of jute production using artificial neural networks. *Biosyst. Eng.* **105** (3) 350–356. <https://doi.org/10.1016/j.biosystemseng.2009.12.005>
- RAMOS MC, NACCI S and PLA I (2003) Effect of raindrop impact and its relationship with aggregate stability to different disaggregation forces. *Catena* **53** (4) 365–376. [https://doi.org/10.1016/S0341-8162\(03\)00086-9](https://doi.org/10.1016/S0341-8162(03)00086-9)
- ROSS TJ (2005) *Fuzzy Logic with Engineering Applications* (2nd edn). Wiley, Hoboken.
- SIHAG P, TIWARI NK and RANJAN S (2017) Estimation and inter-comparison of infiltration models. *Water Sci.* **31** (1) 34–43. <https://doi.org/10.1016/j.wsj.2017.03.001>
- SINGH B, SIHAG P and DESWAL S (2019) Modelling of the impact of water quality on the infiltration rate of the soil. *Appl. Water Sci.* **9** (1) 15–23. <https://doi.org/10.1007/s13201-019-0892-1>
- SINGH B, SIHAG P and SINGH K (2018) Comparison of infiltration models in NIT Kurukshetra campus. *Appl. Water Sci.* **8** (2) 63–70. <https://doi.org/10.1007/s13201-018-0708-8>
- SMETTEM KRJ (1987) Characterization of water entry into a soil with a contrasting textural class: spatial variability of infiltration parameters and influence of macroporosity. *Soil Sci.* **144** (3) 167–174.
- SURYOPUTRO N, SOETOPPO W, SUHARTANTO ES and LIMANTARA LM (2018) Evaluation of infiltration models for mineral soils with different land uses in the tropics. *J. Water Land Dev.* **37** (1) 153–160. <https://doi.org/10.2478/jwld-2018-0034>
- TAKAGI T and SUGENO M (1985) Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst. Man Cybernetics* **SMC-15** (1) 116–132. <http://dx.doi.org/10.1109/TSMC.1985.6313399>
- TRAN LT, RIDGLEY MA, DUCKSTEIN L and SUTHERLAND R (2002) Application of fuzzy logic-based modeling to improve the performance of the revised universal soil loss equation. *Catena* **47** (3) 203–226. [https://doi.org/10.1016/S0341-8162\(01\)00183-7](https://doi.org/10.1016/S0341-8162(01)00183-7)
- VAND AS, SIHAG P, SINGH B and ZAND M (2018) Comparative evaluation of infiltration models. *KSCCE J. Civ. Eng.* **22** (10) 4173–4184. <https://doi.org/10.1007/s12205-018-1347-1>
- VAN-LEEKWIJCK W and KERRE EE (1999) Defuzzification: criteria and classification. *Fuzzy Sets Syst.* **108** (2) 159–178. [https://doi.org/10.1016/S0165-0114\(97\)00337-0](https://doi.org/10.1016/S0165-0114(97)00337-0)
- WHITING D, WILSON C and CARD A (2005) Managing soil tilth: Texture, structure and pore space. Gardening Series. Colorado Master Gardener; no. 7.723. <https://hdl.handle.net/10217/183047>
- ZADEH LA (1965) Fuzzy sets. *Inf. Control* **8** (3) 338–353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
- ZADEH LA (1975) The concept of a linguistic variable and its application to approximate reasoning—I. *Inf. Sci.* **8** (3) 199–249. [https://doi.org/10.1016/0020-0255\(75\)90036-5](https://doi.org/10.1016/0020-0255(75)90036-5)
- ZANGENEH M, OMID M and AKRAM A (2012) A comparative study between parametric and artificial neural networks approaches for economical assessment of potato production in Iran. *Span. J. Agric. Res.* **9** (3) 661–671. <https://doi.org/10.5424/sjar/20110903-371-10>

Experimental investigation of time-dependent local scour downstream of a stepped channel

Aysegul Ozgenc Aksoy¹ and Mustafa Dogan^{1*}

¹Civil Engineering Dept., Dokuz Eylul University, Izmir, Turkey

ABSTRACT

In this study, temporal variation of local scour occurring at the downstream part of the stepped channel were investigated experimentally. The experimental tests were carried out in a stepped flume with a height of 2.4 m. The width of the rectangular flume was 0.10 m and the length of the stilling basin was 2.12 m. Bed material was placed in a sediment box with a height of 24 cm and length of 2.48 m, without any compaction. Experiments were carried out by using bed material of 4 different grain size distributions, 2 different sill heights and 6 different flow rates. Two empirical equations which include Shields parameter (θ) and densimetric Froude particle number (F_p) were proposed by using the experimental findings to predict the temporal variation of the scour depth. The R^2 (coefficient of determination) values were computed for both proposed equations as 0.866 and 0.865. The scatter index (SI) values were also determined and computed as 8.73% and 8.25%. The fit of the equations was also determined by means of Fisher's test.

Keywords: stepped channel, local scour, temporal variation

INTRODUCTION

Stepped channels are currently preferred because they provide substantial energy dissipation and accordingly reduce the size of the downstream stilling basin. Stepped channels also result in reduced construction time and so have a large number of applications in engineering. Flow regimes above stepped channels can be described as nappe flow, transition flow and skimming flow. In the nappe flow regime, the water proceeds in a series of plunges from one step to another. In a skimming flow regime, the water flows down the stepped face as a coherent stream skimming over the steps and cushioned by the recirculating fluid trapped between them (Chanson, 1994).

At the downstream end of stepped channels, supercritical flow generally occurs and subcritical flow is observed at the end of the downstream stilling basin. Local scour around hydraulic structures is considered to be an important phenomenon affecting their stability. Estimation of local scour downstream of stepped spillways is an important task for hydraulic engineering. To date, many scour experiments have been carried out and empirical equations developed to estimate the scour depth that occurs downstream of spillways (Hassan and Narayanan, 1985; Breusers and Raudkivi, 1991; Oliveto and Comiello, 2009; Oliveto et al., 2011).

Tuna and Emiroglu (2011) investigated the scour hole profile downstream of stepped chutes. They also presented equilibrium scour depth values, the location of the maximum scour depth and the length of the scour hole.

Tuna (2012) studied the impact of the offtake channel base angle of stepped spillways on the scour hole, by means of a physical model. According to the experimental results it was revealed that a take-off angle of 30° is the optimum angle which gives the minimum longitudinal area and maximum depth of the scour hole.

Tuna and Emiroglu (2013) investigated the effect of step geometry on the dynamics of local scour processes downstream of a stepped chute. Experimental results showed

that the equilibrium depth of scour is highly dependent on the step geometry.

Farhoudi and Shayan (2014) investigated local scour downstream of adverse-slope stilling basins. Experiments were conducted for different stilling basin slopes. The experimental results revealed that the length of the scour profiles and the volume of eroded materials increased in accordance with the slope of the basin. The time evolution of scour hole dimensions and the equilibrium state were also defined.

Aminpour et al. (2016) investigated the time scale of local scour evolution downstream of stepped spillways. Experiments were conducted using stepped spillways with a height of 45 cm and 60 cm. The results show that the dimensions of the scour hole increase with densimetric Froude particle number. It was also revealed that the energy dissipation achieved by stepped spillways under specific conditions is higher than that of ogee spillways.

Elnikhely (2017) investigated local scour downstream of a spillway by means of cylinder blocks fixed on the back slope of the spillway. Based on the experimental results it was concluded that cylinder blocks can be used as an extra element over existing spillway structures for minimizing of scour.

The temporal evolution of scour depth (d_s) is also an important element in understanding the scour process. Few studies have focused on the temporal evolution of scour depth. The parameters having an effect on the scour process are the flow parameters, such as velocity (V), tailwater depth (h_t), fluid density and viscosity (ρ and μ), bed material (sediment) characteristics, and time.

The main objective of this study was to investigate the temporal variation of scour downstream of the stilling basin and to propose an empirical equation based on experimental data.

DIMENSIONAL ANALYSIS

In the local scour process, the effective dimensional parameters are: density of the water (ρ), kinematic viscosity of the water (ν), density of the bed material (ρ_s), median grain size of the bed material (d_{50}), step height (h), step width (b), stilling basin length (l), sill height (h_s), tailwater depth (h_t), velocity of the

*Corresponding author, email: mustafa.dogan@deu.edu.tr

Received 5 July 2018; accepted in revised form 13 June 2019

water (V), acceleration due to gravity (g) and time (t). Thus the value of the scour depth can be written as:

$$ds = f_1(\rho, \nu, \rho_s, d_{50}, h, b, l, h_s, ht, V, g, t) \quad (1)$$

The independent parameters ρ , ρ_s and g can be combined as g' where $g' = [(\rho_s - \rho)/\rho]g$ (Dey and Raikar, 2005) and under turbulent flow conditions the effect of ν can be neglected. In addition, Densimetric Froude particle number (F_d) has an important role in the temporal dimension of the scour process (Oliveto and Hager, 2002) and is defined as $F_d = V/g'd_{50}$. Dimensionless time parameter can be expressed as (Oliveto et al., 2011):

$$T = (\sqrt{g'd_{50}}/h) * t$$

The non-dimensional parameters were obtained by means of Buckingham π theorem as:

$$\frac{d_s}{h} = f_2\left(\frac{d_{50}}{h}, \frac{b}{h}, \frac{l}{h}, \frac{h_s}{h}, \frac{h_t}{h}, F_d, \frac{\sqrt{g'd_{50}}}{h}t\right) \quad (2)$$

Median grain size of the bed material (d_{50}) is considered in F_d and the non-dimensional time term. The tailwater depth (h_t) depends on the sill height (h_s). The values of (b/h) and (l/h) are constant for the present study. Thus the non-dimensional effective parameters can be written as:

$$\frac{d_s}{h} = f_3\left(\frac{h_t}{h}, F_d, \frac{\sqrt{g'd_{50}}}{h}t\right) \quad (3)$$

In sediment transport, the initiation of motion of sediment in a fluid flow can be determined by using the dimensionless Shields parameter (θ) which is defined as:

$$\theta = u_*^2 / (g'd_{50}) \quad (4)$$

So θ can also be used instead of F_d in Eq. 3.

EXPERIMENTAL SETUP

The experimental tests were carried out in a stepped flume with a height of 2.4 m as shown in Fig. 1. The flume was constructed as a 1/10 scale partial physical model of the Cine Adnan Menderes Dam stepped spillway in the Hydraulic Laboratory of the Civil Engineering Department at Dokuz Eylul University. The sidewalls were made of plexiglass material to observe the flow and scour process. The width of the rectangular flume was 0.10 m and the length of the stilling basin was 2.12 m. Bed material was placed in a sediment box with a height of 24 cm and length of 2.48 m. The water was pumped from a 30 m³ main reservoir to the stilling tank of the experimental set-up using a centrifugal pump.

Experiments were carried out by using bed materials of 4 different grain size distributions (0.55–5.45 mm), 2 different sill heights (5.0–7.5 cm) and 6 different flow rates (2.739–8.402 cm³/s). The step height was 12 cm during the experiments. Bed materials were placed without compaction or any other treatment. The disturbed bed was flattened to obtain a horizontal plane surface before starting a new experiment. Particle size distributions of the bed materials used in the experiments are given in Fig. 2.

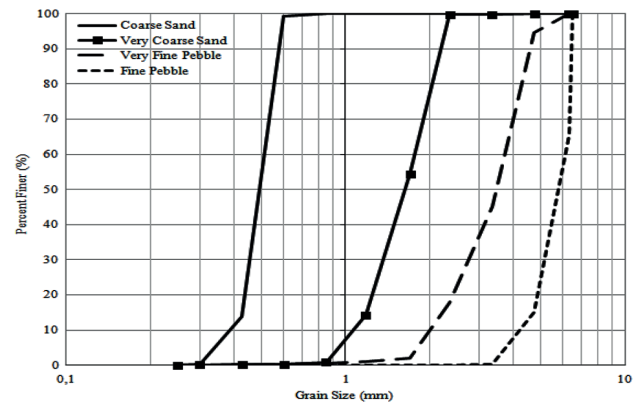


Figure 2. Particle size distributions of the bed materials

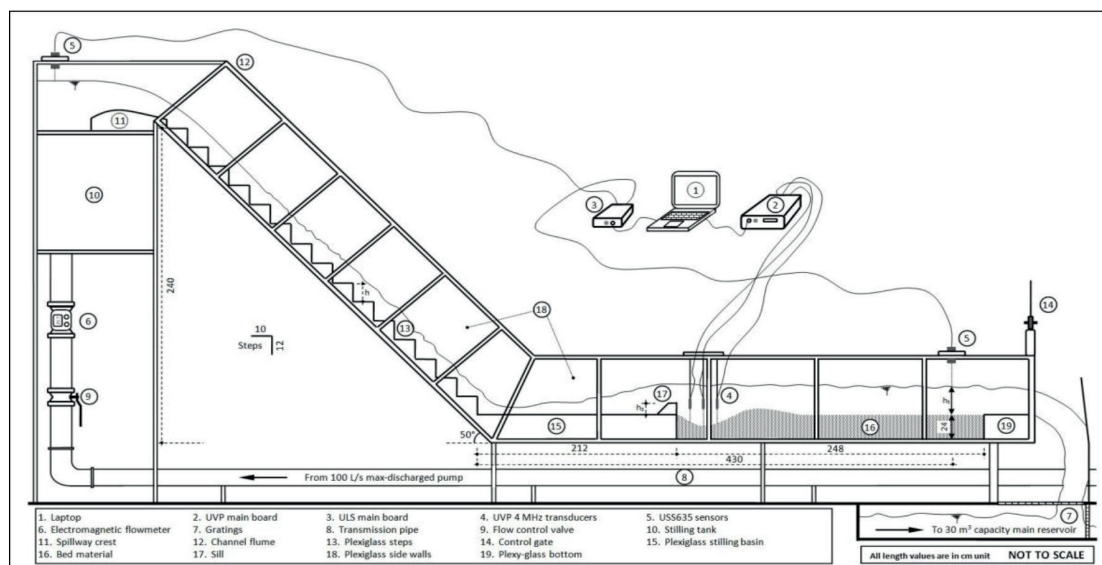


Figure 1. Schematic illustration of the experimental setup

Table 1. Characteristic parameters of the experimental variables: q is the unit flow rate, h_t is the tail-water depth, h_s is the sill height, V is the mean water velocity, u_* is the shear velocity and d_{50} is the median grain size of the bed material

Exp. No.	q (273.9–840.2) ($\text{cm}^3\cdot\text{s}^{-1}\cdot\text{cm}^{-1}$)	h_t (14.5–24.5) (cm)	h_s (5.0–7.5) (cm)	d_{50} (0.55–5.45) (mm)	V ($\text{cm}\cdot\text{s}^{-1}$)	u_* ($\text{cm}\cdot\text{s}^{-1}$)
E-1	273.9	14.5	5.0	0.55	18.89	1.07
E-2	478.5	17.0	5.0	0.55	28.15	1.56
E-3	690.3	19.0	5.0	0.55	36.33	1.98
E-4	738.3	20.5	5.0	0.55	36.01	1.95
E-5	822.7	21.2	5.0	0.55	38.81	2.09
E-6	840.2	21.5	5.0	0.55	39.08	2.10
E-7	273.9	16.0	7.5	0.55	17.12	0.96
E-8	478.5	18.5	7.5	0.55	25.86	1.42
E-9	690.3	20.0	7.5	0.55	34.52	1.87
E-10	738.3	21.0	7.5	0.55	35.16	1.89
E-11	822.7	23.0	7.5	0.55	35.77	1.90
E-12	840.2	24.5	7.5	0.55	34.29	1.81
E-13	273.9	14.5	5.0	1.85	18.89	1.29
E-14	478.5	17.0	5.0	1.85	28.14	1.87
E-15	690.3	19.0	5.0	1.85	36.33	2.38
E-16	738.3	20.5	5.0	1.85	36.01	2.33
E-17	822.7	21.2	5.0	1.85	38.81	2.49
E-18	840.2	21.5	5.0	1.85	39.08	2.51
E-19	273.9	16.0	7.5	1.85	17.12	1.15
E-20	478.5	18.5	7.5	1.85	25.86	1.70
E-21	690.3	20.0	7.5	1.85	34.52	2.24
E-22	738.3	21.0	7.5	1.85	35.16	2.26
E-23	822.7	23.0	7.5	1.85	35.77	2.27
E-24	840.2	24.5	7.5	1.85	34.29	2.15
E-25	273.9	14.5	5.0	3.45	18.89	1.45
E-26	478.5	17.0	5.0	3.45	28.15	2.09
E-27	690.3	19.0	5.0	3.45	36.33	2.65
E-28	738.3	20.5	5.0	3.45	36.01	2.59
E-29	822.7	21.2	5.0	3.45	38.81	2.77
E-30	840.2	21.5	5.0	3.45	39.08	2.78
E-31	273.9	16.0	7.5	3.45	17.12	1.29
E-32	478.5	18.5	7.5	3.45	25.86	1.89
E-33	690.3	20.0	7.5	3.45	34.52	2.49
E-34	738.3	21.0	7.5	3.45	35.16	2.51
E-35	822.7	23.0	7.5	3.45	35.77	2.52
E-36	840.2	24.5	7.5	3.45	34.29	2.39
E-37	273.9	14.5	5.0	5.45	18.89	1.60
E-38	478.5	17.0	5.0	5.45	28.14	2.31
E-39	690.3	19.0	5.0	5.45	36.33	2.92
E-40	738.3	20.5	5.0	5.45	36.01	2.85
E-41	822.7	21.2	5.0	5.45	38.81	3.05
E-42	840.2	21.5	5.0	5.45	39.08	3.06
E-43	273.9	16.0	7.5	5.45	17.12	1.42
E-44	478.5	18.5	7.5	5.45	25.86	2.09
E-45	690.3	20.0	7.5	5.45	34.52	2.74
E-46	738.3	21.0	7.5	5.45	35.16	2.77
E-47	822.7	23.0	7.5	5.45	35.77	2.76
E-48	840.2	24.5	7.5	5.45	34.29	2.62

During the experiments three different measurements were performed via electromagnetic and ultrasonic techniques (Fig. 1). Flow rates were measured precisely by using the

electromagnetic flow meter. Water levels were measured at 2 points along the experimental set-up by means of ULS (Ultrasonic Level Sensors (labelled 4 in Fig.1) – USS635 sensors (labelled 5 in Fig.1)). The locations of the USS635 sensors are shown in Fig. 1. The temporal scour depths were measured by using the ultrasonic velocity profiler (UVP). Three UVP transducers were located on the sediment box to detect the location of the maximum scour depth. Although, the UVP device is designed mainly to obtain the velocity profile, in this study it was used to determine the time-varying bed material motion in an indirect way by placing UVP transducers vertically. A detailed explanation can be found in the study of Guney et al. (2013).

The characteristic values of the experimental parameters are given in Table 1. The efficient and substantial non-dimensional parameters are shown in Table 2.

RESULTS AND DISCUSSION

Experiments were carried out to determine temporal scour variation at the downstream of the stepped channel under different experimental conditions. In order to investigate the effect of the bed material, sill height and flow rate, 4 different bed materials, 2 different sill heights and 6 different flow rates were used. According to the experimental results, scour depth is a function of sill height, tailwater depth and bed material characteristics. Bed material was varied between 0.55 mm and 5.45 mm. Temporal variation of scour depth measured at the centreline of the flume for constant sill height and grain size of the bed material under different flow rates and tailwater depths is given in Fig. 3. According to the experimental results it is revealed that the scour depth increased with the mean flow velocity in all experiments. Maximum scour depth occurred in the case of minimum grain size of the bed material, maximum rate of the flow velocity and lower sill height. The scour depth decreased with an increase in both sill height and median grain size of bed material, as expected. The scour depth builds up over time and when the scour depth reaches its maximum value it fluctuates around the maximum value. The final scour depth values observed at the end of each experiment are given in Table 3.

Least squares method was used to obtain the best-fitting curve by minimizing the sum of the squares of the residuals between observed and calculated values. Equation 5 and Equation 6 are proposed by using the experimental data obtained from the present study. The powers of non-dimensional parameters in the proposed equations were obtained by means of the least squares method. The differences between these two equations are the third terms. Shields parameter (θ) was used in Eq. 5 while densimetric Froude particle number (F_d) was used in Eq. 6. To calculate the value of the Densimetric Froude particle number only mean water velocity (V) is used, but to calculate the value of the Shields parameter, shear velocity (u_*) should first be calculated. The calculated scour depths obtained by using these equations were compared with those measured during the experiments. Figure 4 represents measured and computed temporal variation of scour depth. Based on Fig. 4 one can say that the measured and computed scour depths are in good agreement.

$$\frac{d_s}{h} = 0.47 \left(\frac{h_t}{h} \right)^{0.65} \theta^{0.1} (1 - e^{-0.1T}) \quad (5)$$

Table 2. Summary of the non-dimensional parameters. Re_d is the grain Reynolds number and it can be defined as $u_* d_{50}/\nu \cdot \theta$ and F_d are the Shields parameter and the densimetric Froude particle number, respectively.

Experiment Number	Re_d	θ	F_d	$\frac{h_t}{h}$
E-1	5.9	0.0129	2.0	1.21
E-2	8.6	0.0273	3.0	1.42
E-3	10.9	0.0442	3.9	1.58
E-4	10.7	0.0425	3.8	1.71
E-5	11.5	0.0489	4.1	1.77
E-6	11.5	0.0494	4.1	1.79
E-7	5.3	0.0103	1.8	1.33
E-8	7.8	0.0225	2.7	1.54
E-9	10.3	0.0393	3.7	1.67
E-10	10.4	0.0402	3.7	1.83
E-11	10.5	0.0407	3.8	1.92
E-12	9.9	0.0367	3.6	2.04
E-13	23.9	0.0056	1.1	1.21
E-14	34.7	0.0117	1.6	1.42
E-15	44.0	0.0189	2.1	1.58
E-16	43.0	0.0181	2.1	1.71
E-17	46.1	0.0208	2.2	1.77
E-18	46.3	0.0210	2.3	1.79
E-19	21.3	0.0044	1.0	1.33
E-20	31.4	0.0096	1.5	1.54
E-21	41.4	0.0167	2.0	1.67
E-22	41.8	0.0171	2.0	1.83
E-23	42.0	0.0172	2.1	1.92
E-24	39.8	0.0155	2.0	2.04
E-25	49.9	0.0037	0.8	1.21
E-26	72.2	0.0078	1.2	1.42
E-27	91.3	0.0125	1.5	1.58
E-28	89.2	0.0120	1.5	1.71
E-29	95.6	0.0137	1.6	1.77
E-30	96.0	0.0139	1.7	1.79
E-31	44.4	0.0030	0.7	1.33
E-32	65.3	0.0064	1.1	1.54
E-33	85.9	0.0111	1.5	1.67
E-34	86.7	0.0113	1.5	1.83
E-35	86.8	0.0113	1.5	1.92
E-36	82.3	0.0102	1.5	2.04
E-37	92.2	0.0028	0.6	1.21
E-38	132.9	0.0057	0.9	1.42
E-39	167.7	0.0091	1.2	1.58
E-40	163.7	0.0087	1.2	1.71
E-41	175.3	0.0100	1.3	1.77
E-42	176.0	0.0101	1.3	1.79
E-43	81.8	0.0022	0.6	1.33
E-44	120.0	0.0047	0.8	1.54
E-45	157.7	0.0081	1.1	1.67
E-46	159.1	0.0082	1.2	1.83
E-47	159.0	0.0082	1.2	1.92
E-48	150.6	0.0074	1.1	2.04

$$\frac{d_s}{h} = 0.30 \left(\frac{h_t}{h}\right)^{0.55} F_d^{0.15} (1 - e^{-0.1T}) \quad (6)$$

Table 3. Final scour depth values

Experiment Number	d_s^f (mm)	Experiment Number	d_s^f (mm)
E-1	46.3	E-25	41.2
E-2	51.9	E-26	46.5
E-3	57.0	E-27	48.8
E-4	61.7	E-28	52.6
E-5	64.0	E-29	53.2
E-6	64.9	E-30	54.3
E-7	44.6	E-31	40.6
E-8	48.3	E-32	45.6
E-9	55.0	E-33	47.7
E-10	57.0	E-34	49.7
E-11	58.5	E-35	51.0
E-12	59.4	E-36	52.3
E-13	42.3	E-37	40.5
E-14	48.0	E-38	44.0
E-15	50.5	E-39	48.0
E-16	55.7	E-40	51
E-17	57.2	E-41	52.3
E-18	58.2	E-42	53.0
E-19	41.0	E-43	39.7
E-20	46.5	E-44	43.0
E-21	48.5	E-45	46.8
E-22	51.9	E-46	48.8
E-23	52.5	E-47	50.1
E-24	53.5	E-48	51.5

where T is the dimensionless time parameter which can be expressed as follows:

$$T = \frac{\sqrt{g d_{50}}}{h} t \quad (7)$$

where t is time.

Equations 5 and 6 were evaluated in terms of scatter index (SI) defined as follows:

$$SI (\%) = \frac{\sqrt{\frac{\sum_{i=1}^n (d_{s,measured_i} - d_{s,computed_i})^2}{n}}}{d_{s,measured}} \cdot 100 \quad (8)$$

The SI values were 8.73% and 8.25% for Eq. 5 and Eq. 6, respectively. According to the SI (%) values it is revealed that both equations can be used to predict temporal scour depth values.

The dimensionless scour depth values (d_s^f/h) obtained at the end of the experiments were plotted versus Shields parameter (θ) and densimetric Froude particle number (F_d) and given in Figs 5 and 6. Dimensionless scour depth increases with both Shields parameter (θ) and densimetric Froude particle number (F_d).

Maximum scour depths, measured after each experiment, were also compared with those calculated using Eqs 5 and 6. The results are given in Fig. 7.

The R^2 (coefficient of determination) values were computed as 0.866 and 0.865 for Eq. 5 and Eq. 6, respectively. These results also support the results obtained by using SI (%) values.

To check the validity of the proposed equations the Fisher (f) test was also applied. The parameter f is defined as follows:

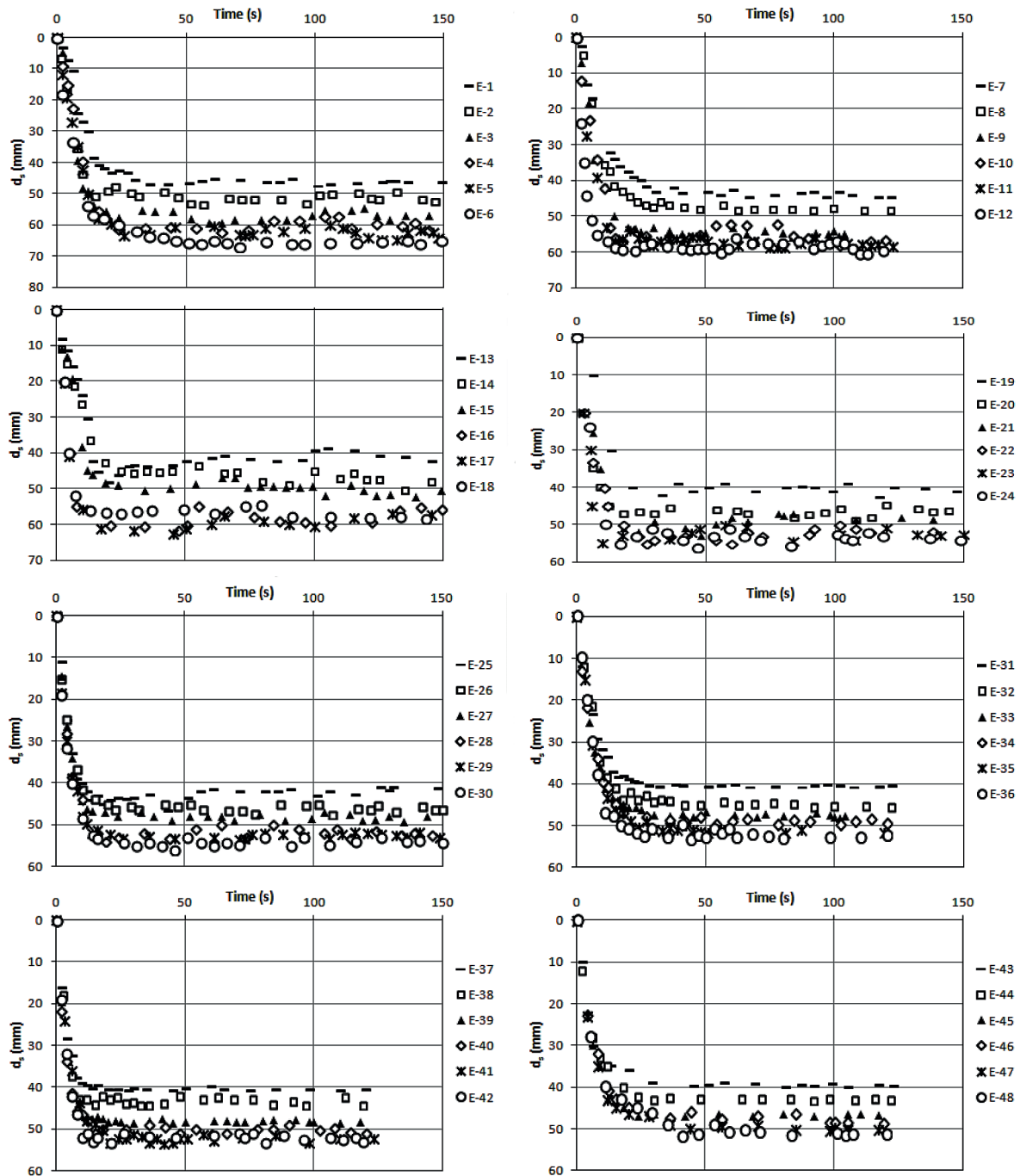


Figure 3. Temporal variation of scour depth measured during the experiments

$$f = \frac{SSR/v_1}{SSE/v_2} \quad (9)$$

where SSR is the sum of squared residuals, SSE is the sum of squares for error, v_1 is the number of the independent variables (k), and $v_2 = n - k - 1$, (n is the number of data points). SSR and SSE can be calculated from the following equations:

$$SSR = \sum_{i=1}^n (d_{si}^{computed} - \overline{d_{si}^{measured}})^2 \quad (10)$$

$$SSE = \sum_{i=1}^n (d_{si}^{measured} - d_{si}^{computed})^2 \quad (11)$$

where $\overline{d_{si}^{measured}}$ and $\overline{d_{si}^{computed}}$ are the arithmetic mean of the measured and computed scour depth values, respectively.

For the selected 0.01 significance level, the critical value of f is 4.22. If the computed f value is greater than the critical value that means that the fit of the proposed equation is statistically significant. The f values were computed as 66.3 and 65.1 for Eqs 4 and 5, respectively. Therefore, one can say that the proposed equations were acceptable.

As a consistent estimator, sum of squared errors (SSE) was computed for Eq. 5 and Eq. 6. SSE is the error between the measured and computed values:

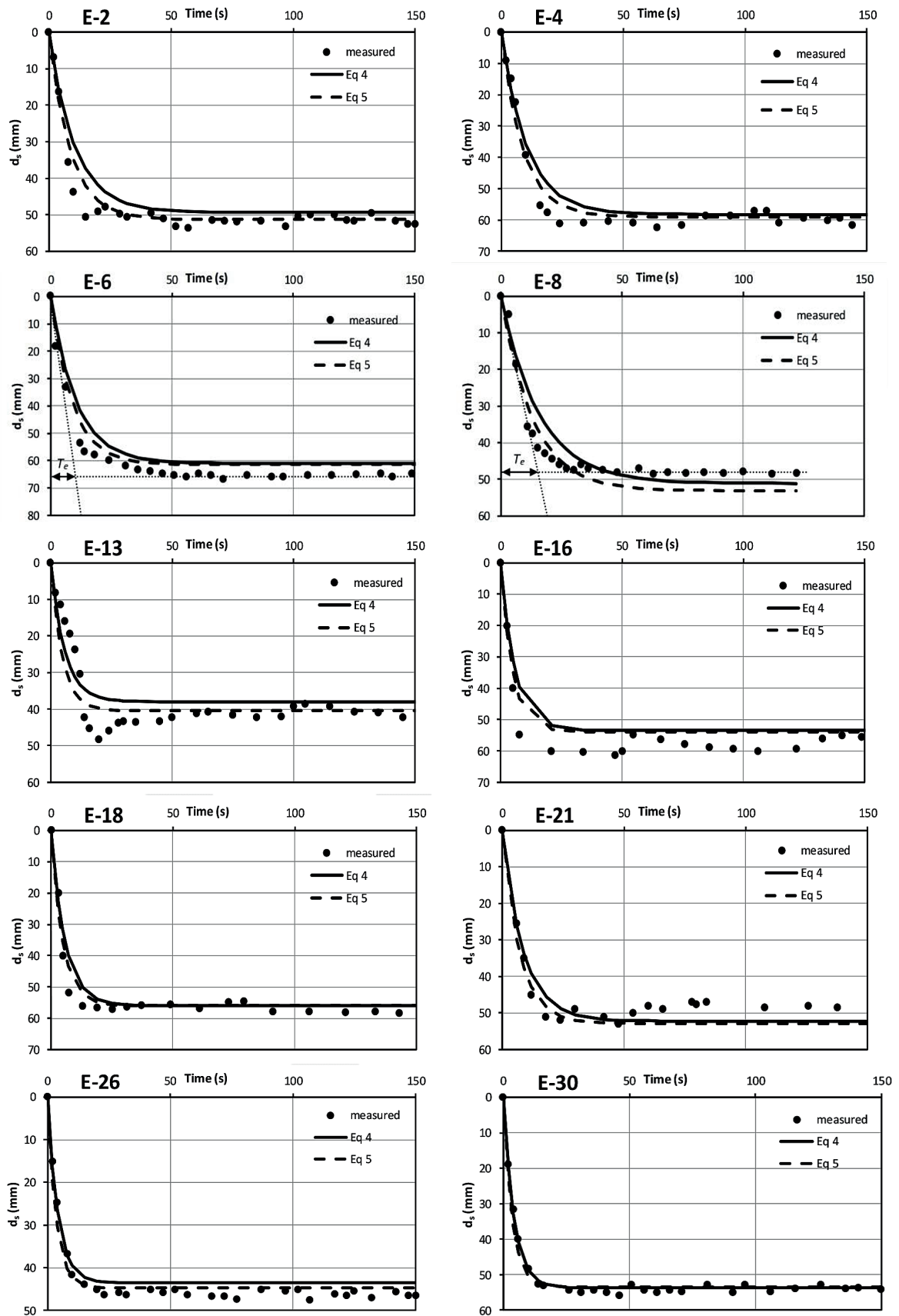


Figure 4. Measured and calculated scour depth values

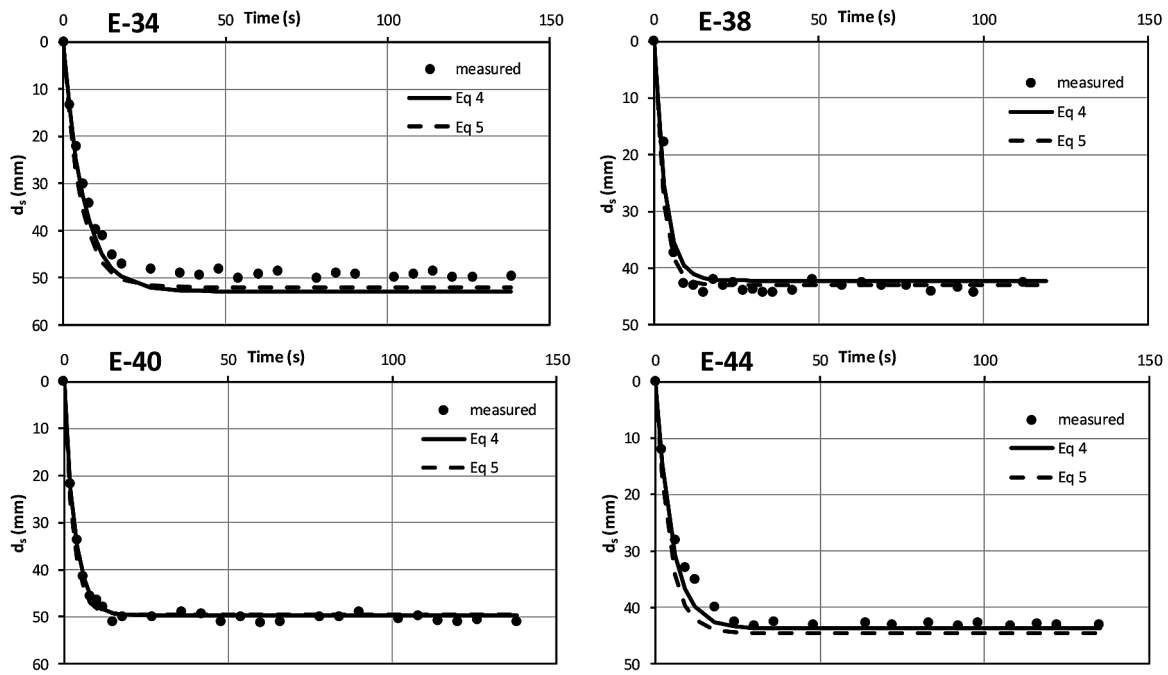


Figure 4. (cont.)

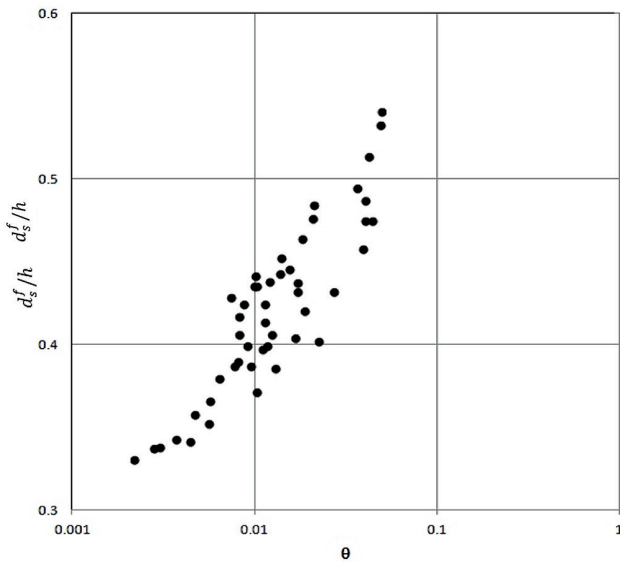


Figure 5. Dimensionless scour depth values (d_s^f/h) versus Shields parameter (θ)

$$SSE(\%) = \frac{\sum (d_{s,measured} - d_{s,computed})^2}{\sum (d_{s,measured})^2} \cdot 100 \quad (12)$$

The SSE values were calculated as 0.24% and 0.23% for Eqs 5 and 6, respectively. According to the statistical results it was revealed that the proposed equations are acceptable.

In the present study, the time scale parameter (T_e) was also investigated. The time scale can be described as the time period in which the scour depth reaches the equilibrium stage. The values of the time scale (T_e) were determined from the time-dependent scour depth graphics for each experiment. The non-dimensional time scale (T_e^*) parameter can be defined as:

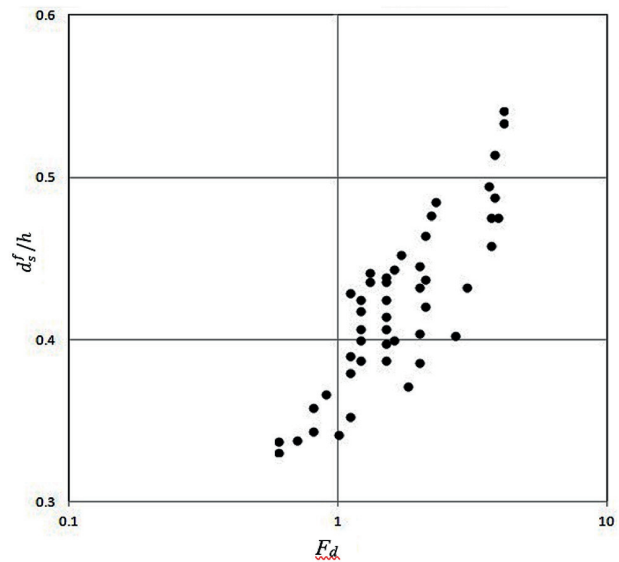


Figure 6. Dimensionless scour depth values (d_s^f/h) versus densimetric Froude particle number (F_d)

$$T_e^* = \frac{\sqrt{g'd_{50}}}{h} T_e \quad (13)$$

Based on the experimental findings two relationships between T_e^* versus θ and T_e^* versus F_d are proposed. Equations 14 and 15 are the empirical equations to estimate the time scale parameter for the scour process at the downstream part of the stepped channels. The powers of these relations were also obtained by means of the least squares method. So these powers give the best fit.

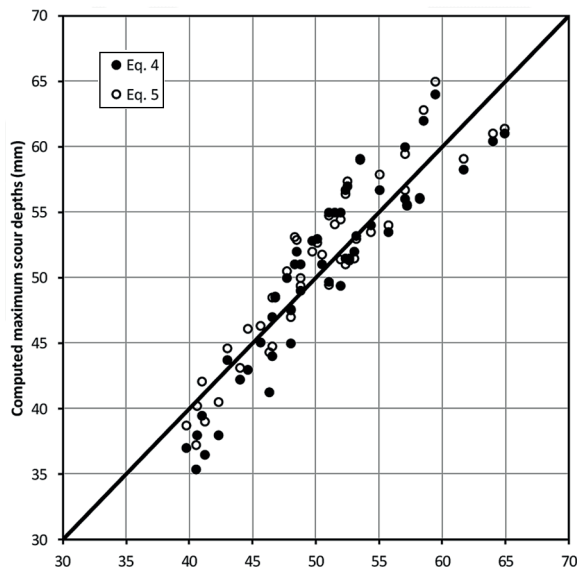


Figure 7. Measured and computed scour depth values

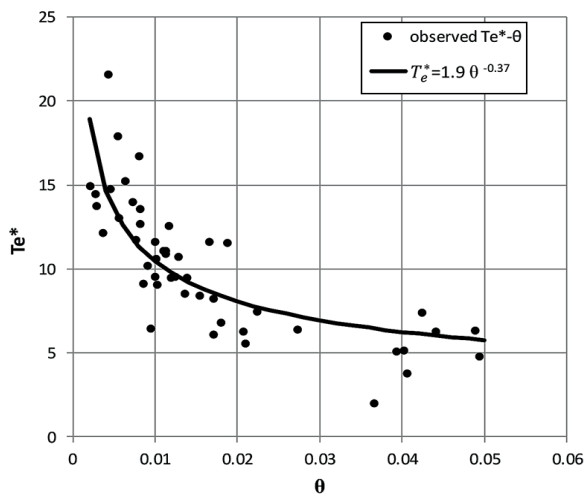


Figure 8. Non-dimensional time scale values versus Shields parameter

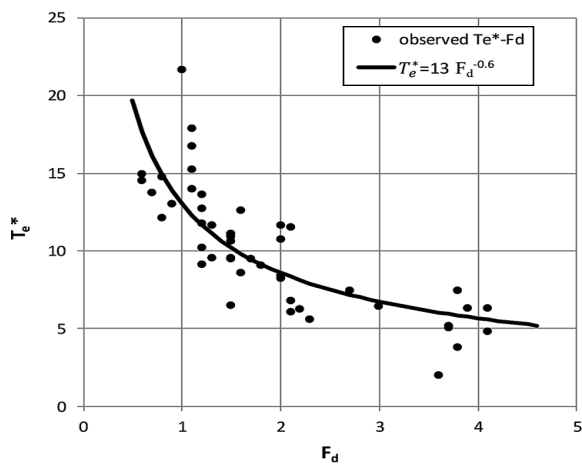


Figure 9. Non-dimensional time scale values versus densimetric Froude particle number

$$T_e^* = 1.9\theta^{-0.37} \quad (14)$$

$$T_e^* = 13 F_d^{-0.6} \quad (15)$$

The observed non-dimensional time scale (T_e^*) values were compared with those calculated by using Eq. 14 and Eq. 15, as shown in Fig. 8 and Fig. 9, respectively.

CONCLUSIONS

In this study, temporal local scour variations at the downstream part of a stepped channel were investigated experimentally. Four different bed materials, two different sill heights and six different flow rates were used. Two empirical equations predicting the temporal variation of scour depth were proposed using the experimental findings. The validity of the equations was also investigated. The following conclusions can be drawn from the results of this study:

- The scour depth increased with the mean flow velocity and the maximum scour depth occurred in the case of minimum grain size of the bed material, maximum rate of the flow and lower sill height
- The scour depth decreased with an increase in both sill height and median grain size of bed material
- The dimensionless scour depth (d_s/h) increases with both Shields parameter (θ) and densimetric Froude particle number (F_d)
- The temporal variation of scour depth was calculated by means of Eq. 5 and Eq. 6, which include Shields parameter (θ) and densimetric Froude particle number (F_d), respectively. The results of both equations are in good agreement with experimental findings.
- To indicate the best-fit equation scatter index (SI%) values were computed for both equations by using the observed and calculated temporal variations of the scour depth. There is not a significant difference between the SI values and it is revealed that both equations can be used to predict temporal scour depth variation.
- A Fisher (f) test was performed to check the statistical significance of the proposed equations. The computed f values for both equations are larger than the critical value of f determined for the selected 0.01 significance level. Therefore, the significance of the proposed equations was confirmed.
- The SSE% values were calculated as 0.24% and 0.23% for Eq. 5 and Eq. 6, respectively. According to the statistical results it is revealed that the proposed equations are acceptable.
- According to the time-dependent experimental results; Eq. 14 and Eq. 15 were proposed to estimate the non-dimensional time scale values which is the time period during the equilibrium scour develops.

REFERENCES

- AMINPOUR Y, FARHOUDI J, SHAYAN HK and ROSHAN R (2016) Characteristics and time scale of local scour downstream stepped spillways. *Sci. Iranica* 25 (2) 532–542. <https://doi.org/10.24200/sci.2017.4187>
- BREUSERS HNC and RAUDKIVI AJ (1991) *Scouring Hydraulic Structures Design Manual*. A. A. Balkema, Rotterdam.
- CHANSON H (1994) Hydraulics of skimming flows over stepped channels and spillways. *J. Hydraul. Res.* 32 3 445–460. <https://doi.org/10.1080/00221689409498745>
- DEY S and RAIKAR RV (2005) Scour in long contractions. *J.*

- Hydraul. Eng.* **131** (12) 1036–1049. [https://doi.org/10.1061/\(asce\)0733-9429\(2005\)131:12\(1036\)](https://doi.org/10.1061/(asce)0733-9429(2005)131:12(1036))
- ELNIKHELY EA (2017) Investigation and analysis of scour downstream of a spillway. *Ain Shams Eng. J.* <https://doi.org/10.1016/j.asej.2017.03.008>
- FARHOUDI J and SHAYAN HK (2014) Investigation on local scour downstream of adverse stilling basins. *Ain Shams Eng. J.* **5** 361–375. <https://doi.org/10.1016/j.asej.2014.01.002>
- GUNEY MS, BOMBAR G, OZGENC AKSOY A and DOGAN M (2013) Use of Uvp to investigate the evolution of bed configuration. *KSCE J. Civ. Eng.* **17** (5) 1–10. <https://doi.org/10.1007/s12205-013-0131-5>
- HASSAN MK and NARAYANAN R (1985) Local scour downstream of an apron. *J. Hydraul. Eng.* **111** (11) 1371–1385.
- OLIVETO G and COMUNIELLO V (2009) Local scour downstream of positive-step stilling basins. *J. Hydraul. Eng.* **135** (10) 846–851. <https://doi.org/10.1080/00221686.2010.538593>
- OLIVETO G, COMUNIELLO V and BULBULE T (2011) Time-dependent local scour downstream of positive-step stilling basins. *J. Hydraul. Res.* **49** (1) 105–112. <https://doi.org/10.1080/00221686.2010.538593>
- OLIVETO G and HAGER WH (2002) Temporal evolution of clear-water pier and abutment scour. *J. Hydraul. Eng.* **128** (9) 811–820. [https://doi.org/10.1061/\(asce\)0733-9429\(2002\)128:9\(811\)](https://doi.org/10.1061/(asce)0733-9429(2002)128:9(811))
- TUNA MC and EMIROGLU ME (2011) Scour profiles at downstream of cascades. *Sci. Iranica* **18** (3) 338–347. <https://doi.org/10.1016/j.scient.2011.05.040>
- TUNA MC (2012) Effect of offtake channel base angle of stepped spillway on scour hole. *IJST, Trans. Civ. Eng.* **36** (C2) 239–251.
- TUNA MC and EMIROGLU ME (2013) Effect of step geometry on local scour downstream of stepped chutes. *Arab. J. Sci. Eng.* **38** 579–588. <https://doi.org/10.1007/s13369-012-0335-x>
-

Assessment of CFD turbulence models for free surface flow simulation and 1-D modelling for water level calculations over a broad-crested weir floodway

Ahmed M Helmi^{1*}

¹Irrigation and Hydraulics Department, Faculty of Engineering, Cairo University, Egypt

ABSTRACT

Floodways, where a road embankment is permitted to be overtopped by flood water, are usually designed as broad-crested weirs. Determination of the water level above the floodway is crucial and related to road safety. Hydraulic performance of floodways can be assessed numerically using 1-D modelling or 3-D simulation using computational fluid dynamics (CFD) packages. Turbulence modelling is one of the key elements in CFD simulations. A wide variety of turbulence models are utilized in CFD packages; in order to identify the most relevant turbulence model for the case in question, 96 3-D CFD simulations were conducted using Flow-3D package, for 24 broad-crested weir configurations selected based on experimental data from a previous study. Four turbulence models (one-equation, $k-\varepsilon$, RNG $k-\varepsilon$, and $k-\omega$) were examined for each configuration. The volume of fluid (VOF) algorithm was adopted for free water surface determination. In addition, 24 1-D simulations using HEC-RAS-1-D were conducted for comparison with CFD results and experimental data. Validation of the simulated water free surface profiles versus the experimental measurements was carried out by the evaluation of the mean absolute error, the mean relative error percentage, and the root mean square error. It was concluded that the minimum error in simulating the full upstream to downstream free surface profile is achieved by using one-equation turbulence model with mixing length equal to 7% of the smallest domain dimension. Nevertheless, for the broad-crested weir upstream section, no significant difference in accuracy was found between all turbulence models and the one-dimensional analysis results, due to the low turbulence intensity at this part. For engineering design purposes, in which the water level is the main concern at the location of the flood way, the one-dimensional analysis has sufficient accuracy to determine the water level.

Keywords: broad-crested weirs, floodway, turbulence, free surface flow, Flow-3D, CFD, VOF, HEC-RAS

INTRODUCTION

Broad-crested weirs are simple flow control structures widely used in open channels, in addition to their use in flow measurements. The broad-crested weir was defined by Henderson (1966 p. 211 and Chanson (2004 p. 48), as 'a flat-crested structure with a crest length large compared to the flow thickness for the streamlines to be parallel to the crest invert and the pressure distribution to be hydrostatic'. Floodways or Irish crossings (Fig. 1) are designed as broad-crested weirs when a portion of a road embankment is permitted to be overtopped by floodwater. Road cross-section elements have a wide range of variations depending on the geometric design and backfill material properties, which leads to difficulty in accurate determination of discharge, Cd. The determination of the flow depth above the road embankment may require physical modelling or numerical simulation.

Hargreaves et al. (2007) and Ahmed and Mohamed (2011) validated the adequacy of the Volume of Fluid (VOF) algorithm for free surface calculations. Yazdi et al. (2010) used the VOF algorithm and $k-\omega$ turbulence model in 3-D simulation of free surface flow around a spur dike, and it was concluded that sufficient accuracy was achieved when the 3-D model results were compared to the flume results. Sarker and Rhodes (2004) verified the efficiency of the $k-\omega$ turbulence model to simulate the free surface flow over a

rectangular broad-crested weir. Maghsoodi et al. (2012) used the VOF algorithm, and $k-\varepsilon$ turbulence model in 3-D simulation of free surface flow over rectangular and broad-crested submerged weirs with variable crest widths and upstream/downstream-facing slopes. The simulation was found to be in good agreement with the experimental results, which has also been verified by Joongcheol and Nam (2015). Stefan et al. (2011) compared SSIIM-2, and Flow-3D CFD Packages in simulating the flow over broad-crested weirs and concluded that Flow-3D achieved the same accuracy as SSIIM-2 with a lower number of grid cell and a lower requirement for computer resources, due to the use of higher-order integration to compute the movement of the water surface. Hossein and Sayed (2013) constructed a physical experiment as well as a 3-D simulation for flow over broad-crested weirs. Their study was for a single geometric setting of a rectangular broad-crested weir with rounded corners, and 4 values of discharge (2, 3, 4, and 6 L/s); they concluded that the RNG $k-\varepsilon$ model has the least errors in obtaining the water surface profile. Tanase et al. (2015) verified the efficiency of the RNG $k-\varepsilon$ turbulence model to simulate the free surface flow over a rectangular broad-crested weir with patterned and smooth top surface. Shaymaa et al. (2017a) tested the adequacy of the CFD turbulence model to simulate the free surface profile for a rectangular and stepped broad-crested weir, for one discharge for each geometry, and found that $k-\varepsilon$ gives the least error among the tested models. Shaymaa et al. (2017b) compared the use of 2-D and 3-D numerical simulation using the $k-\varepsilon$ turbulence model to simulate the

*Corresponding author, email: Ahmed.helmi@eng.cu.edu.eg
Received 14 April 2018; accepted in revised form May 2019

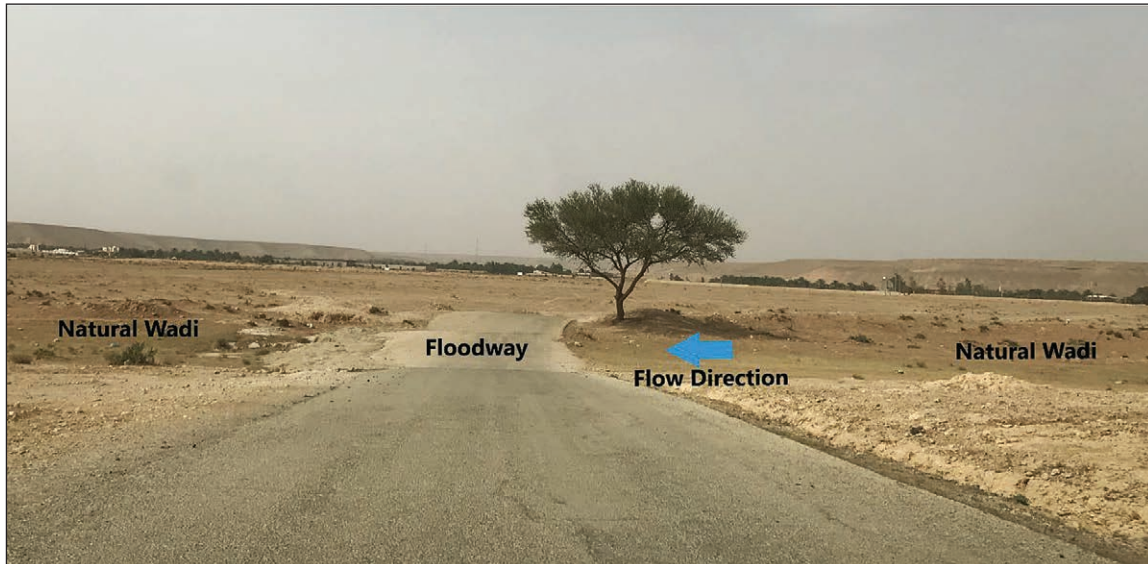


Figure 1. Floodway crossing

free surface flow over the broad-crested weir and almost the same accuracy was obtained. In addition to the studies mentioned above, the flow above a broad-crested weir has been studied by Kirkgov et al. (2008), Hoseini et al. (2013), Kassaf et al. (2016) and Shaker and Sarhan (2017), whether experimentally or by using CFD techniques, with no achieved agreement achieved on the suitable turbulence model for simulation of the case in question.

The coefficient of discharge of broad-crested weirs has been experimentally studied by Fritz and Hager (1998) without considering the effect of the weir's upstream-facing slope. The impact of the weir's upstream-facing slope has been studied by Sargison and Percy (2009), Shaymaa et al. (2015), Aysegul and Mustafa (2016), and Jiang et al. (2018). It was agreed by all researchers that the coefficient of discharge increases when decreasing the angle of the weir upstream-facing slope.

It is clearly shown from all previous studies that the trend of using CFD is increasing and that CFD has been progressively developed with regard to the algorithms used, based on the development of computational and data storage resources. A general agreement between most of the researchers is that the VOF algorithm is sufficient for simulating the free surface water profile.

The key elements of CFD are the grid generation, the algorithm, and turbulence modelling. The first two components have precise mathematical theories, achieving an accurate mathematical precision for the turbulence, which is a rapid spatial and temporal random fluctuation of the various flow characteristics, by using mass and momentum conservation equations. It is mandatory to use a very small time step and grid cell resolution to capture the details of the fluid properties fluctuation spectrum. Up until now this has not been possible due to the limitations of available computer processing and storage (Flow-3D, 2016). Many simplified models are available to describe the turbulence effect on flow characteristics.

There is no consensus among the various literature sources reviewed regarding the most appropriate turbulence model to be used in the CFD simulation of free surface problems. The aim of this research was therefore to:

- Investigate the most popular turbulence models, starting with their original forms, and how they are presented in the latest CFD packages, to validate the adequacy of the turbulence models to simulate free surface flow in the case of flood-ways, by validating the simulated water surface profile against a previous experimental study.
- Test the adequacy of free one-dimensional software (HEC-RAS) to be used by drainage engineers during the design of floodway crossings

Turbulence models

Turbulence models are classified into: algebraic (zero-equation), one-equation, two-equation, and stress-transport models (Wilcox, 2006).

The turbulence kinetic energy equation is the basis of the one-equation turbulence models. The incomplete part in these models is the relation between the turbulence length scale and the domain dimensions. The turbulence kinetic energy per unit mass (k) was chosen by Prandtl (1945) as the basis of the velocity scale (Eq. 1):

$$k = \frac{1}{2}(U'^2 + V'^2 + W'^2) \quad (1)$$

The one-equation model describing the transport equation for the turbulence kinetic energy is given by:

$$\frac{\partial k}{\partial t} + U_j \frac{\partial k}{\partial x_j} = \tau_{ij} \frac{\partial U_i}{\partial x_j} - \varepsilon + \frac{\partial}{\partial x_j} \left\{ \left(\nu + \frac{\nu_t}{\sigma_k} \right) \frac{\partial k}{\partial x_j} \right\} \quad (2)$$

$$\varepsilon = C_D \frac{k^{3/2}}{l} \quad (3)$$

$$\nu_t = C_D \frac{k^2}{\varepsilon} = k^{1/2} l \quad (4)$$

$$\tau_{ij} = 2\nu_t S_{ij} - \frac{2}{3} k \delta_{ij} \quad (5)$$

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (6)$$

The closure coefficient (σ_k), C_D and the length scale (l) have to be specified prior to application of the equation. The values

of $\sigma_k = 1.0$, C_d ranges between 0.07 and 0.09, and a mixing length similar to that of the mixing length model by Prandtl (1925), was tested by Emmons (1954), and Glushko (1965) successfully. The turbulence length scale (l) can be reasonably assumed to equal to 7% of the smallest domain calculation dimension (Fard and Boyaghchi, 2007).

Kolmogorov (1942) proposed the $k-\omega$ turbulence model. This was the first two-equation turbulence model. As did Prandtl (1945), Kolmogorov selected the turbulence kinetic energy as one of his turbulence parameters, and the second parameter (ω) represents the rate of dissipation of energy per unit of volume and time, where β and σ are the closure coefficients of the model (Eq. 7):

$$\frac{\partial \omega}{\partial t} + u_j \frac{\partial \omega}{\partial x_j} = -\beta \omega^2 + \frac{\partial}{\partial x_j} \left(\sigma v_j \frac{\partial \omega}{\partial x_j} \right) \quad (7)$$

Wilcox (1988, 1998), Speziale et al. (1992), Peng et al. (1997), Kok (2000) and Hellsten (2005) added the production term to the Kolmogorov ($k-\omega$) model. Wilcox (2006) represented the model in the following form:

$$\frac{\partial k}{\partial t} + U_j \frac{\partial k}{\partial x_j} = \tau_{ij} \frac{\partial U_i}{\partial x_j} - \beta^* \rho k \omega + \frac{\partial}{\partial x_j} \left\{ \left(\nu + \sigma^* \frac{k}{\omega} \right) \frac{\partial k}{\partial x_j} \right\} \quad (8)$$

$$\frac{\partial \omega}{\partial t} + U_j \frac{\partial \omega}{\partial x_j} = \alpha \frac{\omega}{k} \tau_{ij} \frac{\partial U_i}{\partial x_j} - \beta \omega^2 + \frac{\sigma_d}{\omega} \frac{\partial k}{\partial x_j} \frac{\partial \omega}{\partial x_j} + \frac{\partial}{\partial x_j} \left\{ \left(\nu + \sigma \frac{k}{\omega} \right) \frac{\partial \omega}{\partial x_j} \right\} \quad (9)$$

The closure coefficients and auxiliary relations for the $k-\omega$ model:

$$\beta^* = \frac{9}{100}, \quad \sigma^* = \frac{3}{5}, \quad \alpha = \frac{13}{25}, \quad \beta = \beta_o f_\beta, \quad \sigma = \frac{1}{2} \quad (10)$$

$$\sigma_d = \begin{cases} 0 & \frac{\partial k}{\partial x_j} \frac{\partial \omega}{\partial x_j} \leq 0 \\ \frac{1}{8} & \frac{\partial k}{\partial x_j} \frac{\partial \omega}{\partial x_j} > 0 \end{cases} \quad (11)$$

$$\beta_o = 0.0708, \quad f_\beta = \frac{1 + 85 \chi_\omega}{1 + 100 \chi_\omega}, \quad \chi_\omega = \left| \frac{\Omega_{ij} \Omega_{jk} S_{ki}}{(\beta^* \omega)^2} \right| \quad (12)$$

$$\Omega_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right), \quad S_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (13)$$

The $k-\varepsilon$ model was initially developed by Chou (1945), Davydov (1961) and Harlow and Nakayama (1968), and was in widespread use after Launder and Spalding (1972), and adjustment of the turbulence model closure coefficient proposed by Launder and Sharma (1974) (Eqs. 14-17).

$$x_i = \{x_{i1}, x_{i2}, x_{i3}, \dots, x_{im}\} \quad (14)$$

$$\frac{\partial \varepsilon}{\partial t} + U_j \frac{\partial \varepsilon}{\partial x_j} = C_{\varepsilon 1} \frac{\varepsilon}{k} \tau_{ij} \frac{\partial U_i}{\partial x_j} - C_{\varepsilon 2} \frac{\varepsilon^2}{k} + \frac{\partial}{\partial x_j} \left\{ \left(\nu + \frac{\nu_t}{\sigma_\varepsilon} \right) \frac{\partial \varepsilon}{\partial x_j} \right\} \quad (15)$$

$$\nu_t = C_\mu \frac{k^2}{\varepsilon} \quad (16)$$

The closure coefficient for the ($k-\varepsilon$) model:

$$C_{\varepsilon 1} = 1.44, C_{\varepsilon 2} = 1.92, C_\mu = 1.09, \sigma_k = 1.0, \sigma_\varepsilon = 1.3 \quad (17)$$

By using the renormalization group theory Yakhot and Orszag (1986), and Yakhot et al. (1992) developed the RNG $k-\varepsilon$ model. The ($C_{\varepsilon 2}$) closure coefficient was modified by:

$$C_{\varepsilon 2} \equiv \widetilde{C}_{\varepsilon 2} + \frac{C_\mu \lambda^3 \left(1 - \frac{\lambda}{\lambda_0} \right)}{1 + \beta \lambda^3} \quad (18)$$

$$\lambda \equiv \frac{k}{\varepsilon} \sqrt{2 S_{ij} S_{ji}} \quad (19)$$

The closure coefficients for (RNG $k-\varepsilon$) model are:

$$C_{\varepsilon 1} = 1.42, \widetilde{C}_{\varepsilon 2} = 1.68, C_\mu = 0.085, \sigma_k = 0.72 \quad (20)$$

$$\sigma_\varepsilon = 0.72, \beta = 0.012, \lambda_0 = 4.38 \quad (21)$$

METHODS

Experimental data

Experimental data were acquired from Aysegul and Mustafa (2016). Their experiment was conducted under steady-state conditions in a rectangular flume 8 m long, 15 cm wide, and 40 cm in height, at the Hydraulics Laboratory of the Civil Engineering Department of Dokuz Eylul University. The rate of discharge was calculated by a 90° V-notch located at the downstream part of the flume, and the flow depth was collected with a ± 0.1 mm accuracy ultrasonic level sensor (USL). The discharge ranged between 1 258 and 3 177 cm³/s due to pump capacity. Figure 2 illustrates the different parameters used in the experimental study. 24 experiments were conducted for 8 different broad-crested weir configurations (weir width $B = 10$ and 15 cm), and symmetrical side slopes (0.5:1, 1:1, 2:1, and 3:1) for each weir width; 3 discharges (1 258, 2 007, and 3 177 cm³/s) for each physical configuration were tested as shown in Table 1. The measured coefficient of discharge (C_D) based on the experimental results has a wide range, from 0.30 up to 0.60.

The coefficient of discharge defined by Fritz and Hager (1998) did not consider the effect of weir upstream-facing slope, as shown in Eq. 22:

$$C_D = \left[0.43 + 0.06 \sin \left\{ \pi \left(\frac{H_o}{H_o + L_w} - 0.55 \right) \right\} \right] \quad (22)$$

Aysegul and Mustafa (2016) proposed an update to the C_D equation proposed by Fritz and Hager (1998), to take into account the effect of the approach angel (θ), as shown in Eq. 23:

$$C_D = \left[0.43 + 0.06 \sin \left\{ \pi \left(\frac{H_o}{H_o + L_w} - 0.55 \right) \right\} \right]^{\theta^{0.2}} \quad (23)$$

Figure 3 shows a good correlation ($r = 0.87$) between the proposed equation output and the experimentally obtained values for C_D .

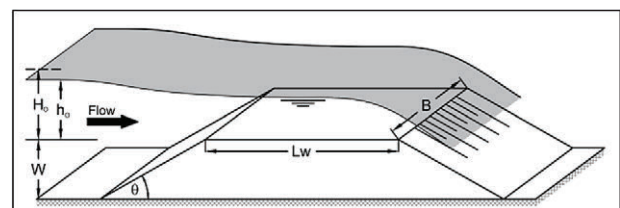


Figure 2. Weir and flow parameters

Table 1. Summary of experiments (Aysegul and Mustafa, 2016)

No.	Experiments			Measured		Fritz and Hager (1998)	
	Slope	L_w cm	Q cm ³ /s	C_D	Average C_D	C_D	Average C_D
1	0.5H:1V	10	3 177	0.409		0.393	
2	0.5H:1V	10	2 007	0.397	0.388	0.385	0.386
3	0.5H:1V	10	1 258	0.359		0.380	
4	1H:1V	10	3 177	0.460		0.391	
5	1H:1V	10	2 007	0.413	0.405	0.384	0.385
6	1H:1V	10	1 258	0.342		0.380	
7	2H:1V	10	3 177	0.434		0.392	
8	2H:1V	10	2 007	0.448	0.450	0.383	0.384
9	2H:1V	10	1 258	0.468		0.377	
10	3H:1V	10	3 177	0.460		0.391	
11	3H:1V	10	2 007	0.533	0.533	0.381	0.382
12	3H:1V	10	1 258	0.605		0.375	
13	0.5H:1V	15	3 177	0.367		0.384	
14	0.5H:1V	15	2 007	0.300	0.336	0.380	0.379
15	0.5H:1V	15	1 258	0.342		0.374	
16	1H:1V	15	3 177	0.441		0.381	
17	1H:1V	15	2 007	0.380	0.382	0.377	0.378
18	1H:1V	15	1 258	0.325		0.375	
19	2H:1V	15	3 177	0.393		0.383	
20	2H:1V	15	2 007	0.413	0.408	0.377	0.377
21	2H:1V	15	1 258	0.420		0.373	
22	3H:1V	15	3 177	0.424		0.382	
23	3H:1V	15	2 007	0.457	0.446	0.376	0.377
24	3H:1V	15	1 258	0.457		0.373	

Numerical simulations

In the present study, both three-dimensional (3-D) simulation using Flow-3D CFD package, and one-dimensional (1-D) simulation using HEC-RAS were conducted for the experimental configurations (geometric dimensions, and discharge values), as described by Aysegul and Mustafa (2016).

1-D simulations

The U.S. Army Corps of Engineers. River Analysis System (HEC-RAS version 5.0.6) is free software used to perform the one-dimensional analysis in this study. HEC-RAS simulation is based on the one-dimensional energy equation, where the friction energy loss is based on Manning's equation. In the case of rapidly varied flow the momentum equation is used (HEC-RAS 5.0.6, 2018).

24 models were simulated for the 8 different broad-crested weir configurations, and 3 discharges were used in the 3-D simulations and were similar to the experimental simulations.

The weir geometry is generated by changing the bed elevation at the weir location as shown in Fig. 4. The cross-sections are generated each 2.5 cm. The Manning coefficient (n) was selected to be 0.01 to represent the plexi-glass material of the flume. Normal depth boundary condition and known water surface boundary condition were selected at the upstream and downstream end, respectively.

3-D simulations

Flow-3D is a CFD package developed by Flow Science Incorporated of Los Alamos. Flow-3D was selected based

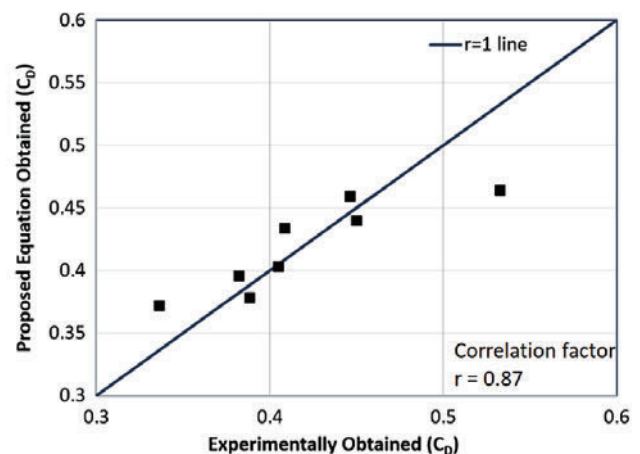


Figure 3. Comparison between experimentally obtained values and proposed equation values of the coefficient of discharge (C_D) (From: Aysegul and Mustafa, 2016)

on its capabilities in tracking free surface flow. The Flow-3D package was tested for several free surface flow problems with satisfactory outcomes (Stefan et al., 2011).

Ninety-six (96) models were simulated: 8 different broad-crested weir configurations (weir width $B = 10$ and 15 cm), and symmetrical side slopes (0.5:1, 1:1, 2:1, and 3:1) for each weir width, 3 discharges (1 258, 2 007, and 3 177 cm³/s) for each physical weir configuration, and 4 turbulence models (one-equation, $k-\epsilon$, RNG $k-\epsilon$, and $k-\omega$).

To avoid conflicts, the run coding is [B (weir width)-Slope ($H: V$)- Q (discharge)], and the legend of the turbulence model is shown on each comparison figure. Ex: 15-3:1-1258 is the run

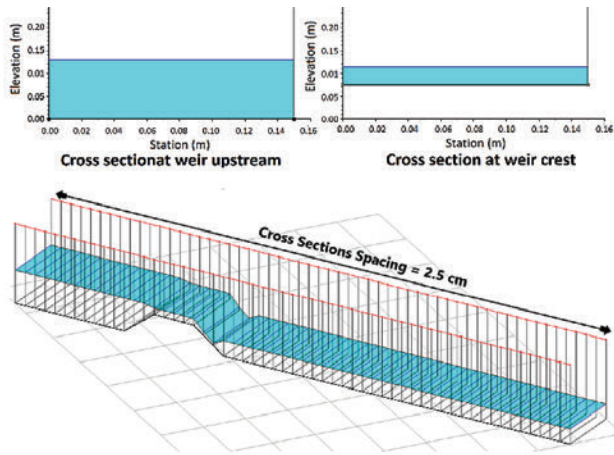


Figure 4. HEC-RAS cross sections

for broad-crested weir top width = 15 cm, with side slopes 3:1, and discharge 1 258 cm³/s.

Geometric presentations and grid type

A stereolithographic (STL) file is prepared for each geometric variation in the broad-crested weir dimensions. Solid objects prepared by CAD programs are converted to STL format by approximating the solid surfaces with triangles, as shown in Fig. 5. Flow-3D uses simple rectangular orthogonal elements in planes and hexahedrals in volumes. This type of grid element eases the mesh generation, reduces the required memory resources and, due to its uniformity, enhances the numerical accuracy. The grid used is a non-adaptive grid and remains fixed throughout the calculations. The boundaries of the fluid domain in the simulation are defined by Fractional Area Volume Obstacle Representation (FAVOR), which computes the fractions of surfaces or volumes occupied by obstacles for each control volume.

Flow governing equations

The flow governing equations are three-dimensional Reynolds Averaged Navier-Stokes equations (RANS). The equations are formulated to comply with FAVOR.

The continuity equation in Cartesian coordinates:

$$V_F \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u A_x) + \frac{\partial}{\partial y}(\rho v A_y) + \frac{\partial}{\partial z}(\rho w A_z) = 0 \quad (24)$$

The momentum equations in Cartesian coordinates:

$$\frac{\partial u}{\partial t} + \frac{1}{V_F} \left\{ u A_x \frac{\partial u}{\partial x} + v A_y \frac{\partial u}{\partial y} + w A_z \frac{\partial u}{\partial z} \right\} = -\frac{1}{\rho} \frac{\partial P}{\partial x} + G_x + f_x \quad (25)$$

$$\frac{\partial v}{\partial t} + \frac{1}{V_F} \left\{ u A_x \frac{\partial v}{\partial x} + v A_y \frac{\partial v}{\partial y} + w A_z \frac{\partial v}{\partial z} \right\} = -\frac{1}{\rho} \frac{\partial P}{\partial y} + G_y + f_y \quad (26)$$

$$\frac{\partial w}{\partial t} + \frac{1}{V_F} \left\{ u A_x \frac{\partial w}{\partial x} + v A_y \frac{\partial w}{\partial y} + w A_z \frac{\partial w}{\partial z} \right\} = -\frac{1}{\rho} \frac{\partial P}{\partial z} + G_z + f_z \quad (27)$$

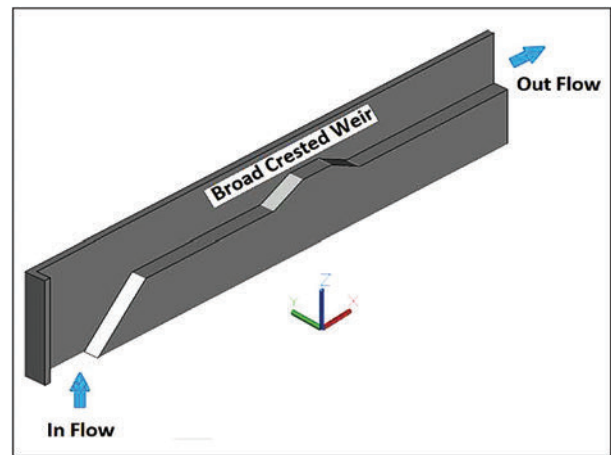


Figure 5. CAD solid geometric file

(VOF) Free surface determination

The flow free surface is a boundary with discontinuity conditions. Hirt and Nichols (1981) developed the VOF algorithm, in which a function (F) is developed to represent each grid cell occupancy with fluid. The F value varies between zero, when the grid cell contains no fluid, and unity when the grid cell is fully occupied with fluid, as shown in Fig. 6. A free surface must be in cells having F values between unity and zero. Since F is a step function, the normal direction to the cutting line represents the free surface inside the grid cell, which is perpendicular to the direction of rapid change in F values.

$$\frac{\partial F}{\partial t} + \frac{1}{V_F} \left\{ \frac{\partial}{\partial x}(F A_x u) + \frac{\partial}{\partial y}(F A_y v) + \frac{\partial}{\partial z}(F A_z w) \right\} = 0 \quad (28)$$

Turbulence models

In the Flow-3D package, the turbulence model equations differs slightly from other formulations to include FAVOR, and represent a slight change in the closure coefficients FLOW-3D (2010).

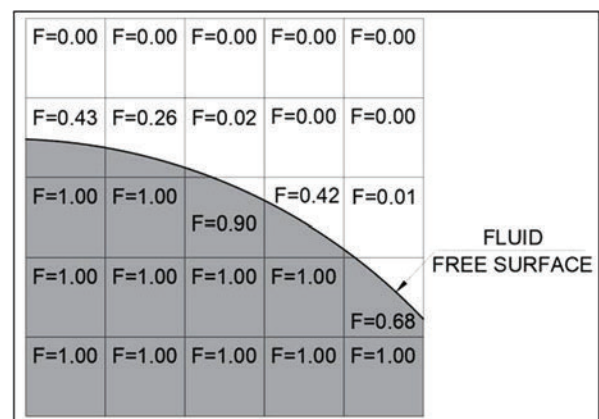


Figure 6. VOF free surface assessment

One-equation model

$$\frac{\partial k}{\partial t} + \frac{1}{V_f} \left\{ u A_x \frac{\partial k}{\partial x} + v A_y \frac{\partial k}{\partial y} + w A_z \frac{\partial k}{\partial z} \right\} = P_t + Diff_k - \varepsilon \quad (29)$$

$$P_t = \frac{\mu}{\rho V_f} \left\{ \begin{aligned} &2A_x \left(\frac{\partial u}{\partial x} \right)^2 + 2A_y \left(\frac{\partial v}{\partial y} \right)^2 + 2A_z \left(\frac{\partial w}{\partial z} \right)^2 \\ &+ \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) \left\{ A_x \frac{\partial v}{\partial x} + A_y \frac{\partial u}{\partial y} \right\} \\ &+ \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right) \left\{ A_z \frac{\partial u}{\partial z} + A_x \frac{\partial w}{\partial x} \right\} \\ &+ \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right) \left\{ A_z \frac{\partial v}{\partial z} + A_y \frac{\partial w}{\partial y} \right\} \end{aligned} \right\} \quad (30)$$

$$Diff_k = \frac{1}{V_f} \left\{ \frac{\partial}{\partial x} \left(v_k A_x \frac{\partial k}{\partial x} \right) + \frac{\partial}{\partial y} \left(v_k A_y \frac{\partial k}{\partial y} \right) + \frac{\partial}{\partial z} \left(v_k A_z \frac{\partial k}{\partial z} \right) \right\} \quad (31)$$

$$\varepsilon = 0.14 \frac{k^{3/2}}{l} \quad (32)$$

k-ε model

$$\frac{\partial \varepsilon}{\partial t} + \frac{1}{V_f} \left\{ u A_x \frac{\partial \varepsilon}{\partial x} + v A_y \frac{\partial \varepsilon}{\partial y} + w A_z \frac{\partial \varepsilon}{\partial z} \right\} = \frac{C1 \varepsilon}{k} P_t + Diff_\varepsilon - C2 \frac{\varepsilon^2}{k} \quad (33)$$

$$Diff_\varepsilon = \frac{1}{V_f} \left\{ \frac{\partial}{\partial x} \left(v_\varepsilon A_x \frac{\partial \varepsilon}{\partial x} \right) + \frac{\partial}{\partial y} \left(v_\varepsilon A_y \frac{\partial \varepsilon}{\partial y} \right) + \frac{\partial}{\partial z} \left(v_\varepsilon A_z \frac{\partial \varepsilon}{\partial z} \right) \right\} \quad (34)$$

Where C1 and C2 are the closure parameters of the equation equal to 1.44 and 1.92, respectively.

RNG k-ε model

The closure factors in RNG *k-ε* have been changed from the normal *k-ε* model. C1 = 1.42, and C2 is calculated from the turbulent kinetic energy and the turbulent production terms.

k-ω model

$$\frac{\partial k}{\partial t} + \frac{1}{V_f} \left\{ u A_x \frac{\partial k}{\partial x} + v A_y \frac{\partial k}{\partial y} + w A_z \frac{\partial k}{\partial z} \right\} = P_t + Diff_k - \beta^* k \omega \quad (35)$$

$$\beta^* = \beta_o^* f_{\beta^*} \quad (36)$$

$$f_{\beta^*} = \begin{cases} 1 & \chi_k \leq 0 \\ \frac{1 + 680 \chi_k^2}{1 + 680 \chi_k^2} & \chi_k > 0 \end{cases} \quad (37)$$

$$\chi_k = \frac{1}{\omega^3} \left(\frac{\partial k}{\partial x} \frac{\partial \omega}{\partial x} + \frac{\partial k}{\partial y} \frac{\partial \omega}{\partial y} + \frac{\partial k}{\partial z} \frac{\partial \omega}{\partial z} \right) \quad (38)$$

$$\frac{\partial \omega}{\partial t} + \frac{1}{V_f} \left\{ u A_x \frac{\partial \omega}{\partial x} + v A_y \frac{\partial \omega}{\partial y} + w A_z \frac{\partial \omega}{\partial z} \right\} = \alpha \frac{\partial \omega}{\partial t} (P_t + Diff_\omega) - \beta \omega^2 \quad (39)$$

$$Diff_\omega = \frac{1}{V_f} \left\{ \frac{\partial}{\partial x} \left(v_\omega A_x \frac{\partial \omega}{\partial x} \right) + \frac{\partial}{\partial y} \left(v_\omega A_y \frac{\partial \omega}{\partial y} \right) + \frac{\partial}{\partial z} \left(v_\omega A_z \frac{\partial \omega}{\partial z} \right) \right\} \quad (40)$$

$$\beta = \frac{9}{125} f_\beta \quad (41)$$

$$f_\beta = \frac{1 + 70 \chi_\omega}{1 + 10 \chi_\omega} \quad (42)$$

$$\chi_\omega = \left| \frac{\Omega_{ij} \Omega_{jk} S_{ki}}{(\beta_o^* \omega)^2} \right| \quad (43)$$

Boundary conditions

To achieve accurate results, boundary conditions should be appropriately defined. Based on the measured experimental values it is clear that the flow over the broad-crested weir is a free flow, and the weir upstream generated head is not affected by the tail water conditions. In order to simulate both the head upstream of the weir, and the generated hydraulic jump at the weir downstream, the outlet boundary condition is defined as a specific pressure boundary condition type as per Flow-3D terminology, where the water level was defined as measured from the experiments. To avoid defining two parameters at the upstream end of the broad-crested weir, (head, and discharge,) the inlet configuration was simulated as the flume inlet as shown in Fig. 7, and the only defined boundary condition is the discharge from the bottom. The top boundary of the calculation domain is defined as a specific pressure boundary condition type with assigned atmospheric pressure. The bottom and sides were defined as walls.

Time step adjustment

The Courant number represents the portion of a grid cell in the direction of flow, that will be travelled by the flow during the simulation time step. Adjustable variable time step was selected to perform the CFD analyses, based on a Courant number stability criterion equal to unity.

$$\delta_t < \min \left(\frac{V_F \delta_{x_i}}{A_x u}, \frac{V_F \delta_{y_j}}{A_y v}, \frac{V_F \delta_{z_k}}{A_z w} \right) \quad (44)$$

Based on the selected criteria the calculation time step varied between 0.0011 and 0.0028 seconds.

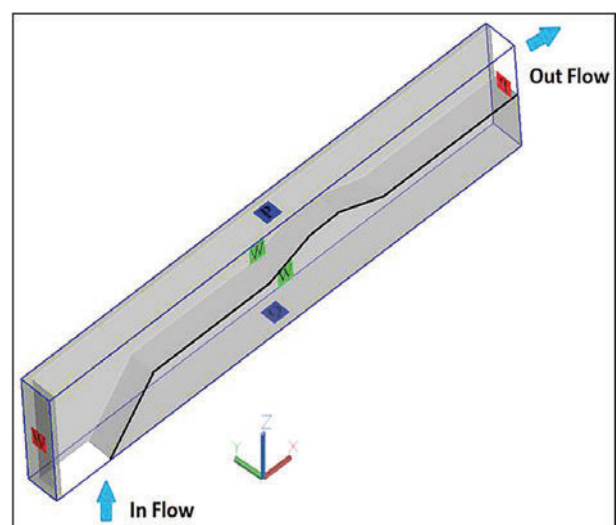


Figure 7. Domain, and boundary conditions

Grid generation and sensitivity analysis

Figure 8 shows the mesh generation which is one of the CFD key elements. In order to assess the adequacy of the selected mesh size for the simulation, a grid sensitivity analysis versus the parameters used in the study should be carried out. In the current study a sensitivity analysis for grid sizes 3, 2, 1, and 0.5 cm, versus 750, 1 258, 2 007, 1 258, 3 177 and 4 000 cm³/s discharges, using *k-ε* turbulence model, was conducted for 3 points along the weir crest, as shown in Figs 9 to 11.

The differences between the water surface profile between mesh sizes 1 cm and 0.5 cm are negligible; accordingly a grid size of 1 cm was selected for the current study.

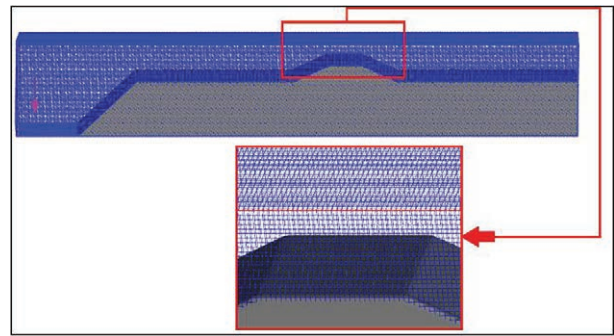


Figure 8. The generated 3-D grid

RESULTS

Simulation output

Figures 12 to 16 show a sample of the output from the 3-D simulation run No. 15-2:1-3177, using the one-equation turbulence model, while Fig. 17 shows the surface water profile obtained from the one-dimensional analysis using HEC-RAS. Figures 18–25 show the comparison between the

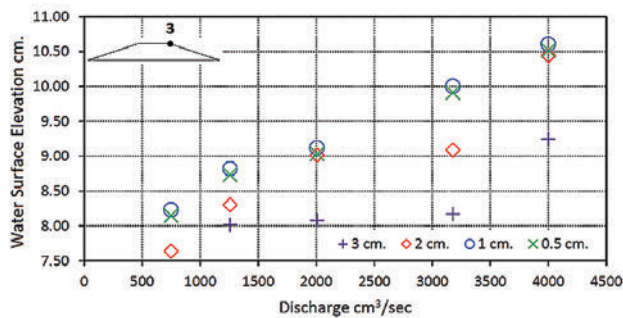


Figure 11. Variation of water surface elevation at Point 3

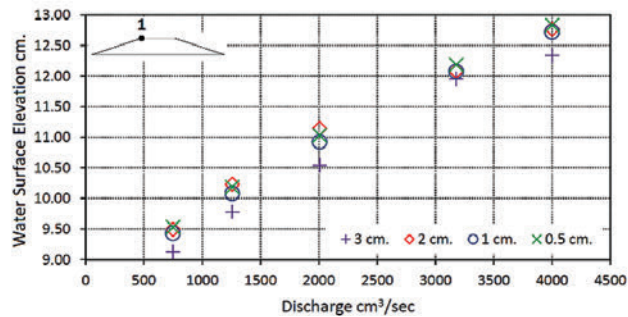


Figure 9. Variation of water surface elevation at Point 1

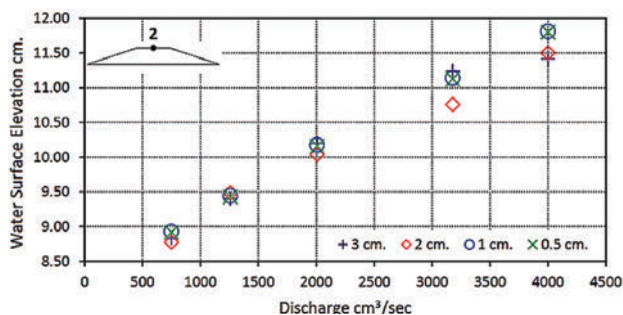


Figure 10. Variation of water surface elevation at Point 2

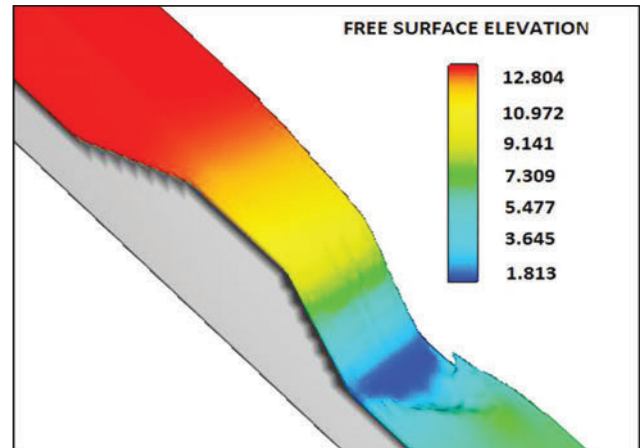


Figure 12. 3-D free surface elevation (cm)

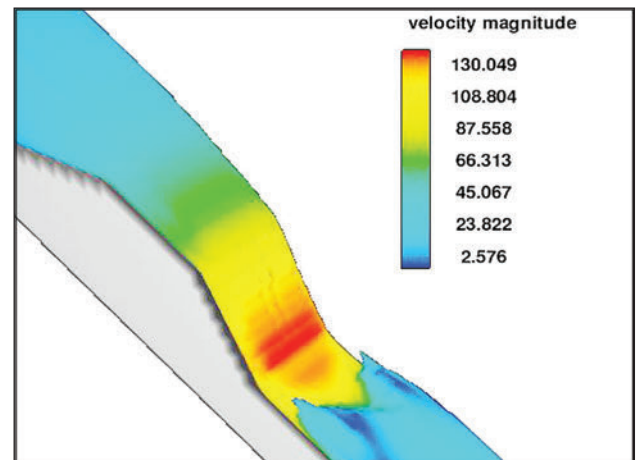


Figure 13. 3-D velocity magnitude (cm/s)

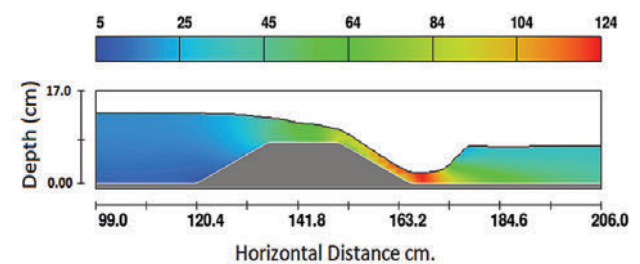


Figure 14. 2-D X-velocity magnitude (cm/s) at flume centreline

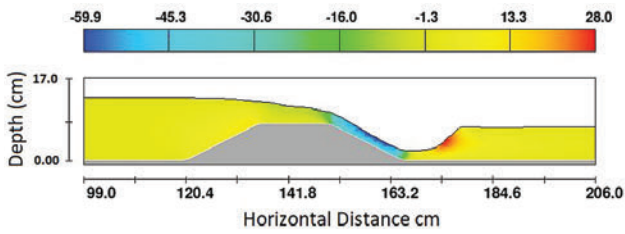


Figure 15. 2-D Z-velocity magnitude (cm/s) at flume centreline

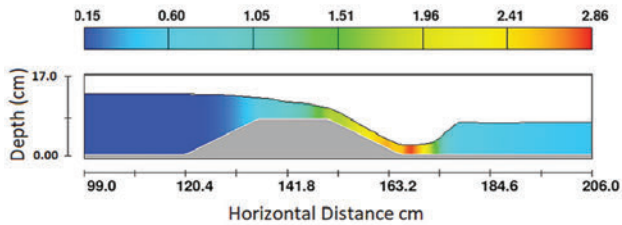


Figure 16. Variation of Froude number at flume centreline

water surface profiles obtained from each 3-D simulation using Flow-3D, and 1-D simulation using HEC-RAS, against the experimental measured values.

Analysis of results

In order to accurately assess the adequacy of each turbulence model to simulate the free surface water level (WL), mean

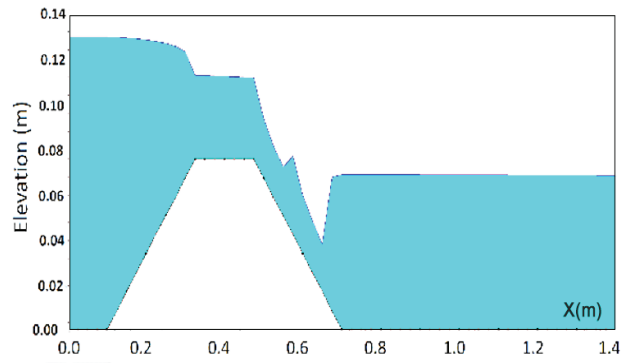


Figure 17. Sample from HEC-RAS output profile

absolute error (MAE), mean relative error percentage (RE), and root mean square error (RMSE) were calculated for each turbulence model according to the following equations:

$$MAE = \frac{1}{N} \sum_{i=1}^N |WL_{experimental} - WL_{numerical}| \quad (45)$$

$$RE = \frac{100}{N} \sum_{i=1}^N \left| \frac{WL_{experimental} - WL_{numerical}}{WL_{numerical}} \right| \quad (46)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (WL_{experimental} - WL_{numerical})^2} \quad (47)$$

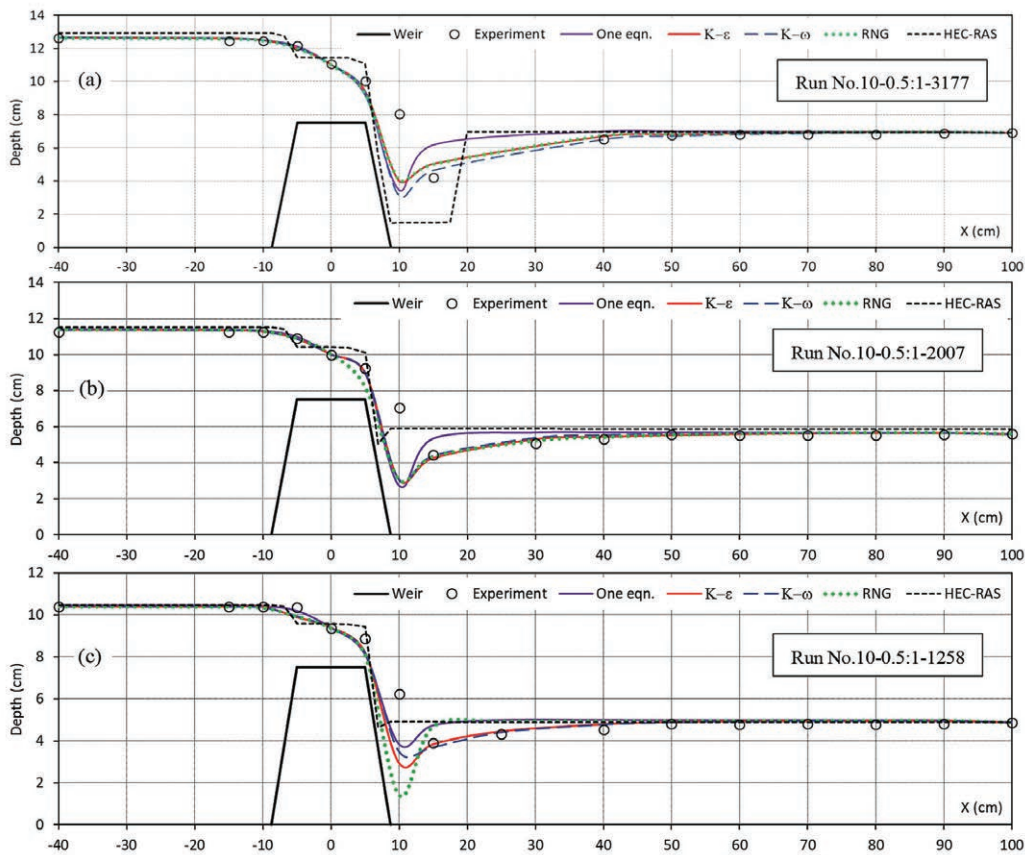


Figure 18. Comparison between 3-D, 1-D simulated water surface, and experimental measured data for broad-crested weir top width = 10 cm, with side slopes 0.5:1, and discharge (a) 3 177, (b) 2 007, and (c) 1 258 cm³/s.

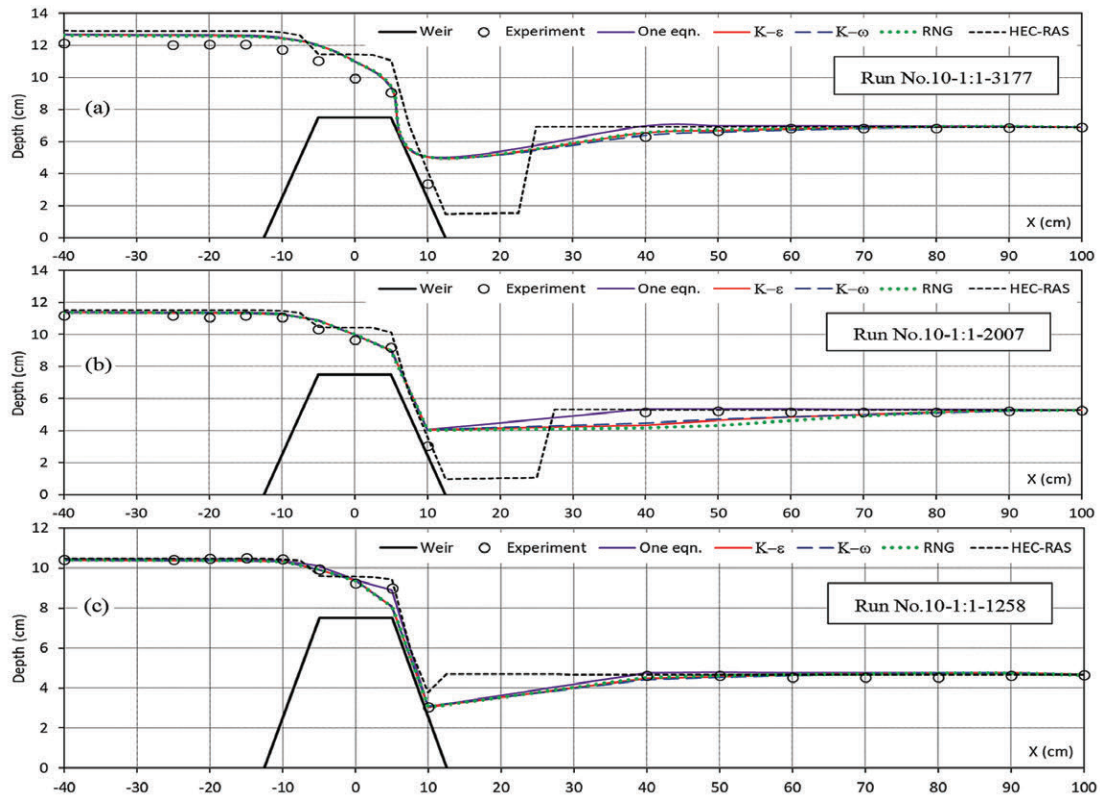


Figure 19. Comparison between 3-D, 1-D simulated water surface, and experimental measured data for broad-crested weir top width = 10 cm, with side slopes 1:1, and discharge (a) 3 177, (b) 2 007, and (c) 1 258 cm³/s.

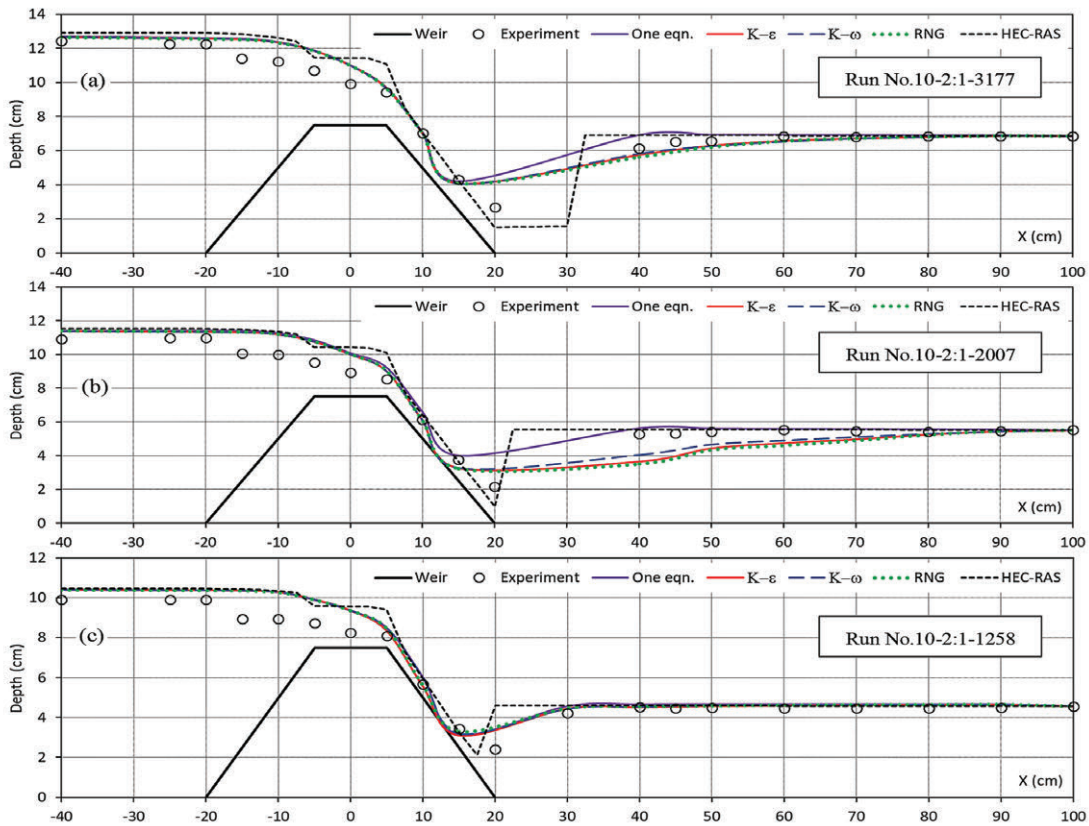


Figure 20. Comparison between 3-D, 1-D simulated water surface, and experimental measured data for broad-crested weir top width = 10 cm, with side slopes 2:1, and discharge (a) 3 177, (b) 2 007, and (c) 1 258 cm³/s.

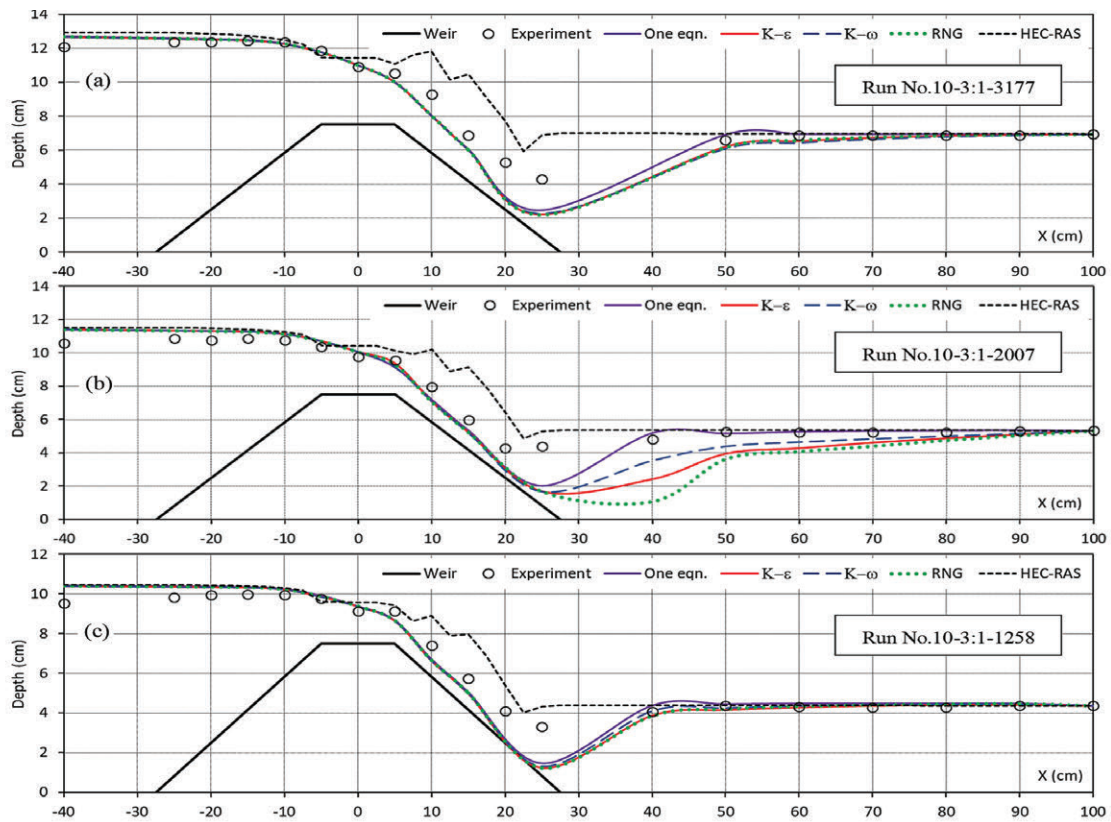


Figure 21. Comparison between 3-D, 1-D simulated water surface, and experimental measured data for broad-crested weir top width = 10 cm, with side slopes 3:1, and discharge (a) 3 177, (b) 2 007, and (c) 1 258 cm³/s.

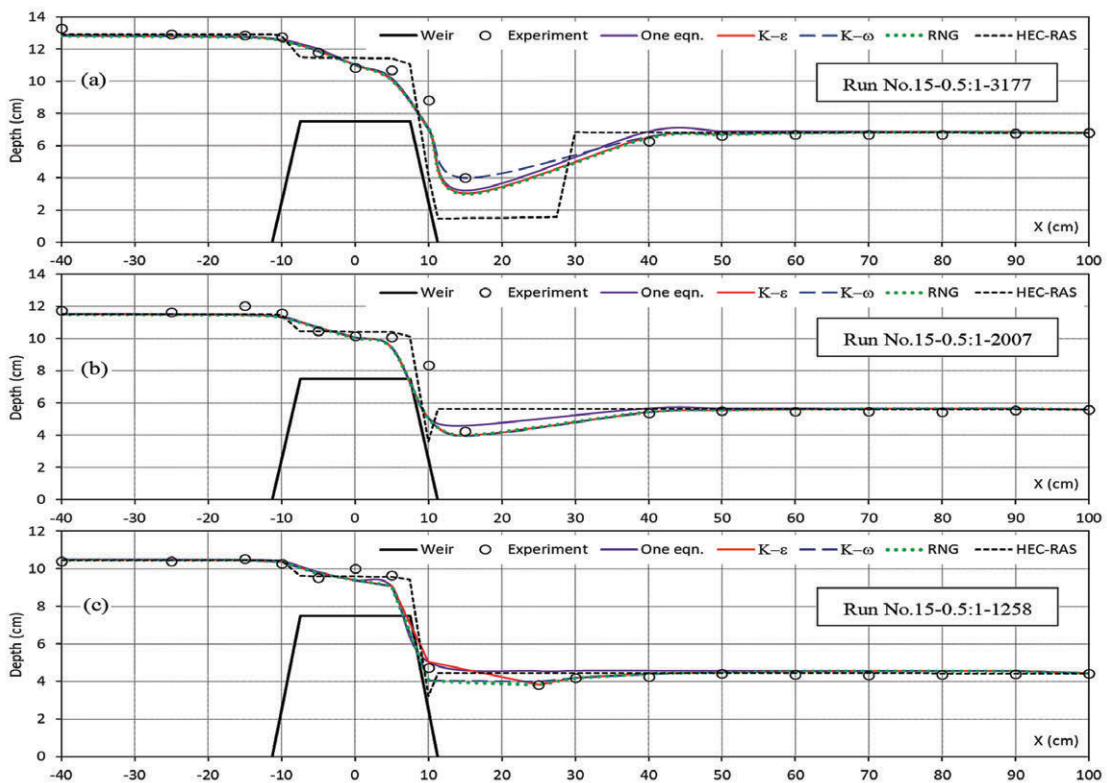


Figure 22. Comparison between 3-D, 1-D simulated water surface, and experimental measured data for broad-crested weir top width = 15 cm, with side slopes 0.5:1, and discharge (a) 3 177, (b) 2 007, and (c) 1 258 cm³/s.

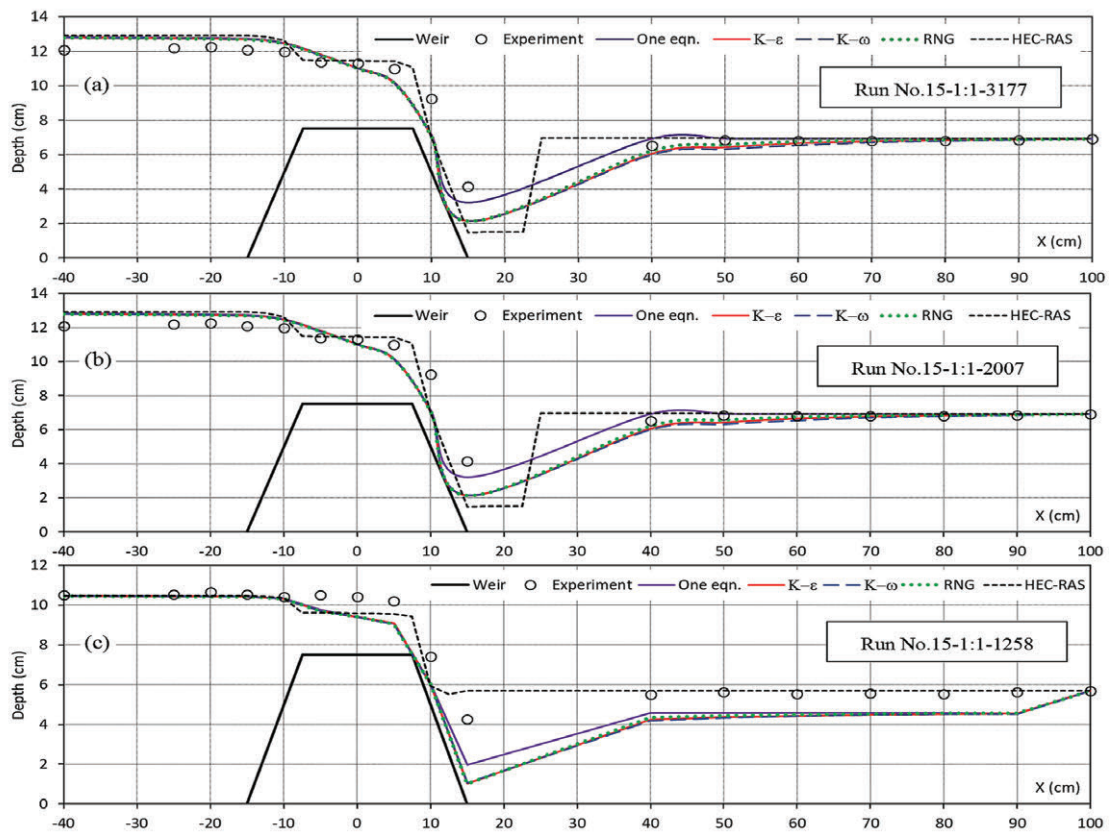


Figure 23. Comparison between 3-D, 1-D simulated water surface, and experimental measured data for broad-crested weir top width = 15 cm, with side slopes 1:1, and discharge (a) 3 177, (b) 2 007, and (c) 1 258 cm³/s.

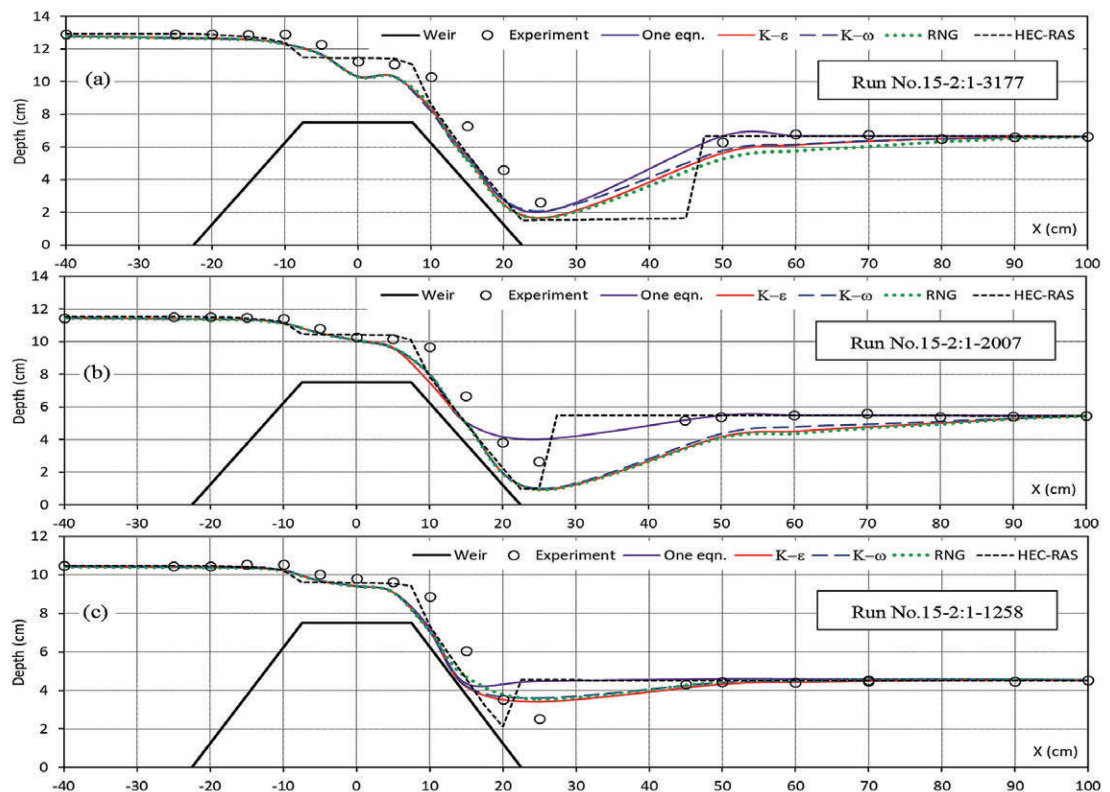


Figure 24. Comparison between 3-D, 1-D simulated water surface, and experimental measured data for broad-crested weir top width = 15 cm, with side slopes 2:1, and discharge (a) 3 177, (b) 2 007, and (c) 1 258 cm³/s.

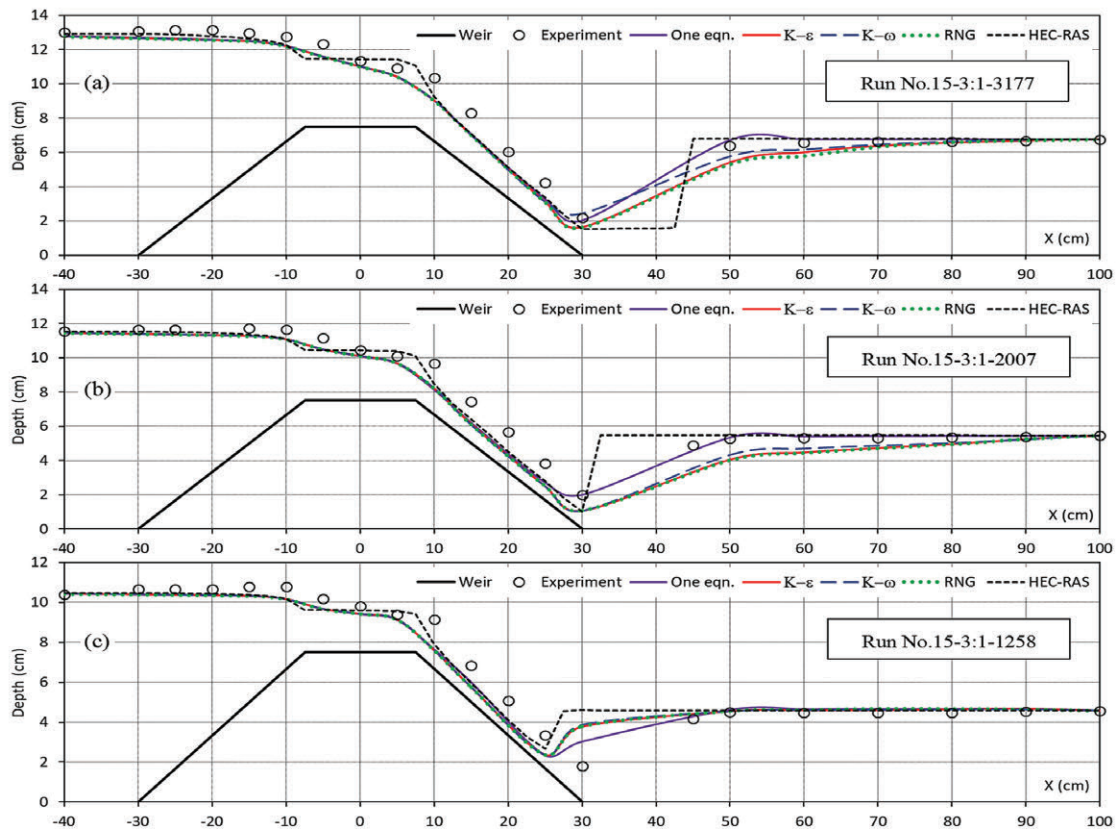


Figure 25. Comparison between 3-D, 1-D simulated water surface, and experimental measured data for broad-crested weir top width = 15 cm, with side slopes 3:1, and discharge (a) 3 177, (b) 2 007, and (c) 1 258 cm³/s.

The error assessment was conducted for two zones of flow; the first is the full flow spectrum as shown in Table 2 from upstream to downstream of the broad-crested weir, and the second for the upstream part of the weir starting from the crest centerline as shown in Table 3. Figure 26 illustrates the assessment zones.

Table 2. Relative error comparison between turbulence 3-D models, and the HEC-RAS 1-D MODEL for the whole flow spectrum (upstream to downstream)

Turbulence model	RE	MAE	RMSE
$k-\epsilon$	7.75	0.49	0.76
$k-\omega$	7.37	0.47	0.74
RNG $k-\epsilon$	8.18	0.51	0.79
One-equation	6.94	0.45	0.67
HEC-RAS	7.75	0.48	0.78

Table 3. Relative error comparison between turbulence 3-D models, and the HEC-RAS 1-D MODEL for the upstream part of flow

Turbulence model	RE	MAE	RMSE
$k-\epsilon$	3.88	0.41	0.58
$k-\omega$	3.87	0.41	0.58
RNG $k-\epsilon$	3.87	0.41	0.58
One-equation	3.93	0.41	0.59
HEC-RAS	3.90	0.42	0.58

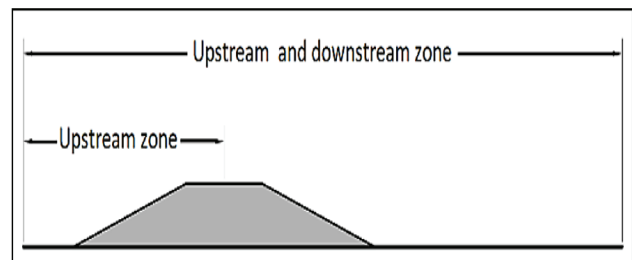


Figure 26. Upstream and downstream zones

CONCLUSION

In the present study, several floodway broad-crested weirs were numerically simulated; (3-D simulation using Flow-3D CFD package, and 1-D using HEC-RAS). The Volume of Fluid (VOF) algorithm was used to define the free water surface.

Assessment of the simulation result accuracy was based on verification against the experimental results obtained for the water surface profile by Aysegul and Mustafa (2016). The relative error approach (MAE, RE, and RMSE) was utilized to check the accuracy of the simulated results.

The error assessment was conducted for two zones of flow; the first was the full flow spectrum from upstream to downstream of the broad-crested weir, and the second for the upstream part of the weir starting from the crest centerline.

The one-equation turbulence model with mixing length equal to 7% of the smallest domain dimension has the minimum error value in simulating the full spectrum free surface flow above a broad-crested weir.

For the upstream part of flow starting from the broad-crested weir centreline, no significant difference was found in accuracy between all turbulence models and the one-dimensional analysis, due to the low turbulence intensity in this section.

For engineering design purposes while the water level is the main concern at the location of the floodway the one-dimensional analysis has sufficient accuracy to define the maximum water elevation.

ACKNOWLEDGEMENTS

The authors would like to thank Eng. Kaleemullah Mehrabi for help in editing the text of this manuscript.

NOTATION

A_x, A_y, A_z : area fraction open to flow in (x, y, and z) directions
Diff: diffusion

f_x, f_y, f_z : viscous acceleration in (x, y, and z) directions

G_x, G_y, G_z : acceleration in (x, y, and z) directions

i, j, k : unit vector in x, y, and z directions

k : kinetic energy of turbulent fluctuation per unit mass

l : turbulence length scale; a characteristic of eddy size

P : pressure

P_t : turbulence kinetic energy production

t : time

U_i : mean velocity in tensor notation

U, V, W : instantaneous velocity components in x, y, and z directions

U', V', W' : fluctuating velocity components in x, y, and z directions

V_f : opened fraction volume to flow

ws_x, ws_y, ws_z : wall shear stress in x, y, and z directions

x_i : position tensor in tensor notation

x, y, z : rectangular Cartesian coordinates

ϵ : dissipation per unit mass

ω : specific dissipation rate

μ : dynamic viscosity

S_{ij} : mean strain rate tensor

δ_{ij} : Kronecker delta tensor

δ_t : time step

Ω_{ij} : mean rotation tensor

ν : kinematic viscosity

ν_t : kinematic eddy viscosity

ν_k : kinetic energy diffusion coefficient

ν_ϵ : ϵ diffusion coefficient

ν_ω : ω diffusion coefficient

τ_{ij} : specific Reynolds stress tensor

λ : Taylor microscale in RNG K- ϵ model

ρ : fluid density

REFERENCES

- AHMED MH and MOHAMED HE (2011) Experimental and numerical investigation of flow through free double baffled gates. *Water SA* 37 (2) 245–254. <https://doi.org/10.4314/wsa.v37i2.65871>
- AYSEGUL OA and MUSTAFA D (2016) Experimental investigation of the approach angle effect on the discharge efficiency for broad-crested weirs. *Sci. Technol.* 17 (2) 279–286. <https://doi.org/10.18038/btda.48930>
- CHANSON H (2004) *The Hydraulics of Open Channel Flows: An Introduction* (2nd edn). Butterworth-Heinemann, Oxford, UK. 650 pp.
- CHOU PY (1945) On the velocity correlations and the solution of the equations of turbulent fluctuation. *Q. Appl. Math.* 3 (1) 38–54.

- DAVYDOV BI (1961) On the statistical dynamics of an incompressible fluid. *Doklady Akademiyi Nauk SSSR* 136 (1) 47–50.
- EMMONS HW (1954) Shear Flow Turbulence. In: *Proceedings of the 2nd U.S. Congress of Applied Mechanics-American Society of Mechanical Engineers*, 14–18 June 1954, Michigan.
- FARD MH, and BOYAGHCHI FA (2007) Studies on the influence of various blade outlet angles in a centrifugal pump when handling viscous fluids. *Am. J. Appl. Sci.* 4 (9) 718–724. <https://doi.org/10.3844/ajassp.2007.718.724>
- FRITZ HM and HAGER WH (1998) Hydraulics of embankment weirs. *Hydraul. Eng.* 124 (9) 963–971. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1998\)124:9\(963\)](https://doi.org/10.1061/(ASCE)0733-9429(1998)124:9(963))
- FLOW-3D (2016), User Manual V11.2. Flow Science INK, Santa Fe.
- GLUSHKO G (1965) Turbulent boundary layer on a flat plate in an incompressible fluid. *Izvestia Acad. Nauk. SSSR Mekh.* 4 (1) 13–23.
- HARGREAVES DM, MORVAN HP and WRIGHT NG (2007) Validation of the volume of fluid method for free surface calculation: The broad-crested weir. *Eng. Appl. Comput. Fluid Mech.* 1 (2) 136–146. <https://doi.org/10.1080/19942060.2007.11015188>
- HARLOW FH and NAKAYAMA PI (1968) Transport of Turbulence Energy Decay Rate. Los Alamos Scientific Laboratory of the University of California Report LA-3854. 11 pp. <https://www.osti.gov/servlets/purl/4556905>
- HEC-RAS 5.0.6 (2018) Hydraulic reference manual. URL: <https://www.hec.usace.army.mil/software/hec-ras/documentation.aspx> (Accessed December 2018).
- HELLSTEN A (2005) New advanced k- ω turbulence model for high-lift aerodynamics. *AIAA* 43 (9) 1857–1869. <https://doi.org/10.2514/1.13754>
- HENDERSON FM (1966) *Open Channel Flow*. Macmillan, New York. 522 pp.
- HIRT CW and NICHOLS BD (1981) Volume of Fluid (VOF) Method for the dynamics of free boundaries. *Comput. Phys.* 39 (1) 201–225.
- HOSEINI SH, JAHROMI SH, and VAHID MS (2013) Determination of discharge coefficient of rectangular broad-crested side weir in trapezoidal channel by CFD. *Int. J. Hydraul. Eng.* 2 (4) 64–70.
- HOSSEINI A and SAYED HH (2013) Experimental and 3-D numerical simulation of flow over a rectangular broad-crested weir. *Int. J. Eng. Adv. Technol.* 2 (6) 214–219.
- JIANG L, DIAO M, SUN H and REN Y (2018) Numerical modelling of flow over a rectangular broad-crested weir with a sloped upstream face. *Water* 10 (11) 1–12. <https://doi.org/10.3390/w10111663>
- JOONGCHEOL P and NAM JI (2015) Numerical modelling of free surface flow over a broad-crested rectangular weir. *Korea Water Resour. Assoc.* 48 (4) 281–290.
- KASSAF SI, ATTIYAH AN and YOUSIFY HA (2016) Experimental investigation of compound side weir with modelling using computational fluid dynamic. *Int. J. Energ. Environ.* 7 (2) 167–178.
- KIRKGOV MS, AKOZ MS and ONER AP (2008) Experimental investigation of compound side weir with modelling using computational fluid dynamic. *Can. J. Civ. Eng.* 35 (9) 975–986. <https://doi.org/10.1139/L08-036>
- KOK JC (2000) Resolving the dependence on freestream values for the k- ω turbulence model. *AIAA* 38 (7) 1292–1295. <https://doi.org/10.2514/2.1101>
- KOLMOGOROV AN (1942) Equations of turbulent motion of an incompressible fluid. *Izvestia Acad. Sci. USSR.* 6 (1–2) 56–58.
- LAUNDER BE and SHARMA BI (1974) Application of the energy dissipation model of turbulence to the calculation of flow near a spinning disc. *Lett. Heat Mass Transfer* 1 (1) 131–138. [https://doi.org/10.1016/0094-4548\(74\)90150-7](https://doi.org/10.1016/0094-4548(74)90150-7)
- LAUNDER BE and SPALDING DB (1972) *Mathematical Models of Turbulence*. Academic Press, London. 169 pp.
- MAGHSOODI R, ROOZGAR MS, SARKARDEH H and AZAMATULLA HM (2012) 3D-simulation of flow over submerged weirs. *Int. J. Model. Simul.* 32 (4) 237–243. <https://doi.org/10.2316/Journal.205.2012.4.205-5656>
- PENG SH, DAVIDSON L and HOLMBERG S (1997) A modified low

- Reynolds number k-w model for recirculating flows. *Fluids Eng.* **119** (4) 867–875. <https://doi.org/10.1115/1.2819510>
- PRANDTL L (1925) Uber die ausgebildete Turbulenz (About the trained turbulence). *ZAMM Appl. Math. Mech.* **5** (2) 136–139.
- PRANDTL L (1945) Uber ein neues Formelsystem fur die ausgebildete Turbulenz (About a new formula system for the deducted turbulence). In: *Nachrichten Akademie der Wissenschaften in Gottingen Mathematisch-Physikalische Klasse*. Springer-Verlag, Berlin.
- SARGISON JE and PERCY A (2009) Hydraulics of broad-crested weirs with varying side slopes. *Irrig. Drain. Eng.* **135** (1) 115–118. [https://doi.org/10.1061/\(ASCE\)0733-9437\(2009\)135:1\(115\)](https://doi.org/10.1061/(ASCE)0733-9437(2009)135:1(115))
- SARKER MA and RHODES DG (2004) Calculation of free-surface profile over a rectangular broad-crested weir. *Flow Meas. Instrum.* **15** (4) 215–219. <https://doi.org/10.1016/j.flowmeasinst.2004.02.003>
- SHAKER AJ and SARHAN AS (2017) Performance of flow over a weir with sloped upstream face. *ZNCO J. Pure Appl. Sci.* **29** (3) 43–54. <https://doi.org/10.21271/ZJPAS.29.3.6>
- SHAYMAA AM, HUDA MM and THAMEEN NN (2017a) Experimental and numerical simulation of flow over broad-crested weir and stepped weir using different turbulence models. *Eng. Sustainable Dev.* **12** (2) 28–45.
- SHAYMAA AM, HUDA MM, RASUL MK, THAMEEN NN and NADHIR AA (2017b) Flow over broad-crested weirs: comparison of 2d and 3d models. *Civ. Eng. Archit.* **11** (1) 769–779. <https://doi.org/10.17265/1934-7359/2017.08.005>
- SHAYMAA AM, SADIQ AS and HUDA MM (2015) determination of discharge coefficient of rectangular broad-crested weir by CFD. In: *Proceedings of The Second International Conference on Buildings, Construction and Environmental Engineering*, 17–18 October 2015, Beirut, Lebanon.
- SPEZIALE CG, ABID R and ANDERSON EC (1992) Critical evaluation of two-equation models for near-wall turbulence. *AIAA* **3** (2) 324–331.
- STEFAN H, NILS RB and ROBERT F (2011) numerical modelling of flow over trapezoidal broad-crested weir. *Eng. Appl. Comput. Fluid Mech.* **5** (3) 397–405. <https://doi.org/10.1080/19942060.2011.11015381>
- TANASE NO, BROBOANA D and BALAN B (2015) Free surface flow over the broad-crested weir. In: *Proceedings of the Ninth International Symposium on Advanced Topics in Electrical Engineering*, 7–9 May 2015, Bucharest, Romania.
- WILCOX DC (1988) Reassessment of the scale determining equation for advanced turbulence models. *AIAA* **26** (11) 1299–1310. <https://doi.org/10.2514/3.10041>
- WILCOX DC (1998) *Turbulence Modelling for CFD*. (2nd edn). DCW Industries, Inc., La Canada, California. 540 pp.
- WILCOX DC (2006) *Turbulence Modelling for CFD*. (3rd edn). DCW Industries, Inc., La Canada, California. 522 pp.
- YAKHOT V and ORSZAG SA (1986) Renormalization group analysis of turbulence: 1. Basic theory. *Sci. Comput.* **1** (1) 3–51.
- YAKHOT V, ORSZAG SA, THANGAM S, GATSKI TB and SPEZIALE CG (1992) Development of turbulence models for shear flows by a double expansion technique. *Phys. Fluids A: Fluid Dyn.* **4** (7) 1510–1520. <https://doi.org/10.1063/1.858424>
- YAZDI J, SARKARDEH H, AZAMATULLA HM and GHANI AA (2010) 3D simulation of flow around a single spur dike with free surface flow. *Int. J. River Basin Manag.* **8** (1) 55–62. <https://doi.org/10.1080/15715121003715107>

Development of a design and implementation process for the integration of hydrokinetic devices into existing infrastructure in South Africa

CM Niebuhr^{1*}, M Van Dijk¹ and JN Bhagwan²

¹Department of Civil Engineering, University of Pretoria, Pretoria, 0001, South Africa

²Water Research Commission, Private Bag X03, Gezina 0031, South Africa

ABSTRACT

In South Africa there is currently no notable use of modern small-scale hydrokinetic (HK) energy systems, mainly due to formerly low-cost coal-powered electricity. This renewable energy option makes use of the kinetic energy from flowing water, rather than potential energy, which is more often used in conventional hydropower. Updated refined versions of this technology are now being investigated and manufactured due to the global drive towards reducing carbon emissions and increasing energy efficiency. These modular units allow for installation of HK turbines into existing water infrastructure with very little civil works. The study's objective was to develop a simplified design and implementation process for HK devices within the South African legislative and regulatory environment. Approximately 66% of South Africa's water supply is used by the agricultural sector with more than 6 500 km of canal systems running through many areas which could benefit from alternative energy sources. The recent electricity crisis in the country allowed for problem resolution through funding opportunities and thereby an introduction of an innovative and sustainable technology to provide renewable electricity where otherwise not feasible. A pilot HK project was implemented in an applicable section on the Boegoeberg irrigation canal in the Northern Cape Province and tested for optimum functionality and correct application. This process allowed evolution of a development process for the implementation of HK devices in existing water infrastructure.

Keywords: hydrokinetic turbine, water infrastructure, renewable energy, hydropower, small-scale

INTRODUCTION

Fossil fuels are currently the primary resource for electricity generation in South Africa (SA). However, the availability of renewable energy resources is abundant and holds the potential to assist in solving issues such as reducing greenhouse gas emissions and having larger energy security by diversifying the supply (Kusakana and Vermaak, 2013).

In 2012, Kusakana and Vermaak found approximately 6 000–8 000 potential sites for traditional hydropower installation in SA. The Department of Minerals and Energy (DME) of the Government of SA revealed the country to have a considerable potential for small- and large-scale hydropower generation; however, only a handful of small hydropower developments have been undertaken in the past 30 years (DME, 2003; Koko and Kusakana, 2014). Currently about 3 700 MW of installed hydropower exists in SA although there remains a great lack of knowledge on hydropower and especially hydrokinetic (HK) systems.

Prior to the case study referred to in this paper there was no existing modern HK installation in SA, although the country was previously reported to have a great potential thereof (Kusakana and Vermaak, 2013). SA is a water-scarce country resulting in highly protected water infrastructure. A challenge for implementation of HK systems lies in the ownership of the infrastructure (canals etc.) onto which such systems are installed versus the mandate of the 'owner' in having such systems installed. A general perception exists that hydropower is not fit for installations in SA due to the lack of availability of components, as well as the drawn out regulatory and

legislative procedures. It is believed that all these problems have arisen due to lack of knowledge of hydropower systems (due to scarcity of existing systems) and therefore the inability to predict any environmental/social impacts thereof.

Hydropower is also misunderstood and, in many cases, believed to consume or pollute water resources. In the case of HK systems, these can be installed in line with existing infrastructure (or designed to be incorporated with new infrastructure in that instance) and do not require the water to be diverted (as required by most conventional hydropower schemes). Furthermore, the impact of the installation of a HK system into existing infrastructure is novel in SA. The purpose of this paper is to clarify such issues and give a broader understanding of HK installations and the possibilities thereof in SA.

In a report on the assessment of HK turbines in open channel applications prepared by the United States Department of Energy it is stated that 'Hydrokinetic energy from flowing water in open channels has the potential to support local electricity needs with lower regulatory or capital investment rather than impounding water with more conventional means (Gunawan et al., 2017) indicating not only local but international relevance to such studies.

Hydrokinetic applicability to South African canal systems

The demand for energy in SA is increasing with a greatly fossil fuel dependent economy resulting in an unsustainable future. According to environmental statistics SA is among the top 20 countries with the highest level of carbon dioxide emissions, also being the highest emitter of greenhouse gases in Africa (MEAAI, 2011). In 2010 the DME released an Integrated

*Corresponding author, email: Chantel.niebuhr@up.ac.za

Received 19 March 2018; accepted in revised form 19 June 2019

Resource Plan (IRP) to increase renewable energy sources to 17.8 GW by 2030 (Koko and Kusakana, 2014). This allows a large opportunity for the development of renewable energy, such as biomass, wind, solar and hydropower. Details of the DME targets and evolution of the electrification goals in SA are summarized in Fig. 1.

Hydropower is a trusted technology with significant potential in SA (White, 2011). It is a form of clean renewable energy making use of the available flow or head in water without consuming or polluting the water itself. Often these hydropower systems have long lifespans and high efficiencies with low operating costs (van Vuuren et al., 2011).

Kusakana et al. (2013) investigated the possibility of HK hydropower development for rural and isolated loads in South Africa. The case studies proved that where adequate water resources are available in South Africa, HK power generation could be the best, most cost-effective supply option in relation to wind, photovoltaic and diesel generators (Kusakana and Vermaak, 2013).

A large network of canal systems exist in South Africa, the Department of Water and Sanitation (DWS) asset management database revealed a total of 47 schemes with a network of more than 6 500 km of canals, as shown in Fig. 2. Additional to the HK potential in straight uniform sections on these canal systems, 21 286 structures such as tunnels, syphons, weirs, control gates, chutes and drops exist, which could also hold unexploited HK (and conventional hydropower) potential.

HK energy extraction is a simple process; therefore the cost of energy extraction is low. Due to its modular nature the initial installation cost and deployment time is short as it does not require construction of significant infrastructure. It also allows an easily scalable energy output but limits any decrease in capital cost per kW output (Gaden and Bibeau, 2006). The cost of generation alone, not including the infrastructural cost

itself, may be lower than traditional hydropower forms and more economical when compared to solar power (Kusakana and Vermaak, 2013).

Development process

The development process of a HK installation varies throughout the world, as each country's specific laws and regulatory processes have to be followed. As HK implementation is a new concept in SA, the development process is not yet streamlined. An outline of a suggested development process for HK integration into existing water infrastructure was compiled through results from global case studies and the pilot study completed in SA, and is depicted schematically in Fig. 3. Some of the processes are described below by means of reference to the case study.

Step 1: Site-selection and pre-feasibility study

The final HK installation site should be carefully selected, as for the case study described later in this paper. This may be a difficult task when optimal velocities are not available. HK turbines usually require a velocity of around 2–3.5 m/s (based on available turbine operational ranges). Sites at existing narrowed sections (such as flume sections or syphon outlets) are usually more likely to have these velocities readily available; examples of these sites are shown in Table 1. When selecting sites such as these, several issues, such as high damming or interruption of the infrastructure operation (before installation) may occur; therefore, selection of a site with a higher velocity along a uniform section, where only a small fraction of the cross section is utilised for the HK installation, remains the best installation option. The site selection procedure shown in Fig. 4 may be used as a guideline to an optimal site selection process.

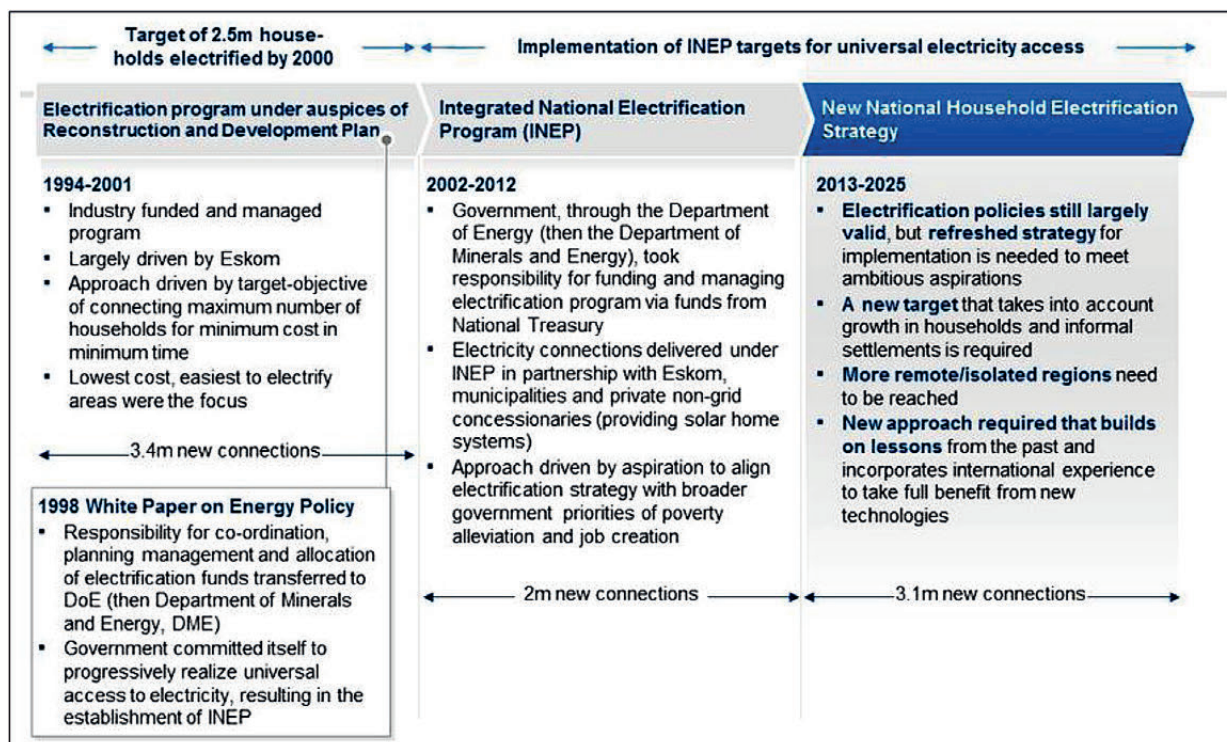


Figure 1. Evolution of electrification in South Africa (IFC, 2013)

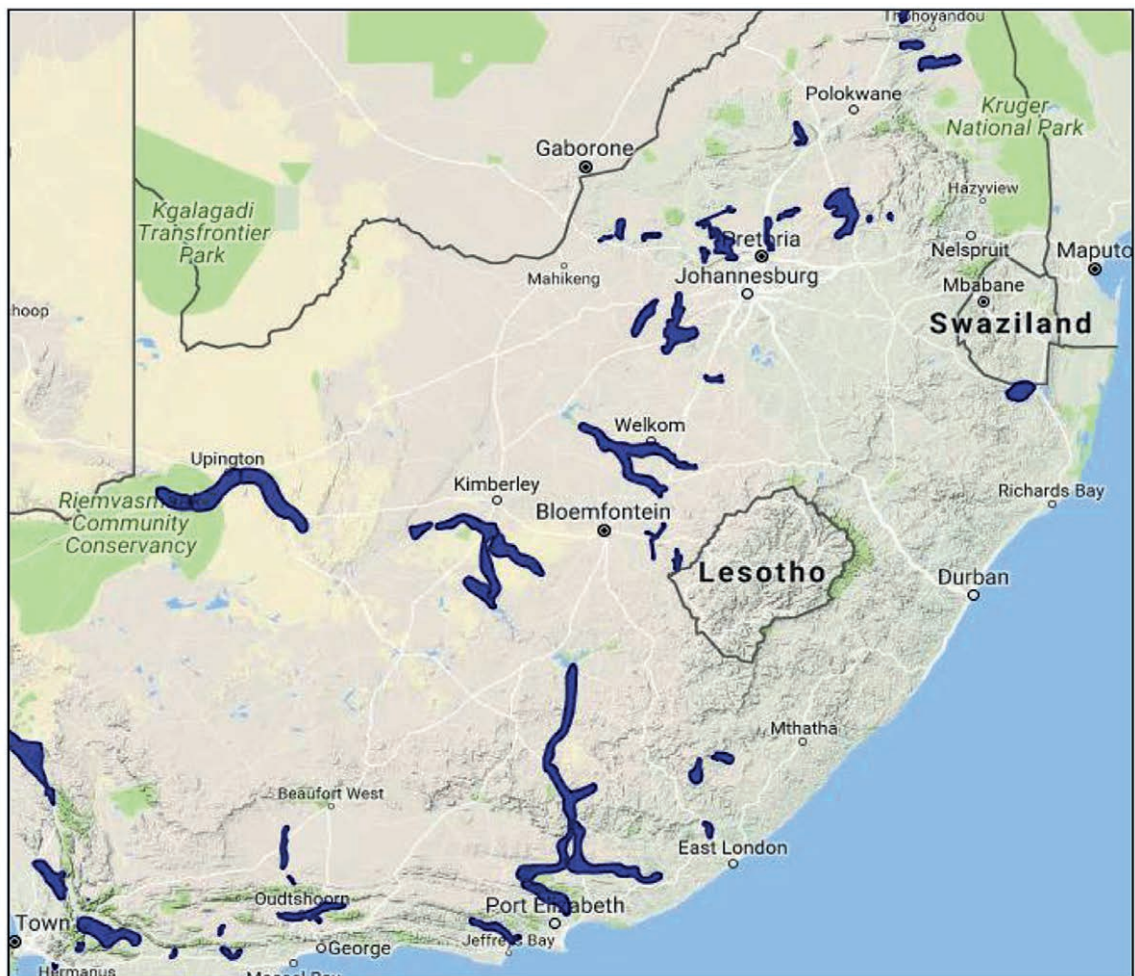





Figure 2. Canal schemes in South Africa (DWS, 2015)



Figure 3. Canal HK system development process

Table 1. Sections with higher velocities (DWS, 2016)

Potential site	Picture
<u>Steep sections:</u> Examples include canal outlets to balancing dams, or sections of higher gradients	
<u>Flume sections:</u> Where a narrowing in the canal occurs, higher velocities may be present	
<u>Siphon/pipe exits:</u> Where siphons are used between canal sections the narrowed section at the exits may be considered as potential sites	

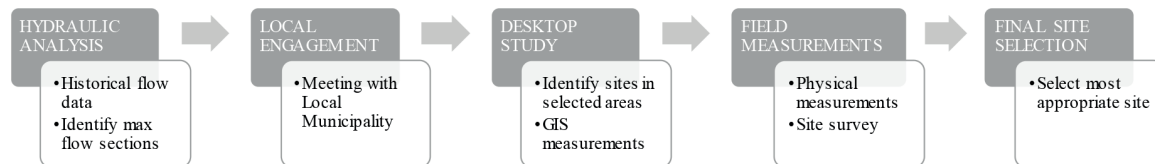


Figure 4. Site selection procedure

In conjunction with site selection, a pre-feasibility study of shortlisted potential sites should be undertaken; this includes:

- Accurate velocity, flow and level measurements
- Identification of electricity use and transmission length
- Identification of connection type and point (grid integration or stand-alone scheme)
- Conceptual design
- Preliminary costing of components for economic evaluation

The pre-feasibility study should be carefully completed, and all costs should be calculated and weighed against the output before final design and implementation of the HK system. The main cost considerations are shown in Fig. 5.

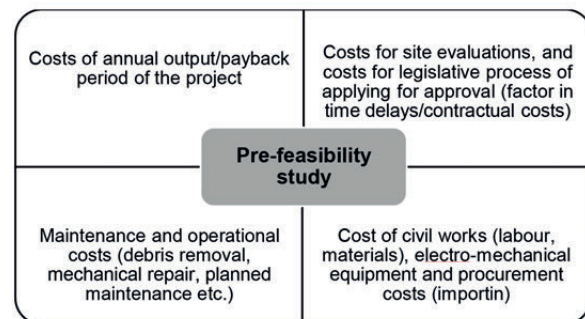


Figure 5. Pre-feasibility study cost considerations

Step 2: Turbine selection

The HK turbine type should be carefully considered before selection. As these devices are not currently readily available in South Africa, high import costs drive up the capital cost. Modular units may be imported, or units may be designed and manufactured locally (where the project owner has a broad knowledge of HK workings). The typical ranges of commonly available turbine units can be seen in Fig. 6.

To suit the design power output the rotor diameters may be adapted. As an indication of the changes experienced in the power curve for a range of rotor diameters, Fig. 7 was included. The exponential behaviour of the curve is emphasized at larger rotor diameters, allowing significantly higher power output when using larger turbines.

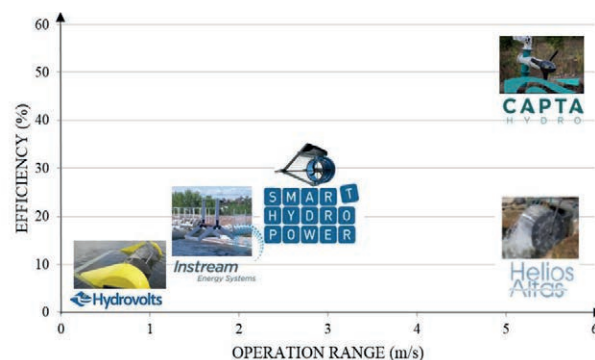


Figure 6. HK turbines (CaptaHydro, 2017)

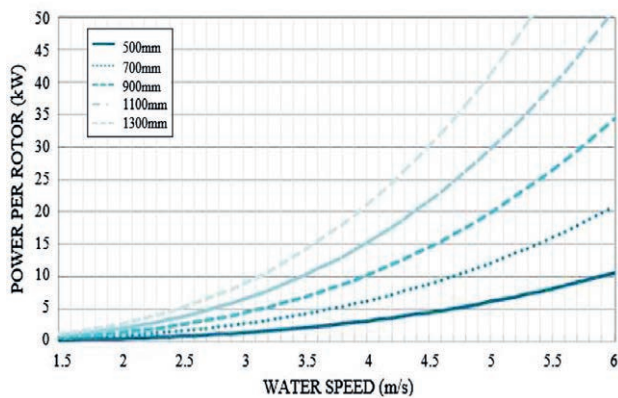


Figure 7. Axial flow rotor diameter comparisons (CaptaHydro, 2017)

Step 3: Grid integration

The control systems for the HK device are designed according to the specified electricity use, which falls within the selected grid integration method:

- Islanded stand-alone method:
 - Control systems including inverters, rectifiers and dump-loads are used to regulate the system
 - This will require the development of a mini grid for the distribution of the power produced
- Grid-connected system:
 - Stand-alone system requiring a grid connection to synchronise frequency
 - Usage of existing transmission lines to feed electricity to the end user (may require additional approvals)

When the grid integration method has been selected the following should be designed or selected as part of the modular turbine unit as supplied by the manufacturer:

- Electric control boards
- Regulation system
- Electric transmission
- Grid connection

Step 4: Legislative assessments and approvals

In each relevant country certain legislative and regulatory requirements may be applicable when installing a HK system in a canal system. The entities which must be considered in a

Table 2. Small-scale HK assessment and approval guidelines

Aspect	Entity	Requirement
Water/infrastructure use	Department of Water and Sanitation (DWS)	[If] DWS is the owner of the infrastructure, the user [may] require written permission from DWS to use the infrastructure to anchor the HK turbines for hydropower generation and to provide access to the land
	Water user association (WUA) or Water Board (WB) involved	The project owner may require the statutory body's written permission to non-consumptively make use of the water in the infrastructure, to access and modify the infrastructure and receive acknowledgment and approval for the project
Environmental	Provincial Department of Environmental Affairs (DEA)	Provision of an environmental authorisation should a listed activity according to the National Environmental Management Act be triggered
Electricity	National Energy Regulator (NERSA)	Needs to be informed of the project and depending on the extent of the electricity generation and distribution infrastructure provided, issue licences for these activities
	Eskom	May be consulted if grid-integration, wheeling or feed-in is considered

South African context are listed in Table 2 (it must be noted that this may change as changes in legislation are introduced and the entities involved may vary as the use of electricity and the size of HK system varies). Although specific to SA, similar entities and aspects may be globally representative. Greater detail on these processes can be seen in (Scharfetter and Van Dijk, 2017).

Step 5: Meeting with stakeholders

When all stakeholders have been identified, meetings should be arranged with each relevant entity to obtain the required approvals and if required arrange for agreements to be formalised to utilise the water infrastructure asset.

Due to the large changes in process and legislation as the project owner varies (municipality/ private entity) it must be noted the development process described assumes the municipality as project/asset owners for the purpose of this paper (and this process may vary in other cases).

Step 6: Capital procurement

This phase involves procurement of all equipment required for installation of a complete HK system. This may include (but is not limited to) the following:

- Procurement of materials and equipment required for civil works:
 - Procurement of labourers for site work
 - Anchor block/foundation material
 - Turbine fastening equipment (e.g. anchor cables)
 - Bridge structure, hoisting mechanism
 - Fencing
 - Control room
- Electrical and transmission equipment:
 - Turbine control equipment (regulators, inverters, dump-loads)
 - Transmission lines / grid connectors / mini grid
 - Measuring and monitoring equipment and cabling
- Mechanical equipment:
 - Turbine (blades, generator)
 - Generator
- Spare parts required for maintenance

Step 7: Project implementation and optimisation

During project implementation the considerations shown in Fig. 8 are of importance and must be considered, after which the project can be implemented and tested for correct functioning.

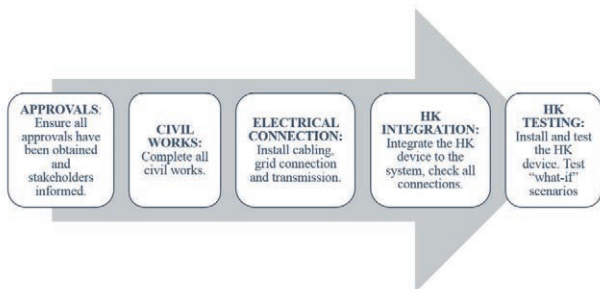


Figure 8. Implementation procedure

During project implementation it may be necessary to optimise the system to obtain a higher output. Some methods of optimisation include:

- Turbine confinement (Shives, 2008)
- Shroud addition (Gaden, 2007)
- Diffuser addition (Gaden, 2007) (Riglin et al., 2014)
- Channel modification (Khan et al., 2008)
- Multiple turbine application (Riglin et al., 2014)

Step 8: Operation and maintenance

The operation and maintenance of the HK installation will govern the lifespan and functioning of the device. The following aspects are of importance for this procedure:

- Training of operation and maintenance staff
- Maintenance plan and schedule of inspections and cleaning
- Issuing maintenance and operating manuals

During the first year of project implementation, additional problems (such as seasonal debris variations) should be evaluated and the relevant mitigation measures put into place. Correct operation and maintenance of the system will result in an efficient system with a long operational life.

Step 9: Monitoring and evaluation

For new technologies such as HK systems in South Africa, monitoring and evaluation of the system over its operational life is an important step to understand system functioning within the South African context. This will result in the construction of more reliable, cost-effective systems in the future once a clearer understanding of the roles of all the variables has been attained.

In addition to this, monitoring the systems will allow risk mitigation by allowing quick response times in blockage or potential canal overflow situations, and will also allow reporting when unplanned maintenance may be required due to observed decreases in system outputs. Some important technical aspects which should be monitored include the following:

- Power output and relevant velocity at specific time intervals
- Damming levels (allowing blockage monitoring)
- For future feasibility studies, monitoring and data collection of the following aspects could be valuable:
- Costs incurred over the operational life (maintenance/operational costs)
- Problems encountered during operational life (blockages/floods/overtopping/theft)

Boegoeberg case study

The first HK installation in SA was designed and implemented through benchmarks highlighted in the developed design and implementation process, which are summarized below. The project was initiated by allocated funding for small-scale hydropower development in the !Kheis Local Municipality (LM) in the Northern Cape Province. This LM has a similar problem to numerous other municipalities throughout the country, this being a lack of revenue due to large sections of the community being unemployed and not able to pay for services.

The Boegoeberg irrigation canal runs adjacent to numerous towns and small settlements in the municipality. Small-scale hydropower schemes along this route could allow a sustainable means of supplementing the income of this municipality and reducing its reliance on Eskom.

Selection of a site in the !Kheis Municipality was based on the following criteria:

i. Technical viability:

- Adequate flow velocity
- Demand for electricity
- Adequate cross section (large enough for turbine and turbulence recovery)

ii. System sustainability:

- Anti-theft/ anti-vandalism
- Identified electricity need
- Community support
- System integration effect on canal functioning (water supply should not be compromised by addition of system)

iii. Site resilience:

- Overflow protection of damming effect
- Accessible to municipal workers in case of emergency (e.g. blockage)

iv. Project specifications:

- A representative canal section of what is available in most areas

The Groblershoop water treatment works (WTW) site lies in the town of Groblershoop in the !Kheis Municipality. This plant supplies water to the majority of the surrounding areas. The WTW pump station is currently partially supplied by a small solar generation plant erected at the WTW facility. The town of Groblershoop lies approximately 38 km downstream of the canal offtake (Boegoeberg Dam). The approximate electricity usage of the pumps at the WTW is 25 kW. The selected site lies on the upper region of the Boegoeberg irrigation canal; the layout of the site can be seen in Fig. 9.

The canal has a rectangular shape with a slightly cambered canal bed, the canal drawings were obtained from the Boegoeberg Irrigation Board in original design drawing format. As shown, the selected section is a combination of two straight



Figure 9. Groblershoop site

sections; this is ideal for installation as there is less turbulence in these sections. The downstream canal section including the bend as well as the typical section parameters can be seen in Fig. 10. Sections further upstream were not considered as these lie at too great a distance from the control room/demand and would result in long transmission cable lengths (and therefore higher costs and increased possibility of vandalism).

A DWS gauging station (D7H17) which has recorded significant amounts of historical flow data lies approximately 63 km upstream of the installation section of the canal. The flow recorded over a ± 10 -year period can be seen in Fig. 11. The trend indicates an annual average of 40 days of zero flow (during maintenance etc.). The average flow during operational periods was found to be $12.5 \text{ m}^3/\text{s}$ (Fig. 11). A localised flow (due to other offtakes) of around $6.6 \text{ m}^3/\text{s}$ was measured at the selected location for the installation of the HK device through the use of a speedy doppler-type velocity sensor and ultrasonic level sensor. With the flow of $6.6 \text{ m}^3/\text{s}$, the flow velocity was expected to be 1.1 m/s and measured to be an average of 1.05 m/s (measured by a doppler velocity sensor placed on the canal bed).

The turbine selected for the site is shown in Fig. 12 and is specifically designed by Smart Hydropower GmbH for canal systems. It is placed directly on the flat canal bed and anchored by steel cables to the sides/bottom of the canal. System specifications can be seen in Table 3. The turbine is equipped with an underwater permanent magnet generator providing AC power to the system. To achieve maximum power output from this specific turbine a flow velocity of 3.1 m/s is required. The output curve of the generator is provided in Fig. 13.

The Groblershoop WTW already had a facility where power could be generated for one of the pumps by means of solar panels which are also grid connected to supply the remaining pumps (however, this was not functioning due to damaged control equipment). This allowed integration of the HK plant with the existing system (and control equipment).

Table 3. Turbine specifications (SHP, 2017)

Output (W)		250–5 000
Dimensions	Length (mm)	2 640
	Width (mm)	1 120
	Height (mm)	1 120
Rotational speed (r/min)		90–230
Weight (kg)		300
Number of rotor blades		3
Rotor diameter (mm)		1 000



Figure 12. Kinetic turbine for deployment in canal (SHP SMART free stream turbine) (SHP, 2017)

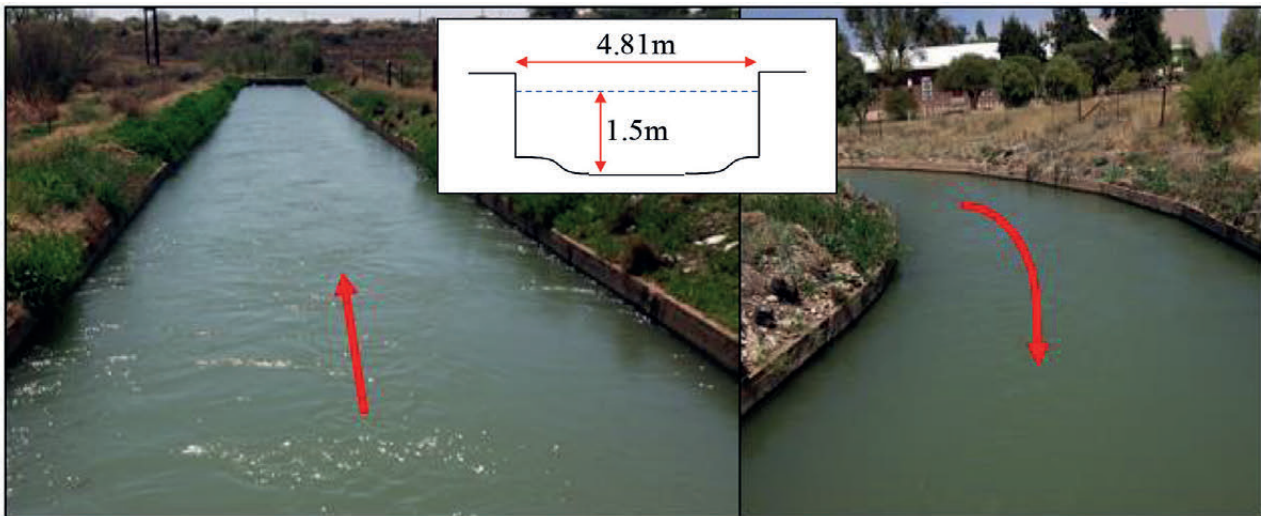


Figure 10. Canal system at the Groblershoop site

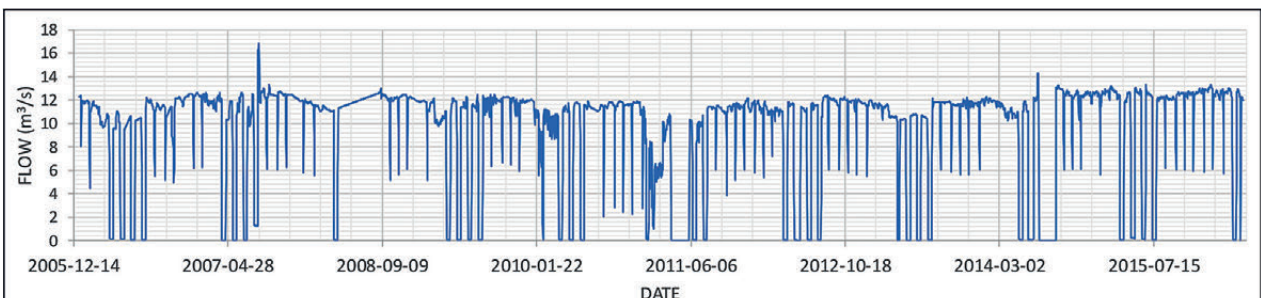


Figure 11. Flowrate in the Boegoeberg Canal Station D7H17 (Dec 2005 to April 2016)

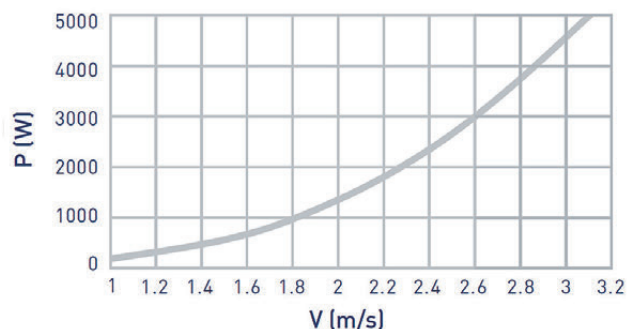


Figure 13. Generator capability (SHP, 2017)

The electronic control system is the intelligence of this renewable energy system (Fig. 14). Integrated inside the electrical cabinet is a reverse control (to remove accumulated debris from the blades), optional auxiliary devices, customizable battery storage interface (it has a built-in battery charger) and integrated monitoring system (monitoring control board). The basic layout of the installation is depicted in Fig. 15.

For hydropower projects developed at local level, the institutional powers, functions, roles and responsibilities of municipalities are of utmost importance. In the context of the four major regulatory requirements necessary to initiate and implement small-scale hydropower projects, namely a Water Use Licence (in terms of the National Water Act (NWA) (RSA, 1998a)), an Environmental Authorisation (in terms of the National Environmental Management Act (NEMA) (RSA, 1998b)), an Electricity Generation Licence and an Electricity Distribution Licence (in terms of the Electricity Regulation Act (ERA) (RSA, 2006)), the DWS (or the regional departmental offices or the catchment management agency, where these are established and have the authority), the provincial office of the Department of Environmental Affairs (DEA) and National Energy Regulator of South Africa (NERSA), would be the primary stakeholders involved in this type of hydropower

project, over and above the project initiator.

In the context of local electrification through small-scale hydropower technologies, the DWS is a primary stakeholder. The DWS has the mandate to protect and manage SA's water resources. Furthermore, the DWS owns waterworks across SA such as weirs, irrigation canals and dams which could be utilised to generate electricity (in this case the Boegoeberg irrigation canal). For this study, the project-specific institutional stakeholders are identified in Table 4. The roles of each stakeholder during the project are shown under project specification. The project formed part of the 'Draft policy on sustainable hydropower generation' which allowed a simpler process of approval.

The infrastructure components required for the installation of the HK turbines are listed in three sections, namely, civil components, electro-mechanical components and electrical components:

i. Electro-mechanical components

- SHP free stream turbine (Equipped with 5 kW underwater generator and rotor)

ii. Civil components

- Concrete block foundation for cable hoist
- Turbine lifting system

iii. Electrical components

- Turbine control equipment
- Flow meter
- Level sensors

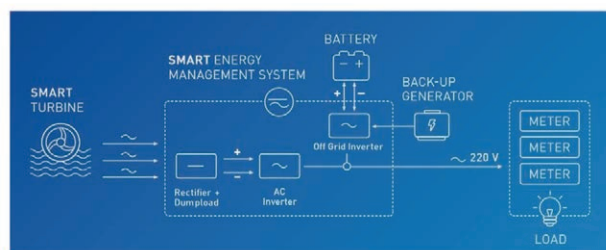


Figure 15. System layout (SHP, 2017)

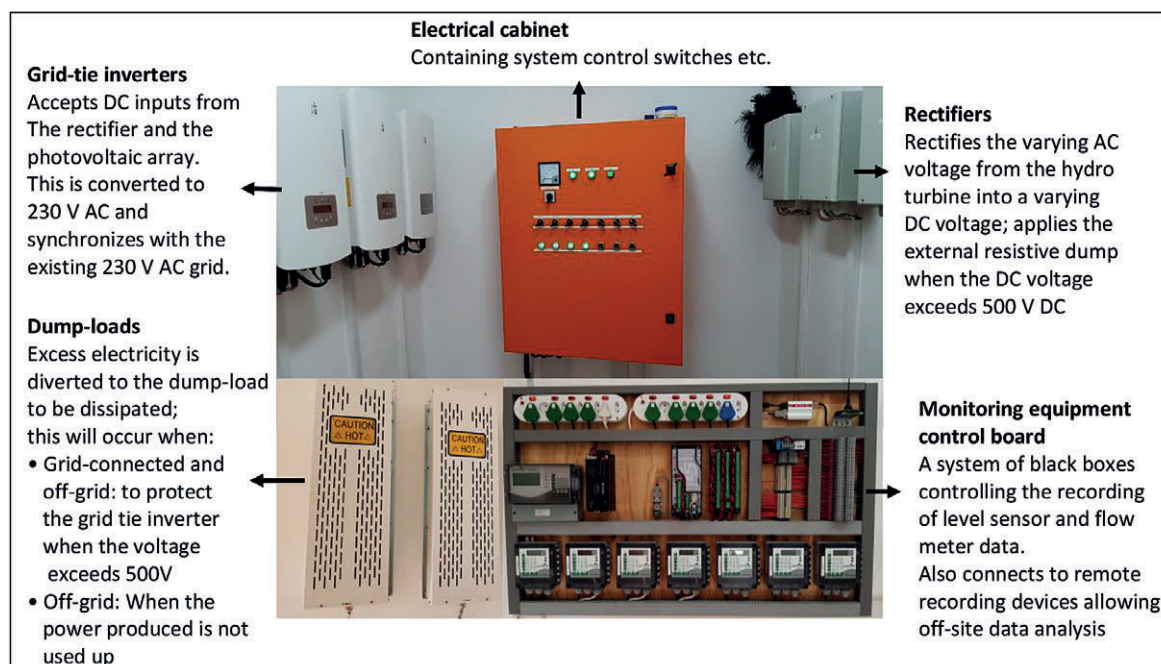


Figure 14. Electrical management system components (Control room)

Table 4. Project-specific institutional stakeholders Groblershoop Water Treatment Works

No.	Stakeholder	Role and responsibility	Project specification
1	Department of Science and Technology (DST)	Funder	Funds obtained for the project duration
2	WRC	Implementing Agent for the DST	Reports submitted
3	DWS	Owner of irrigation canal Owner of land on which the irrigation canal is located Custodian of the water resource	A series of meetings set-up to obtain approval Signee on the memorandum of agreement (MoA) between all entities involved to ascertain approval
4	!Kheis LM	Hydropower project owner (asset will be listed onto the LM's Asset Register) All regulatory submissions would be in the name of the !Kheis LM and for the !Kheis LM Operation & maintenance of hydropower generation plants and the electricity transmission infrastructure	A series of meetings were set-up to obtain approval and knowledge of municipal roles and responsibilities concerning the installation Co-signee of MoA
5	Citizens of the !Kheis LM	Beneficiaries of the project's service delivery outcomes	Local labourers were hired and trained to perform maintenance on such an installation Continuous employment to operate and maintain
6	Boegoeberg WUA	Operator of Boegoeberg canal	Meetings were set-up to inform WUA of all construction plans and obtain approval Notified and requested approval during any modification/alteration to canal Co-signee of MoR
7	NERSA	Electricity regulator	No licence required as the electricity produced is for 'own-use' and small scale of generation
8	Eskom	Electricity distributor	No approval/licensing required, no Eskom owned distribution lines used
9	Northern Cape Provincial DEA	Provision of an environmental authorisation should a listed activity according to the NEMA be triggered	No triggers for this authorisation, therefore not required
10	University of Pretoria	Planning and design engineers	Responsible for project implementation

All equipment required for installation was procured together with the completion of civil works (listed above) to allow turbine installation. The velocity in the canal was measured as around 1.2 m/s at the point of installation which provided a lower range of application on the power curve (Fig. 13). This also proved true at the initial testing phase where a maximum output of 228 W was achieved (placing the turbine into the flow area as specified by manufacturer). Due to this optimisation of the section was considered. To achieve a higher velocity through the turbine the option of canal narrowing at the installation points thus creating a venturi effect and speeding up the flow entering the turbine was considered. This was done through the use of steel plates anchored onto a hinge point on the canal wall.

To allow installation of a HK system in SA more information is required on the impacts such devices could have on the canal functioning and water security as a whole. A numerical model using HEC-RAS was developed for both scenarios with and without the narrowing. The findings of such numerical simulations were compared with the constructed test installation from which data could be collected and analysed.

To allow an indication of 'worst case' infrastructural influences, the numerical modelling was used in the testing process identified in 2 phases. Phase 1 included the 'worst case' of full turbine grid blockage which was modelled by an inline structure. The length of the blockage was chosen as 3 m which is the length of the turbine plus an additional 0.75 m allowing for possible debris build up (trees, shrubs etc.)

The maximum damming level was found to be 40 mm just upstream of the turbine, which then returns to normal flow

depth approximately 2.7 km upstream of the installation point. This large backwater effect resulted from the relatively flat canal bed and subcritical flow pattern. The highest damming levels lie within the available freeboard and therefore no overtopping was expected. The damming levels comparing to the 'no turbine' scenario can be seen in Fig. 16.

Phase 2 including narrowing the canal by a total of 1.2 m per side. To model the narrowed width new cross-sections were added to the HEC-RAS model to simulate the new incrementally narrowed section. The blocked turbine was added to the narrowed section, similarly to Phase 1.

Both scenarios at the narrowing of the canal, first, with no installed turbine and, second, with a fully blocked turbine, can be seen in Fig. 17. The maximum damming level of the Phase 2 canal narrowing with a fully blocked turbine grid was found to be around 290 mm just upstream of the blockage and the recovery distance of damming levels was 6.3 km upstream. For the scenario of narrowing alone without the turbine the damming was found to be around 70 mm with a recovery distance of 4.2 km upstream.

During the blocked turbine Phase 2 modelled scenario with the narrowed width, the canal showed overtopping at distances of 440–1 800 m, which lies about 360–1 700 m upstream of the blockage. Consideration was given to heightening of this section when permanent narrowing was considered; however, due to the original canal drawings being used to create the model, and alterations (in terms of heightening the canal wall) having been made since initial construction, it allowed the flexibility for the damming required for the narrowing (without alterations).

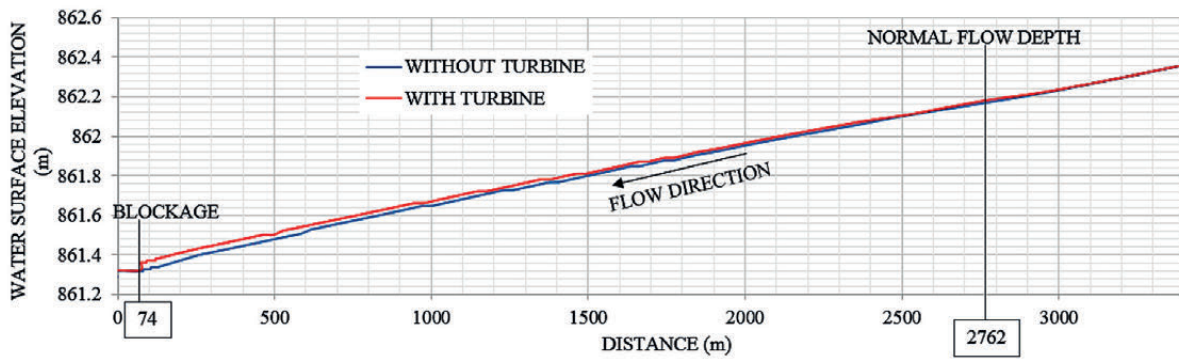


Figure 16. Phase 1 blockage damming levels

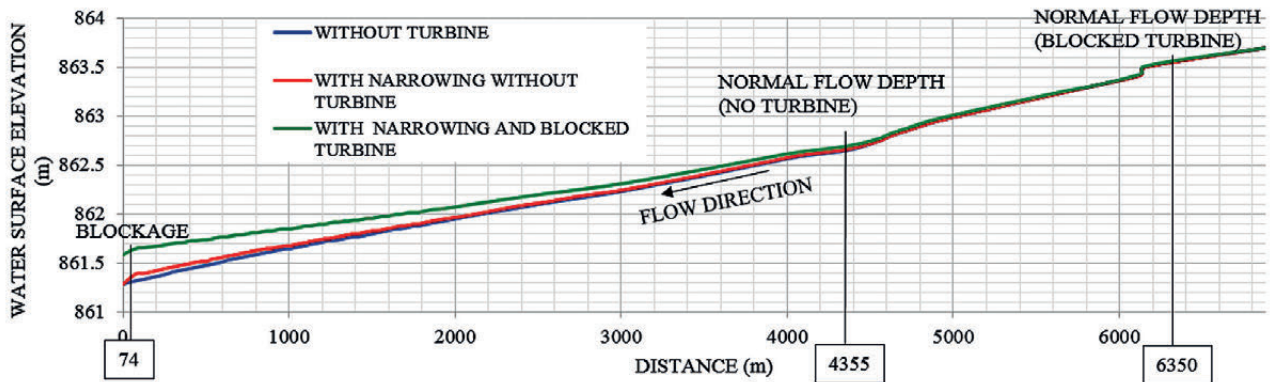


Figure 17. Phase 2 blockage damming levels

The physical testing phase included the construction indicated in Fig. 18. Phase 1 included a single turbine which was temporarily anchored from 8 mm diameter steel cables (lifting capacity of 750 kg) attached to concrete block foundations.

The damming effects of the canal narrowing due to the installation of the HK turbine were measured by means of level sensors at the installation point and a distance upstream (allowing measurement of the upstream effects). The flow velocity was measured at various points by use of a hand-held velocity propeller sensor (wading rod) to allow understanding of the flow characteristics. The results proved a 70% increase in velocity from the canal narrowing with a maximum damming effect of 100 mm (found to be 70 mm with numerical modelling); additional details are shown in Table 5.

For Phase 2 testing the turbine was placed in the flow (within the narrowed flume section) and the water levels,

flow velocity and the impact on water levels further upstream were once again recorded. The freeboard available proved to be sufficient for the damming effects (Table 5). The effects on the canal system were compared to numerical model results and the model was calibrated accordingly. This allowed re-modelling in the numerical model and thus assisted with the design of the final system with greater accuracy.

The final conclusions reached from the correlated physical and numerical testing results can be seen in Table 5.

From the results of the temporary test installation a permanent design was completed. A total of 4 turbines were installed in series along two narrowed sections 45 m apart (2 turbines in each section) with both sections being narrowed to a width of 2.41 m. The final design results in an increased output of 3.9 kW (increased velocity of 2.8 m/s) per turbine and thus a total system production of around 15.6 kW 89% of

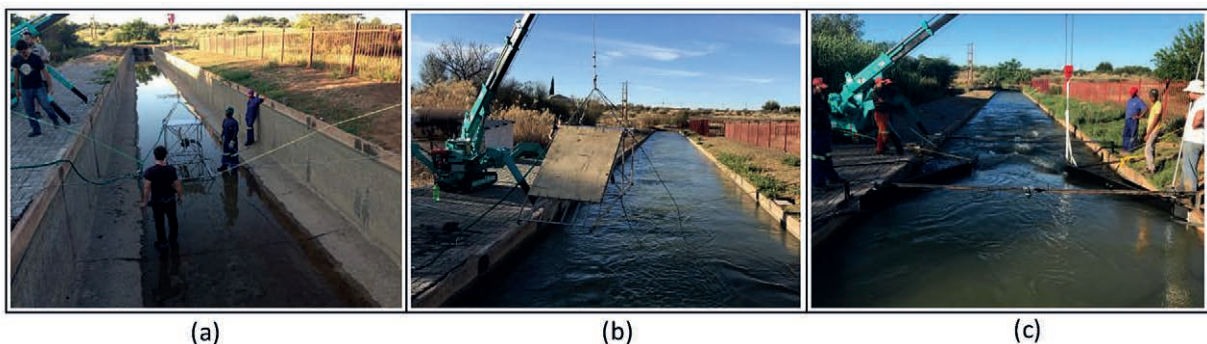


Figure 18. Testing (a) turbine placement testing (b) blocked turbine testing (c) Narrowed section testing

Table 5. Testing results summary

Scenario		Maximum damming effects	Conclusions reached
Phase 1: Cable anchored turbine	Normal operating turbine	30 mm maximum damming recorded at the point of installation and 50 m further upstream	The turbine proved to have minimal effects on the canal; with no visual or noise disturbances and minimal damming
	Blocked turbine (Fig. 18b)	60 mm maximum damming at the point of installation, reducing to 50 mm within 50 m upstream	If the turbine grid is fully blocked by debris, etc., no interruption of the water supply of the Boegoeberg canal will be experienced (no overtopping etc.)
Phase 2: Canal narrowing with cable anchored turbine	Normal operating turbine	100 mm maximum damming at the point of installation, reducing to 90 mm within 50 m upstream	By narrowing the canal section to obtain optimal output the damming effect remains within the available freeboard; therefore, the final installation should function well with minimal disturbance on the system
	Blocked turbine	140 mm maximum damming at the point of installation, reducing to 100 mm within 50 m upstream	If the turbine grid is fully blocked by debris, etc., no significant interruption of the water supply of the Boegoeberg canal will be felt (no overtopping etc.); improvements can be made to allow flow over narrowing plates or raising of the channel walls in this area

the year (accounting for maintenance and canal dry times). This design can be seen in Fig. 19 and is currently under construction. The design of the system allows simplified removal of each turbine if necessary (in case of blockages) and operational flexibility; the entire system is removable and can be altered and adjusted if necessary.

During the design life of a HK device, efficient functioning is crucial to obtaining design outputs. The energy production is predicted and considered during the preliminary design analysis; this is directly related to the production and therefore profitability/success of the installation. During the operational phase testing, routine maintenance was required due to debris build up (as indicated in Fig. 20). The maintenance strategy that was found to be effective for correct functioning of the installation through its design life involved a single operator hired with the following duties:

- Weekly visual checks and clearing of visible blockages
 - Turbine removal and cleaning at 2-week intervals
 - Turbine inspection and clearing at all canal shutdown periods, and first day of full flow (as build up occurs during start-up)
 - Impromptu response to alarm system linked to level sensors indicating sudden increases in water level and thus unplanned clearing/removal and cleaning as required.
- During the HK test installation various risks were identified which could impact future developments and result in delays or additional costs:
- Potential of theft and vandalism of the installation during

construction, which could result in additional expenditure for fencing and additional security measures

- The test installation site is in close proximity to a school in the town of Groblershoop. In summer temperatures rise above 40°C and children tend to swim in the canal systems (although forbidden); measures had to be taken to prevent

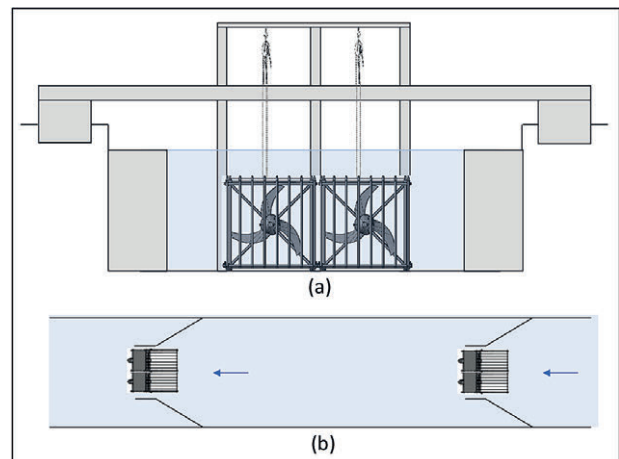


Figure 19. (a) Front view of final design (b) top view of final design with turbines 45 m apart



Figure 20. Debris blockage (a) during maintenance period (b) major debris build up after a 2-week interval (c) turbine removal and cleaning at 2-week interval

swimming in the section (to prevent injury or fatalities from spinning blades) and additional protective casing on the turbine (resulting in additional costs)

- Larger than expected debris, such as deceased livestock floating down the canal, due to livestock and wildlife drinking from the canal (as no barrier exists around the canal) and occasionally falling in. This could potentially result in turbine blockages; therefore, a warning system must be included in the installation to allow for quick removal in such cases.
- The legislation in South Africa is not yet in place to allow entities (municipalities or developers) to utilise DWS infrastructure for hydropower development in any form. This test installation requires special permission and a memorandum of agreement (MoA) between the !Kheis Local Municipality, the Boegoeberg Water User Association (WUA) and the DWS and is a test case for new legislation.
- Due to a lack of specific policy instated in South Africa, which would allow a streamlined process of installation, delays in the process may occur.
- The low velocities in most canal sections result in low efficiency systems due to the exponential behaviour of the power curve; however, as narrower sections or confined sections with a greater flow velocity are selected for installation, additional problems such as complete blockage (due to the turbine covering the majority of the smaller cross section) and large damming effects may occur. Due to this the risk of overflow and thus water loss to the WUA increases and thus the viability of the project decreases (civil works costs increase); it can also result in reluctance of the WUA to agree to the installation.
- Maintenance of the HK system remains a problem in South Africa; municipalities do not have the money and thus not the workforce to complete maintenance of the systems. The debris during short testing periods was evident which could result in large debris build-up over time and with no maintenance this could result in an inefficient system which is prone to blockage stoppages.

SUMMARY AND CONCLUSIONS

South Africa, the country with the second largest installed hydropower capacity in Africa, produces less than 5% of its total electricity demand from hydropower. However, multiple studies have indicated the unexploited potential available (Kusakana and Vermaak, 2013) (Loots et al., 2015). This highlights the lack of and need for innovative hydropower research and installations in Africa. In South Africa, mainly due to a relative scarcity of surface water, there is a prevailing perception that the potential for hydropower development is rather low. The largest percentage (>60%) of water is made available for the agricultural sector, delivered through an extensive water supply network consisting of weirs/intake structures, pipelines, tunnels, siphons, canals, chutes, etc.

Although there is significant HK potential, globally there is little technical literature or validation of successful installations of HK systems. The objective of this study was to address this problem and develop a HK design and implementation process to provide a guideline to HK installations. Literature on small-scale hydropower and specifically HK systems both locally and internationally were reviewed. Although no HK installation had been constructed in South Africa, the potential of these systems in rivers has previously been investigated at a very high level when considering several technologies (Koko and Kusakana,

2014) and a DWS asset management study illustrated the vast network of canals allowing further HK potential in the country.

A pilot project of a HK installation to showcase the installation possibilities was constructed. The working of the device was tested and as a result of the methodology followed during the study the following conclusions have been drawn. These conclusions should be read in the context of the study:

- The modular HK system was installed with all required aspects to operate the system on a permanent basis. This allowed a showcase of the technology in the South African context.
- With the required maintenance met (from debris/blockages) and the correct site selection, HK technology could be used as an additional sustainable renewable energy source in South Africa.
- As the majority of canal systems pose a controlled environment with a constant uniform supply of water (except during short shut-off intervals), with careful site selection and feasibility studies, HK energy systems in canal networks in South Africa can be a feasible alternative energy source. With pilot installations such as these, the way to a streamlined, simplified regulatory and legislative process for the installation of HK systems in South Africa is demonstrated.
- The design, implementation and testing of a HK turbine within the South African environmental context highlighted all important aspects of the design and procedures necessary, thus assisting future installations.
- When designing new canal systems, the potential opportunity to include HK systems should be considered and some design requirements already provided, e.g., designing parts of the canal for faster velocities and higher canal walls in anticipation of damming.

Through the testing and investigation of this modern form of hydropower, a design and implementation process for the integration of HK turbines was formulated, which allows for an improved and streamlined implementation process where uncertainties are reduced.

ACKNOWLEDGEMENTS

This research was made possible by the financial support of the South Africa Department of Science and Technology, Water Research Commission and University of Pretoria whose support is acknowledged with gratitude. Additionally, the project was completed with support from the Boegoeberg Water User Association as a pilot project under the DWS General notice: Draft Policy on Sustainable Hydropower Generation.

REFERENCES

- CAPTAHYDRO (2017) Capta SC: For fast flowing canals. URL: <http://www.captahydro.com/capta-sc.html> (Accessed 9 August 2017).
- DME (Department of Minerals and Energy, South Africa) (2003) *White Paper on Renewable Energy*. Department of Minerals and Energy, Pretoria.
- DWS (Department of Water and Sanitation, South Africa) (2015) Strategic overview of the water services sector in South Africa. Version 4. Department of Water and Sanitation, Pretoria.
- DWS (Department of Water and Sanitation, South Africa) (2016a) Condition Assessment Audit of irrigation scheme infrastructure: Cluster: Central, Scheme: Sand-Vet GWS. Scheme Report. Department of Water and Sanitation, Pretoria.
- DWS (Department of Water and Sanitation, South Africa) (2016b) Condition Assessment Audit of irrigation scheme infrastructure: Cluster: Central, Scheme: Orange Riet GWS. Scheme Report. Department of Water and Sanitation, Pretoria.

- DWS (Department of Water and Sanitation, South Africa) (2016c) Condition Assessment Audit of irrigation scheme infrastructure: Cluster: Central, Scheme: Hartbeespoort GWS. Scheme Report. Volume 1 (2.4). Department of Water and Sanitation, Pretoria.
- GADEN DLF (2007) An investigation of river kinetic turbines: Performance enhancements, turbine modelling techniques and an assessment of turbulence models. MSc thesis, University of Manitoba.
- GADEN DLF and BIBEAU EL (2006) Increasing power density of kinetic turbines for cost-effective distributed power generation. University of Manitoba, Manitoba.
- GUNAWAN B, NEARY VS, MORTENSEN J and ROBERTS JD (2017) Assessing and testing hydrokinetic turbine performance and effects on open channel hydrodynamics: an irrigation canal case study. U.S Department of Energy, Albuquerque. <https://doi.org/10.2172/1367421>
- HYDRO4AFRICA (2017) hydro4africa.net. URL: http://hydro4africa.net/HP_database/country.php?country=South%20Africa (Accessed 05 April 2017).
- IFC (2013) Toward universal energy access: designing a new household electrification strategy for South Africa. International Finance Corporation, South Africa.
- KHAN MJ, BHUYAN G, IQBAL J and QUAICOE JE (2008) Hydrokinetic energy conversion systems and assessment of horizontal and vertical axis turbines for river and tidal applications: A technology status review. *Appl. Energ.* **86** 1823–1835. <https://doi.org/10.1016/j.apenergy.2009.02.017>
- KOKO PS and KUSAKANA K (2014) Techno-economic analysis of an off-grid micro-hydrokinetic river system as a remote rural electrification option. Central University of Technology, Free State, South Africa.
- KUSAKANA K and VERMAAK HJ (2013) Hydrokinetic power generation for rural electricity supply: Case of South Africa. *Renewable Energ.* **55** 467–473. <https://doi.org/10.1016/j.renene.2012.12.051>
- LOOTS L, VAN DIJK M, BARTA B, VAN VUUREN SJ and BHAGWAN JN (2015) A review of low head hydropower technologies and applications in a South African context. *Renewable Sustainable Energ. Rev.* **50** 1254–1268. <https://doi.org/10.1016/j.rser.2015.05.064>
- MEAAI (2011) *Energy Report 2011*. Ministry of Economic Affairs, Agriculture and Innovation, The Netherlands.
- RSA (Republic of South Africa) (1998a) National Water Act. Act No. 36 of 1998. *Government Gazette* 19182. Government Printer, Cape Town.
- RSA (Republic of South Africa) (1998b) National Environmental Management Act. Act No. 107 of 1998. *Government Gazette* 19519. Government Printer, Cape Town.
- RSA (Republic of South Africa) (2006) Electricity Regulation Act. Act No. 4 of 2006. *Government Gazette* 28992. Government Printer, Cape Town.
- RIGLIN J, SCHLEICHER W and OZTEKIN A (2014) Diffuser optimisation for a micro-hydrokinetic turbine. *ASME International Mechanical Engineering Congress and Exposition*, Montreal, Canada. <https://doi.org/10.1115/imece2014-37304>
- SCHARFETTER B and VAN DIJK M (2017) Legislation governing the implementation of small-scale hydropower projects for rural electrification in South Africa. *J. Energ. South. Afr.* **28** (2) 14–28. <https://doi.org/10.17159/2413-3051/2017/v28i2a2005>
- SHIVES MR (2008) Hydrodynamic modelling, optimisation and performance assessment for ducted and non-ducted tidal turbines. Masters dissertation, Carleton University, Ottawa.
- SHP (2017) Smart Hydro Power. URL: <https://www.smart-hydro.de/renewable-energy-systems/hydrokinetic-turbines-river-canal/> (Accessed 15 January 2017).
- VAN VUUREN SJ, BLERSCH CL and VAN DIJK M (2011) Modelling the feasibility of retrofitting hydropower to existing South African dams. *Water SA* **37** (5) 679–692. <https://doi.org/10.4314/wsa.v37i5.5>
- WHITE J (2011) *Viability of Small-Scale Hydropower in South Africa*. University of Cape Town. Cape town, South Africa. <https://doi.org/10.15641/ghi.v2i1.728>

Garden irrigation as household end-use in the presence of supplementary groundwater supply

Bettina Elizabeth Meyer¹ and Heinz Erasmus Jacobs¹

¹Department of Civil Engineering, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa

ABSTRACT

Garden irrigation is a significant and variable household water end-use, while groundwater abstraction may be a notable supplementary water source available in some serviced residential areas. Residential groundwater is abstracted by means of garden boreholes or well points and – in the study area – abstracted groundwater is typically used for garden irrigation. The volume irrigated per event is a function of event duration, frequency of application and flow rate, which in turn are dependent on numerous factors that vary by source – including water availability, pressure and price. The temperature variation of groundwater abstraction pipes at residential properties was recorded and analysed as part of this study in order to estimate values for three model inputs, namely, pumping event duration, irrigation frequency, and flow rate. This research incorporates a basic end-use model for garden irrigation, with inputs derived from the case study in Cape Town, South Africa. The model was subsequently used to stochastically evaluate garden irrigation. Over an 11-d period, 68 garden irrigation events were identified in the sample group of 10 residential properties. The average garden irrigation event duration was 2 h 16 min and the average daily garden irrigation event volume was 1.39 m³.

Keywords: garden irrigation, end-use, groundwater, residential water demand

INTRODUCTION

Residential water consumption is typically categorised into indoor end-uses and outdoor end-uses. Previous studies suggest outdoor use to be seasonal, driven by weather-related variables, whilst indoor use has been found to be relatively constant (Fisher-Jeffes et al., 2015). Outdoor use is also considered more unpredictable than indoor use (Hemati et al., 2016). Howe and Linaweaver (1967), in an early study of residential water demand, reported on the inelastic nature of indoor water use versus the elastic nature of outdoor use, meaning that outdoor use was found to be more sensitive to a change in inputs than indoor use. Jacobs and Haarhoff (2007) used elasticity and a sensitivity parameter to identify pan evaporation, an irrigation factor, lawn surface area (lawn size) and the vegetation crop factor (lawn grass genotype) as the most notable parameters when modelling outdoor water use.

Various parameters describing outdoor use have received attention as part of earlier work, including garden irrigation (Beal et al., 2011), lawn size (Runfola et al., 2013), swimming pools (Fisher-Jeffes et al., 2015), and water use from the outside tap (Makwiza and Jacobs, 2017). Household water leakage was also addressed in earlier work (Britton et al., 2008; Lugoma et al., 2012). The most notable outdoor end-use in an unrestricted scenario is garden irrigation. Garden irrigation is often reported as a notable part of the total per-capita consumption (Willis et al., 2011). It is unsurprising that outdoor use is the primary target during water restrictions, with earlier studies reporting on reduced water use during water restrictions, mainly due to reduced outdoor use (Jacobs et al., 2007).

Despite the attention to various facets of outdoor use in earlier work, end-use studies have paid limited attention to water supply from supplementary household water sources (Nel et al., 2017). This research focuses on modelling garden irrigation as an end-use in an unrestricted scenario, where

groundwater was abstracted from privately owned groundwater abstraction points (GAPs) as supplementary water source. Residential GAPs include garden boreholes and relatively shallow wellpoints.

Consumers may turn to alternative non-potable water sources such as rainwater, groundwater or greywater during stringent water restrictions. The quality of these resources typically limits application to nonpotable uses, such as garden irrigation (MacDonald and Calow, 2009). According to Nel et al. (2017), groundwater use is the most notable supplementary source in terms of the expected supply volume. Many privately owned GAPs are in use across South Africa, with at least one notable case study in the Cape Town region (Wright and Jacobs, 2016). Monitoring of household groundwater abstraction in South Africa is poor and published information regarding yield, flow rate, and/or the pumping event duration of household GAPs is limited.

Garden irrigation as outdoor end-use

The contribution of garden irrigation to the total household water use varies by season (Parker and Wilby, 2013) and also varies from country to country and even from house to house. Garden irrigation tends to be higher during dry, hot seasons, and increases with reduced rainfall (Jacobs and Haarhoff, 2004; Parker and Wilby, 2013) and increased maximum daily temperatures (Rathnayaka, 2015), for example. The garden event duration and number of occurrences are contingent on the method of irrigation. Roberts (2005) identified three main irrigation methods, namely, hand-held hose, manual sprinkler and automated sprinkler. The latter contributed most to garden irrigation volumes from the end-use study conducted by Roberts (2005) in Australia. The same three irrigation methods were found in the study area during this research.

Literature includes various reports of garden irrigation expressed as a percentage of the total household water demand, in order to explain the significant contribution of garden irrigation to total household water use. The perceived

*Corresponding author, email: hejacobs@sun.ac.za

Received 5 Feb 2018; accepted in revised form 27 May 2019

percentage of residential water demand used for garden irrigation in South Africa, based on an annual average, was reported to vary between 0% and 70% (Veck and Bill, 2000). More recent end-use studies conducted in South Africa reported the percentage of average annual household water demand ascribed to garden irrigation as 40% to 60% (Du Plessis and Jacobs, 2015) and 58% (Du Plessis et al., 2018) in different South African study samples.

End-use studies conducted in other parts of the world also report a wide range of values expressing garden irrigation as a percentage of the total household water use. In Australia, the percentage of household water demand used for garden irrigation ranges from 5% (Beal et al., 2011) to 54% (Loh and Coghlan, 2003). Arbon et al. (2014) reported a strong seasonal impact in Adelaide, Australia, with a 2013 winter mean of 153 L/person per day increasing to 498 L/person per day in the summer of 2013/14. The average annual use was 245 L/person per day and 289 L/person per day in 2013 and 2014 respectively, that could indicate a garden irrigation contribution of 50% to 70% of the total annual household demand; a significant shift in the diurnal pattern was noted, with an afternoon peak more prominent during summer. A lower outdoor use contribution of 15% was reported at high-income detached houses by Ghavidelfar et al. (2018) in Auckland, New Zealand. Wasowski (2001) conducted an end-use study in the United States of America and stated that between 40% and 60% of annual average residential water demand is attributed to garden irrigation.

Rationale

Suburban households in the case study area of Cape Town, South Africa, are accustomed to a reliable supply of potable water from the pressurised water distribution system. However, the rising block-based water tariff was relatively high and, also, outdoor water use from the distribution system was banned during water restrictions in the study area – for the period June 2017 to December 2018. Consumers subjected to emergency water restrictions turned to alternative sources of water to maintain gardens during this 18-month period. Little is known about garden water use by consumers with access to groundwater from garden GAPS; the restrictions provided the opportune time to investigate the matter. The main challenge in this study was to obtain data regarding actual groundwater use by private homeowners, who were often reluctant to share any information regarding uncontrolled and unmetered household water sources.

Research problem

An end-use model was needed to assess garden irrigation in relation to supplementary groundwater supply, while populating the model with data that could realistically represent the key unknowns.

METHODS

Parameters describing the quantity and quality of household groundwater abstraction form important inputs to end-use models of household water use. Groundwater use for garden irrigation was modelled in this study, with inputs based on measured values. Data were collected from a relatively small case study site in Cape Town, South Africa. Direct measurement of groundwater abstraction was not considered feasible and an alternative method to assess the volume of

groundwater abstracted for garden irrigation was employed.

Groundwater pumping event start times and durations were derived from continuously recorded pipe wall temperatures at each of the 10 residential properties. Ad hoc volumetric measurements were subsequently conducted at each home to gain insight into flow rates at each study home. Stochastic end-use modelling was employed to estimate the expected garden irrigation event volume of the 10 properties in the research sample. Based on information obtained during the site survey, garden irrigation volume was considered to be equal to the groundwater abstracted from GAPS for all homes in the case study.

Overview of residential end-use models

The focus of this study was on modelling water demand at a small spatial scale of single residential homes – and garden irrigation as a specific end-use of water. Numerous residential end-use models have been developed in the past; however, a model to evaluate garden irrigation in relation to groundwater abstraction as supplementary source has not yet been developed. Some examples of earlier end-use models include the Poisson Rectangular Pulse (PRP) model developed by Buchberger et al. (1996; 2003), the SIMulation of Demand End-Use Model (SIMDEUM) by Blokker et al. (2010) and the Residential End-Use Model (REUM) by Jacobs and Haarhoff (2004). REUM and SIMDEUM incorporate garden irrigation as end-use.

Experimental field tests and data analysis

Study site selection and sample group

A map of verified residential properties with GAPS in the Cape Town Metropolitan area was developed by Wright and Jacobs (2016). The sample group of 10 homes for this study was based on sub-regions where clustering of GAPS (as reported by Wright and Jacobs, 2016) was observed, followed by personal invitation to participate in the study. Relatively small sample sizes are not unusual for end-use studies. Former end-use studies had sample sizes of 28 homes (Butler, 1991), 16 homes (DeOreo et al., 1996), 37 homes (DeOreo et al., 2001), 21 homes (Buchberger et al., 2003), 12 homes (Heinrich, 2007), 10 homes (Jacobs, 2007) and 10 homes (Mead and Aravinthan, 2009).

The manageable sample size in this study also enabled the authors to inspect individual pump installations for leaks and to conduct follow-up inspections. All the houses in the sample were single residential properties, with property plot sizes ranging from 600 m² to 1 400 m². Prominent, well-irrigated gardens and lawns were present at all homes. Two residential properties from the study site each had a swimming pool; however, the homeowners assured that the abstracted groundwater was explicitly used for garden irrigation at the time of this study. The assumption that groundwater supply equalled garden irrigation was thus considered valid for the study sample. The addresses and suburb names of the study homes were omitted for anonymity, in line with ethical requirements.

Data collection methods

Residential water demand patterns should preferably be obtained by measuring actual water use (Scheepers and Jacobs, 2014); however, empirical investigations involving data collection are often faced with several logistical, time and financial constraints. Various data collection methods

were considered for this study in order to collect sensitive information that was needed to assess household groundwater abstraction. A list of empirical measurement methods is presented in Table 1, including the key advantages and disadvantages in each case, as well as a reference to earlier application. Each method was categorised in terms of feasibility as it relates to the case study. Two categories were included, namely: (1) considered for this case study and also implemented in the study; (2) considered for this study, but not used.

Equipment and temperature recording

The project plan involved recording pipe wall temperature at case study homes in an unobtrusive way, with no plumbing requirements, a short installation time and relatively low cost. The DS1922 ThermoChron Hi Resolution iButton was selected for this study, based on the relatively small size, ruggedness, accuracy, cost, and availability. The iButtons were used in this study to measure the variation in temperature of the groundwater pump delivery pipe – that is the delivery pipe of the GAP pump supplying water directly for garden irrigation. The temperature variations were subsequently used to assess water use events by determining the event duration of groundwater pumping, start and stop times.

The iButtons were preconfigured to set the start time and sample rate. ColdChain ThermoDynamics software was used for preconfiguration and to extract and save the recorded data. All the iButtons were programmed to have a sampling rate of 2 min, which was considered sufficient when compared to the relatively long events. The period of 2 min was the shortest interval available when programming the equipment. The iButtons were synchronised to start at the same time on the same date. The internal iButton memory allowed for a total recording duration of 11 d and 9 h (sample count of 8 192 records per iButton). After the iButtons were activated and before the specified start time, the iButtons were installed on the outside wall of the outlet pipe, using adhesive electrical tape. Each GAP was equipped with two synchronous iButtons to record temperature in parallel. The sample included 10 homes and data were recorded during April and May 2016. Subsequently the total data set included 110 test days, representing 11 actual calendar days for each of the 10 homes.

The iButtons were placed in three different environments (A, B, and C). Each environment type was linked to an installation that affected the temperature changes of the iButtons differently. In Environment A the pump and outlet pipes were located in an enclosure that was not exposed to any sunlight. A typical Environment A would be described as a well-insulated concrete pump house with an access door. Due to the insulation, the ambient temperature fluctuation within the enclosure was moderate. Environment B would have the pump and outlet pipes protected from direct sunlight and precipitation by means of a four-walled, wooden or steel enclosure. Access to the equipment was provided via a removable roof. The shape and size of the enclosure is similar to that of a typical medium-sized doghouse. Environment B was found to be relatively similar to Environment A in terms of temperature fluctuation within the enclosure. In Environment C, the pump and outlet pipes were exposed to direct sunlight and therefore experienced more notable ambient temperature changes compared to the other two environments. The sample group had two GAPs located in an Environment A, six GAPs in an Environment B and two GAPs in an Environment C.

Flow intensity measurements

The intensity (flow rate) was determined at each GAP, using on-site volumetric measurements. The measurement entailed filling a container with water at the endpoint of the irrigation pipe. The container was filled for 45 s and subsequently weighed. The container was weighed pre- and post-fill and the flow rate was calculated. The manageable sample size allowed for a sufficient number of volumetric measurements. Each measurement was repeated 10 times at each GAP, resulting in 100 flow rate measurements. The measurements were used to create a distribution graph, representing the flow intensities for the study site. This method was easily executed, cost effective and caused little disturbance to the residents.

Consumer surveys

Surveys have been used in the past as an indication of indoor (Blokker et al., 2010) and outdoor water use (Roberts, 2005;

Table 1. Measurement methods for water end-use data collection

Measurement method/ device	Gathered information	Advantages	Disadvantages	Literature	Applicability to this study*
Consumer Surveys	Any information (within ethical constraints)	Flexibility, relatively simple to implement	Lower accuracy, ethical restrictions, post- processing of data required	Roberts, 2005; Colvin and Saayman, 2007	1
Temperature recorders (iButtons)	Time stamp, temperature	Non-intrusive, relatively low cost, no plumbing changes needed	Post-processing of data required	Chapmin et al., 2014; Massuel et al., 2009	1
Mechanical water meter (no logger)	Consumption, meter reading data	Accuracy	Manual readings, plumbing changes needed, relatively expensive	Turrall et al., 2005	2
Smart water meter with data logger	Flow rate, pressure, time stamp	Accuracy, automated readings	High cost, plumbing changes needed	Ngunyen et al., 2013	2
Watt-hour meter (electrical)	Time stamp, pump power, on-off state	Non-intrusive, no plumbing required	High cost, electrical changes needed	Massuel et al., 2009	2

*Note: (1) feasible and implemented; (2) considered, but not-used

Veck and Bill, 2000). A site survey was conducted as part of this study to obtain relevant information regarding water-use activities, including identification of the irrigation method (hand-held hose, manual sprinkler, automatic sprinkler), system connectivity, pump placement environment and water leakage. Although the method of irrigation was documented in the survey, it was not incorporated into the end-use model due to the limited sample size. The site surveys were also used to confirm that the residents used the irrigation systems at the maximum flow rate in each case. The pump flow rate was assumed equal to the garden irrigation flow rate in each case, with no leakage reported at any site.

Identifying pumping events and durations

Adopting terminology from Jacobs and Haarhoff (2004), the number of events over a given time period was described by using the term 'event frequency', expressed as the number of events per day. The term 'event duration' was used to describe the time lapse from an event start to event end. The recorded pipe wall temperature was analysed in order to identify pumping events and to extract the event frequency and event duration. The procedure was termed temperature variation analysis. Since temperature on each pipe was separately measured and analysed, there were no overlapping events. Each pumping event represented a single garden irrigation occurrence and was characterised by the pump start operation (water flowing through the pipe with corresponding temperature change) and the pump being turned off again. A Visual Basic macro, for implementation in MS Excel, was written to implement the temperature variation calculations. The baseline temperature, needed to identify significant interruptions in the expected graph pattern, was first established. Each interruption (difference between pipe wall temperature and baseline temperature) corresponded to a pumping event.

The daily ambient temperature fluctuated over the study period. The fluctuations varied per installation, because each iButton was placed in a different environment. Consequently, the baseline temperature at each GAP varied. The developed baseline temperature time-series graph at each GAP represented the typical daily temperature cycle per installation. The coefficient of determination (R^2) was used as a measure of similarity in shape between the baseline temperature at each GAP, and the temperature measured on the pipe wall. Thus each GAP had a specific baseline temperature corresponding to the particular environment and the ambient temperature of

the specific day. After the baseline temperature was developed, pumping events and durations were identified. Figure 1 shows an example of the pipe wall temperature measured by an iButton, and the corresponding baseline temperature curve, for one property over a 2-d period. The selected time series shows two events. A pumping event is noticed at about 06:00 on both days.

During time steps where the measured pipe wall temperature and baseline temperature deviated notably, water was likely flowing through the pipe (evidence of an event). Firstly, the temperature noise in each environment had to be separated from notable temperature deviations. Figure 2(a) shows the difference between the derived baseline temperature and the pipe wall temperature. Temperature noise is clearly visible around the zero y -axis value. In order to automatically detect pumping events, a conditional filter, incorporating a threshold temperature, was applied. The threshold value was determined with consideration for the different environments in which the iButtons were placed, being informed by earlier studies. Massuel et al. (2009) used a threshold of 2.6°C to detect pumping events as part of a study in India. In this study, the threshold was set equal to 2.0°C for Environment A and Environment B and to 3.0°C for Environment C. Implementing the threshold allowed for pumping events to be identified with an algorithm, which is significantly less time consuming than manual interpretation of the recorded data. With reference to the temperature noise visible in Fig. 2(a), all values not exceeding the threshold temperature were set equal to 0 and the result is plotted in Fig. 2(b). Figure 2(b) shows the two individual events in the selected time series, excluding temperature noise below the selected threshold values.

All recorded data were analysed in this manner by employing the algorithm. In total, 68 individual events were identified on 59 test days, considering the full data set of 110 test days. Multiple irrigation events per day were detected at only one home. The highest event frequency was 3 events per day, reported only once; 2 events per day were reported 7 times in the full data set at the same home. The limited number of events reported on in this study also allowed for subsequent visual inspection of the temperature difference at each event.

Basic model structure

A rudimentary model was developed to stochastically determine the average daily volume of groundwater pumped for garden irrigation. The model included three independent

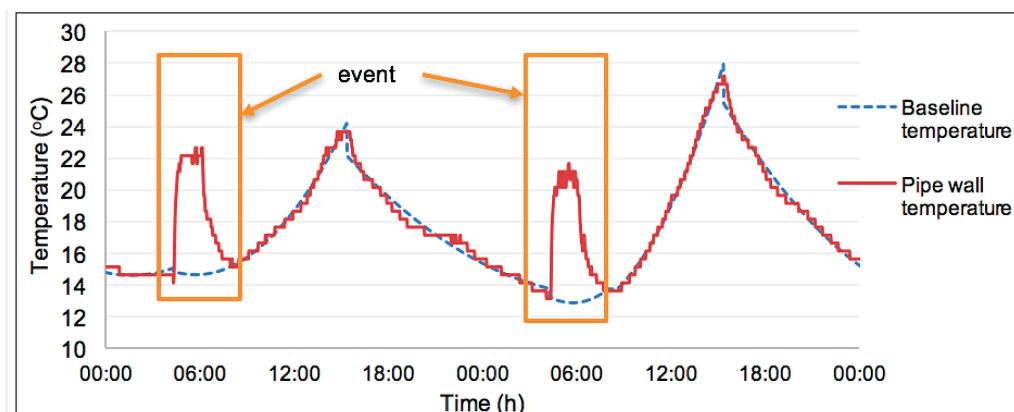


Figure 1. Measured pipe wall and derived baseline temperatures

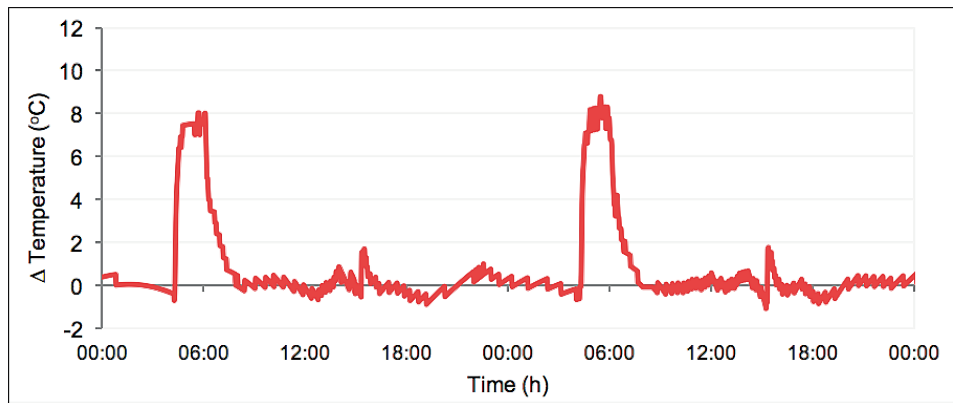


Figure 2 (a). Temperature difference between pipe wall and baseline temperatures

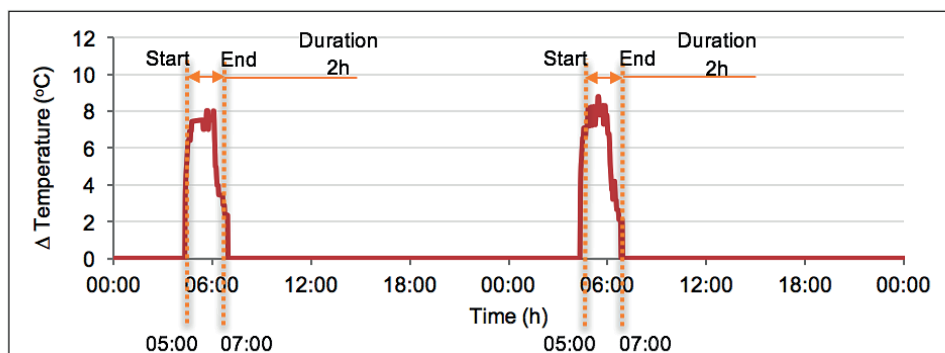


Figure 2 (b). Filtered temperature differences for pumping duration

parameters (duration, frequency, intensity) and was termed the DFI model. The DFI model adopted notation from the SIMDEUM model developed by Blokker et al. (2010), and is based on the assumption that all the input parameters are independent and statistically distributed random variables.

The DFI model structure, for a single residential property, is described by Eq. 1:

$$V_p = D_p \times F_p \times I_p \quad (1)$$

where,

V = average daily garden irrigation event volume (m^3/d)

D = event duration (h/event)

F = event frequency (events/d)

I = flow intensity (flow rate) at GAP (m^3/h)

The subscript p represents the best-fit probability distribution type of the respective variables in the DFI model. The procedure of setting up a stochastic model with the known distributions for each variable involved an evaluation of each model input parameter in terms of suitable statistical distributions.

Stochastic description of parameters

The best-fit statistical distributions for D , F , and I , were selected by implementing goodness-of-fit (GOF) tests to the measured data, using @Risk software. The GOF tests included the Anderson-Darling, Kolmogorov-Smirnov and Chi-Squared statistic. The best-fit distribution was chosen based on a combined scoring system of the GOF tests, similar to the selection tool developed by Masereka et al. (2015). The fitted

statistical distributions used in the DFI model are presented in Table 2, along with the corresponding mathematical descriptions and parameters.

RESULTS

Experimental field test results

Results discussed in this section were obtained from temperature variation analysis and volumetric measurements. A total of 68 irrigation events were identified over the 11-d study period by means of the temperature variation analysis. The average garden irrigation event duration was 2 h 16 min, with a relatively large standard deviation of 1 h 17 min. The longest irrigation event measured was 6 h and 59 min, and the shortest was 22 min. Some events were found to be relatively long in comparison with garden irrigation events reported elsewhere. The consumers of this study sample confirmed that, in some cases, the GAP would be operated until the aquifer was (temporarily) depleted and the event had to be terminated to allow for recharge, resulting in relatively long events.

The probability of an irrigation event occurring on a specific day during the 110 test days in the sample was 54%, meaning that consumers irrigated roughly every second day, on average. The flow intensities at the GAPs ranged between $1.14 \text{ m}^3/\text{h}$ and $1.25 \text{ m}^3/\text{h}$, with a most likely value of $1.16 \text{ m}^3/\text{h}$. These values were considered to be typical for groundwater abstraction at the household scale in South Africa. Local borehole contractors often use a thumb rule of 1 L/s ($3.6 \text{ m}^3/\text{h}$) as a relatively good flow rate from a garden borehole pump. Tennick (2000) reported that garden borehole flow rates in Hermanus, South Africa, ranged between $1 \text{ m}^3/\text{h}$ and $2 \text{ m}^3/\text{h}$.

Table 2. Statistical distribution descriptions

Distribution	Probability distribution function, $f(x)$ Cumulative distribution function, $F(x)$	Parameters
Log-logistic LL(β, α) Continuous	$f(x) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{(1 + (x/\alpha)^\beta)^2}$ $F(x) = \frac{1}{1 + (x/\alpha)^{-\beta}}$	$\alpha > 0$ (scale) $\beta > 0$ (shape) $x \geq 0$
Lognormal LN(μ, σ) Continuous	$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$ $F(x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left[\frac{\ln x - \mu}{\sigma \sqrt{2}} \right]$	$\mu > 0$ $\sigma > 0$ μ = mean σ = standard deviation
PERT PERT(a, m, b) Continuous	$f(x) = \frac{1}{B(\alpha_1, \alpha_2)} \frac{(x-a)^{\alpha_1-1} (b-x)^{\alpha_2-1}}{(b-a)^{\alpha_1+\alpha_2-1}}$ $B(\alpha_1, \alpha_2) = \text{Beta function}$ $\alpha_1 = \frac{4m + b - 5a}{b - a}$ $\alpha_2 = \frac{5b + a - 4m}{b - a}$ $F(x) = \frac{B_z(\alpha_1, \alpha_2)}{B(\alpha_1, \alpha_2)} = I_z(\alpha_1, \alpha_2)$ $z = \frac{x - a}{b - a}$ $B_z(\alpha_1, \alpha_2) = \text{incomplete Beta function}$	$b > a$ (boundary) $b \geq m \geq a$ a = minimum b = maximum m = most likely value
Binomial B(n, p) Discrete	$f(x) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{(1 + (x/\alpha)^\beta)^2}$ $F(x) = \frac{1}{1 + (x/\alpha)^{-\beta}}$	$n > 0$ $0 < p < 1$ n = count p = success probability

Naidoo and Burger (2017) also reported on groundwater abstraction in South Africa. The average pump flow rate was found to range between 0.36 m³/h and 2.7 m³/h (Naidoo and Burger, 2017). Flow intensity values from this study were thus within the range reported earlier.

Stochastic results

Event duration

Garden irrigation event duration D was determined by means of the temperature variation analysis procedure described earlier. Many factors may contribute to the duration of irrigation, including the method of irrigation, property size, rainfall, aquifer yield, ambient temperature and time of day. These factors were not considered in the DFI end-use model; event duration was modelled as an independent variable. If multiple events occurred on the same day, each event duration was analysed separately.

A cumulative distribution function (CDF) of the measured duration is presented in Fig. 3, along with the CDF of the stochastic distribution with the best fit. The loglogistic distribution provided the best fit, slightly outperforming the lognormal distribution that ranked second in terms of fit. However, the parameters of the lognormal distribution (mean and standard deviation) are more readily available than the

shape and scale parameters of the log-logistic distribution. Therefore, the lognormal distribution was selected to model the irrigation duration variable, thus simplifying practical application in the future.

Event frequency

Event frequency F is often described using a discrete statistical distribution (Blokker et al., 2010) and is typically expressed as a

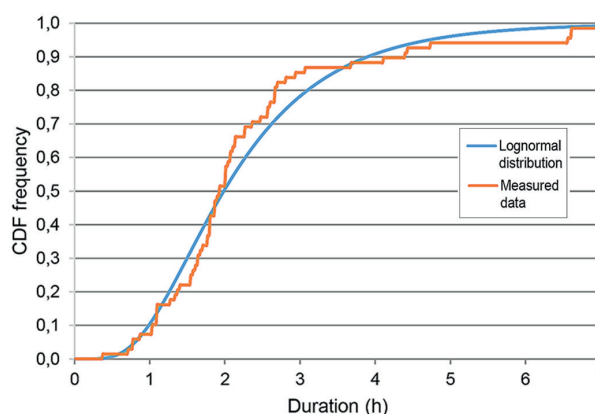


Figure 3. Cumulative distribution function for garden irrigation duration

Poisson distribution, in which case only one parameter is needed (λ = average) to populate the distribution. The event frequency was modelled as the probability of a garden irrigation event occurring on a specific day, with a maximum of 1 pumping event per day. Consequently, the binomial distribution was used to describe event frequency over the 11-d study period. Figure 4 shows the CDF of the measured irrigation frequencies with the fitted binomial distribution curve.

No distribution fitted the measured data well, partly because of the small sample size and relatively sustained consumer habits and/or the use of automated programmed irrigation timers. Many different events from a particular home would thus report the same frequency and would be lumped in the CDF. The significant difference between event irrigation frequencies could also be ascribed in part to consumer behaviour and also to the different types of irrigation systems used at the study homes. The irrigation method was, however, not included as independent variable in the DFI model. Additionally, no rain days occurred during the study period. The binomial distribution provided the best fit to the data and was considered adequate to illustrate application of the model, with appreciation that future research in this regard is needed.

Flow intensity

The site survey confirmed that all GAPS were operated at full capacity while irrigating. Thus the flow intensity I at each GAP was measured at the maximum flow rate. A CDF, containing 100 data points (10 measurements at each of the 10 residential properties) was plotted in Fig. 5. The Beta-Program Evaluation and Review Technique (PERT) distribution, identified as the best fit to the actual data, was superimposed on the actual data. The PERT distribution incorporates three parameters: the minimum, maximum, and the most likely value.

Application of DFI model to study site

Statistical distributions were fitted to measurements obtained from the iButtons and volumetric measurements. Equation 1 was modified to include the identified best-fit distributions for each model variable. Equation 2 represents the stochastic DFI model:

$$V(LN\sim\mu,\sigma) = D(LN\sim\mu,\sigma) \times F(B\sim n,p) \times I(PERT\sim a,m,b) \quad (2)$$

Table 3 summarises the variables of the DFI model, as well as the parameter values of the specific study site. The parameter values in Table 3 represent garden irrigation in autumn for the specific Cape Town study site.

The DFI model was implemented on the study site by populating Eq. 2 with the values presented in Table 3. A total of 1 000 000 iterations were simulated using the Monte Carlo method to stochastically determine the average daily volume (in m^3/d) of groundwater pumped for garden irrigation. The CDF of the average daily garden irrigation event volume supplied from GAPS at the study site is shown in Fig. 6. A comparison of the DFI model's stochastic results (based on GOF tests) and the study site measurements is presented in Fig. 6.

The results presented in Fig. 6 relate to the study site over the study period (April/May) and should not be generalised. The average daily groundwater abstraction for garden irrigation could simply be calculated by multiplying the average values of D , F , and I . The stochastic results also show

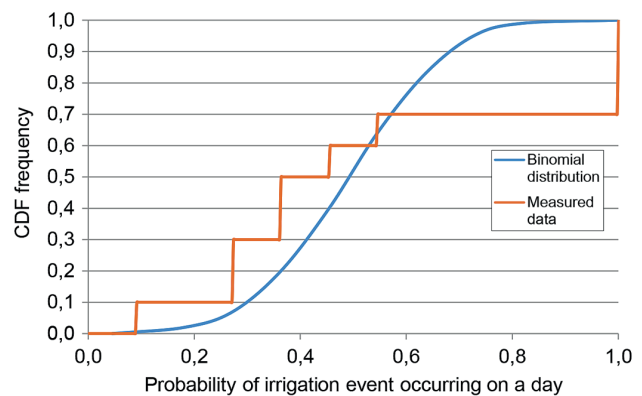


Figure 4. CDF for garden irrigation event frequency

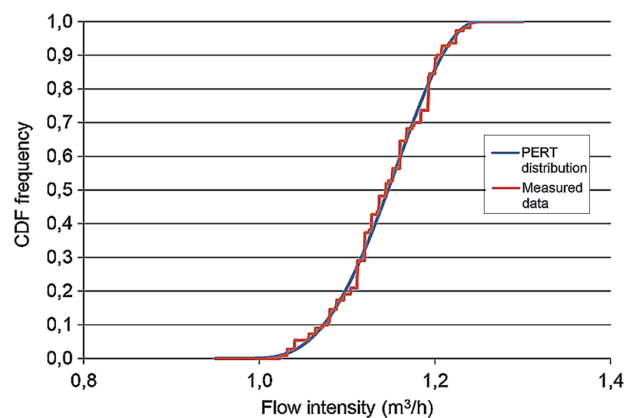


Figure 5. Cumulative distribution function for groundwater flow intensity

Table 3. DFI model input parameters for study site in autumn

Variable	Average (μ)	Distribution	Parameter	Parameter value
Duration (h)	2.273	Lognormal	μ	2.372
			σ	1.289
Frequency (events/d)	0.536	Binomial	n	11.000
			p	0.536
Intensity (m^3/h)	1.143	PERT	a	0.978
			m	1.157
			b	1.252

that a daily average of $1.39 m^3/d$ is used for garden irrigation, as would be expected. Due to the relatively large variation in garden irrigation volume, from one home to the next, one region to the other and by season, the average value alone does not provide sufficient insight. The stochastic results provide more detail. An additional sensitivity analysis was conducted in order to explain the relative contribution of different parameter values. The sensitivity analysis showed that garden irrigation volume was the most sensitive to event duration. The significant contribution of event duration in the model is explained by the notable parameter variability coupled to a relatively wide range in event duration amongst residents.

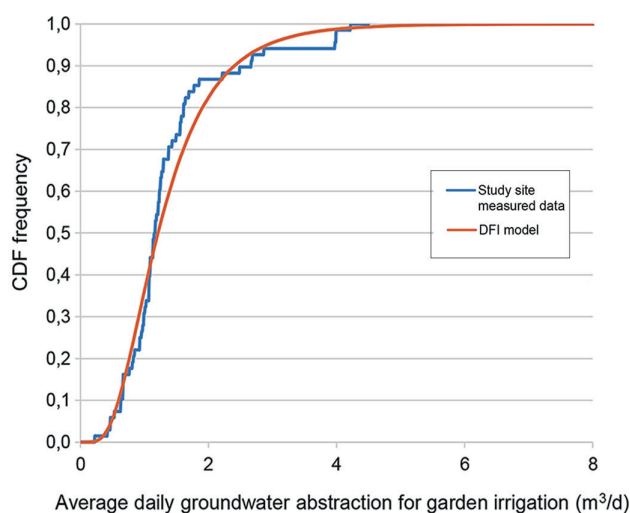


Figure 6. Cumulative distribution function of the average daily garden irrigation event volume

DISCUSSION

Utilising iButtons as indirect method for measuring water usage at privately owned GAPs proved useful. The method was simple, cost effective and caused relatively little disturbance to the homeowners. The average pumping event duration at the study site was 2 h and 16 min, with the shortest event being 22 min. The recording interval of 2 min ensured that irrigation events could successfully be identified, because event duration significantly exceeded the recording interval. Expanding the application of iButtons to include different household end-use components, such as the bath, shower, washing machine and dishwasher could be explored. However, iButtons would be unable to detect events with a relative short duration (less than 2 min), such as basin taps and toilet flushing, or events where the temperature variation is expected to be small. The temperature variation method has been applied to hot water end-uses, such as the shower (Botha et al., 2017), where temperature variation is expected to be relatively large.

The average irrigation duration and frequency measured in this study are higher than values reported by Roberts (2005). This research project focused on groundwater as supplementary water source, meaning that an unrestricted irrigation scenario was considered. Consequently, it could be expected that residents with GAPs (this study) would irrigate more regularly and for longer durations compared to a sample group of residents using water from the potable water distribution system for garden irrigation.

The DFI model can serve as a useful, rudimentary means to investigate garden irrigation by researchers and utility managers. Based on the relatively small sample of 59 measurement points, a probability distribution function (PDF) cannot be defined with sufficient representativeness for longer time periods, or other regions. The combination of literature values and the 59 data points was used in this study to compile PDFs as a means to illustrate the method and obtain results from the study sample. The results are not representative of a larger region, or consumers beyond the study site. However, the DFI model is scalable over different study sites, as the parameters of the distribution curves could be populated with values corresponding to another region, or time, as applicable.

The DFI model could be expanded in the future to incorporate seasonal variability, different irrigation methods and also other types of supplementary household water supply, such as rainwater and greywater.

CONCLUSION

Unique garden irrigation events from groundwater abstraction points were identified by means of temperature variation analysis in the Cape Town case study site. A relatively high garden irrigation event occurrence was observed at all 10 homes and the recorded duration of the 68 detected events was relatively long. The DFI model was based on data measured in the Cape Town study site and was subsequently used to illustrate stochastic modelling of garden irrigation. The temperature variation analysis could be employed elsewhere to populate the DFI model with values for event duration, frequency and intensity (flow rate) in order to assess garden irrigation from groundwater abstraction points in other regions.

ACKNOWLEDGEMENTS

The authors would like to thank the National Research Foundation (NRF) of South Africa and GLS Consulting for financial support towards a post-graduate student bursary.

REFERENCES

- ARBON N, THYER M, HATTON MACDONALD D, BEVERLEY K and LAMBERT M (2014) Understanding and predicting household water use for Adelaide. Technical Report Series No. 14/15. Goyder Institute for Water Research, Adelaide, South Australia.
- BEAL C, STEWART RA, HUANG T and REY E (2011) South East Queensland residential end use study. *J. Aust. Water Assoc.* **38** (1) 80–84.
- BLOKKER EJM, VREEBURG JHG and VAN DIJK JC (2010) Simulating residential water demand with a stochastic end-use model. *J. Water Resour. Plann. Manage.* **136** 19–26. [https://doi.org/10.1061/\(asce\)wr.1943-5452.0000002](https://doi.org/10.1061/(asce)wr.1943-5452.0000002)
- BOTHA BE, JACOBS HE, BIGGS B and ILEMOBADE AA (2017) Analysis of shower water use and temperature at a South African university campus. *CCWI 2017 – Computing and Control for the Water Industry*, 5–7 September 2017, Sheffield, United Kingdom. <https://doi.org/10.15131/shef.data.5364562.v1>
- BRITTON T, COLE G, STEWART R and WISKAR D (2008) Remote diagnosis of leakage in residential households. *J. Aust. Water Assoc.* **35** 89–93.
- BUCHBERGER SG and WELLS GJ (1996) Intensity, duration, and frequency of residential water demands. *J. Water Resour. Plann. Manage.* **112** (1) 11–19. [https://doi.org/10.1061/\(asce\)0733-9496\(1996\)122:1\(11\)](https://doi.org/10.1061/(asce)0733-9496(1996)122:1(11))
- BUCHBERGER SG, CARTER JT, LEE YH, and SCHADE TG (2003) Random demands, travel times, and water quality in dead ends. Rep. No. 90963F. AWWARF, Denver.
- BUTLER JJ (1991) A stochastic analysis of pumping tests in laterally nonuniform media. *Water Resour. Res.* **27** (9) 2401–2414. <https://doi.org/10.1029/91wr01371>
- CHAPMIN T, TODD AS and ZEIGLER MP (2014) Robust, low-cost data loggers for stream temperature, flow intermittency, and relative conductivity monitoring. *Water Resour. Res.* **50** (8) 6542–6548. <https://doi.org/10.1002/2013wr015158>
- COLVIN C and SAAYMAN I (2007) Challenges to groundwater governance: a case study of groundwater governance in Cape Town, South Africa. *Water Polic.* **9** (2) 127–149. <https://doi.org/10.2166/wp.2007.129>
- DEOREO WB, HEANEY JP and MAYER PW (1996) Flow trace analysis to assess water use. *J. Am. Waste Water Assoc.* **88** (1) 79–90.
- DEOREO WB, DIETERMANN A, SKEEL T, MAYER PW, LEWIS DM

- and SMITH J (2001) Retrofit realities. *J. Am. Water Works Assoc.* **93** (3) 58–72.
- DU PLESSIS JL and JACOBS HE (2015) Procedure to derive parameters of stochastic modelling of outdoor water use in residential estates. *Proced. Eng.* **119** 803–812. <https://doi.org/10.1016/j.proeng.2015.08.942>
- DU PLESSIS JL, FAASEN B, JACOBS HE, KNOX AJ and LOUBSER C (2018) Investigating wastewater flow from a gated community to disaggregate indoor and outdoor water use. *J. Water Sanit. Hyg. Dev.* **8** (2) 238–245. <https://doi.org/10.2166/washdev.2018.125>
- FISHER-JEFFES L, GERTSE G and ARMITAGE NP (2015) Mitigating the impact of swimming pools on domestic water demand. *Water SA* **41** (2) 238–246. <https://doi.org/10.4314/wsa.v41i2.9>
- GHAVIDELFAR S, SHAMSELDIN AY and MELVILLE BW (2018) Evaluating spatial and seasonal determinants of residential water demand across different housing types through data integration. *Water Int.* **43** (7) 926–942. <https://doi.org/10.1080/02508060.2018.1490878>
- HEMATI A, RIPPY M, GRANT S, DAVIS K and FELDMAN D (2016) Deconstructing demand: the anthropogenic and climatic drivers of urban water consumption. *Environ. Sci. Technol.* **50** (23) 12557–12566. <https://doi.org/10.1021/acs.est.6b02938>
- HEINRICH M (2007) Water end use and efficiency project (WEEP) - Final report. BRANZ Study report 159. BRANZ, Judgeford, New Zealand.
- HOWE CW and LINAWEAVER FP (1967) The impact of price on residential water demand and its relation to system design and price structure. *Water Resour. Res.* **3** (1) 13–32. <https://doi.org/10.1029/wr003i001p00013>
- JACOBS HE (2007) The first reported correlation between end-use estimates of residential water demand and measured use in South Africa. *Water SA* **33** (4) 549–558.
- JACOBS HE and HAARHOFF J (2007) Prioritisation of parameters influencing residential water use and wastewater flow. *J. Water Supply: Res. Technol.* **56** (8) 495–514. <https://doi.org/10.2166/aqua.2007.068>
- JACOBS HE AND HAARHOFF J (2004) Structure and data requirements of an end-use model or residential water demand and return flow. *Water SA* **30** (3) 293–304. <https://doi.org/10.4314/wsa.v30i3.5077>
- JACOBS HE, GEUSTYN LC, FAIR K, DANIELS J and DU PLESSIS JA (2007) Analysis of water savings: a case study during the 2004/2005 water restrictions in Cape Town. *J. S. Afr. Inst. Civ. Eng.* **49** (3) 16–26.
- LOH M and COGHLAN P (2003) Domestic water use study in Perth, Western Australia (1998–2001). March 2003. Water Corporation, Perth, Australia. ISBN 1 74043 1235
- LUGOMA MFT, VAN ZYL JE and ILEMOBADE AA (2012) The extent of on-site leakage in selected suburbs of Johannesburg. *Water SA* **38** (1) 127–132. <https://doi.org/10.4314/wsa.v38i1.15>
- MACDONALD AM and CALOW RC (2009) Developing groundwater for secure water supplies in Africa. *Desalination* **248** 546–556. <https://doi.org/10.1016/j.desal.2008.05.100>
- MAKWIZA C and JACOBS HE (2017) Sound recording to characterize outdoor tap water use events. *J. Water Supply: Res. Technol.* **66** (6) 392–402. <https://doi.org/10.2166/aqua.2017.120>
- MASEREKA EM, OTIENO FAO, OCHIENG GM and SNYMAN J (2015) Best fit selection of probability distribution models for frequency analysis of extreme mean annual rainfall events. *Int. J. Eng. Res. Dev.* **11** (4) 34–53.
- MASSUEL S, PERRIN J, WAJID M, MASCRE C and DEWANDEL B (2009) A simple low-cost method to monitor duration of groundwater pumping. *Groundwater* **41** (1) 141–145. <https://doi.org/10.1111/j.1745-6584.2008.00511.x>
- MEAD N and ARAVINTHAN V (2009) Investigation of household water consumption using smart metering system. *Desalination Water Treat. J.* **11** (1–3) 115–123. <https://doi.org/10.5004/dwt.2009.850>
- NAIDOO V and BURGER M (2017) Groundwater abstraction assessment for development on the farm Vlakfontein 523 JR Portion 25. Report compiled for JCJ Developments. Geo Pollution Technologies Report No. ICEAS-17-2246. JCJ Developments, Gauteng, South Africa.
- NEL N, JACOBS HE, LOUBSER C and DU PLESSIS JA (2017) Supplementary household water sources to augment potable municipal supply. *Water SA* **43** (4) 553–562. <https://doi.org/10.4314/wsa.v43i4.03>
- PARKER JM and WILBY RL (2013) Quantifying household water demand: A review of theory and practice in the UK. *Water Resour. Manage.* **27** (4) 981–1011. <https://doi.org/10.1007/s11269-012-0190-2>
- RATHNAYAKA K (2015) Seasonal demand dynamics of residential water end-uses. *Water J.* **7** 202–216. <https://doi.org/10.3390/w7010202>
- ROBERTS P (2005) Yarra Valley Water 2004 residential end use measurement study. Final report, June 2005. Yarra Valley Water. Melbourne, Australia.
- RUNFOLA DM, POLSKY C, NICOLSON C, GINER NM, PONTIUS RG, KRAHE J and DECATUR A (2013) A growing concern? Examining the influence of lawn size on residential water use in suburban Boston, MA, USA. *Landscape Urban Plann. J.* **119** 113–123. <https://doi.org/10.1016/j.landurbplan.2013.07.006>
- SCHIEPERS HM and JACOBS HE (2014) Simulating residential indoor water demand by means of a probability based end-use model. *J. Water Supply: Res. Technol.* **63** (6) 476–488. <https://doi.org/10.2166/aqua.2014.100>
- TENNICK FP (2000) Preliminary report on hydrocensus of groundwater resources in the Eastcliff and Northcliff suburbs, Hermanus, Western Cape, South Africa. Compiled by FPT Consulting Services for the Overstrand Municipality, May 2000.
- TURRAL HN, ETCHELLES T, MALANO HMM, WIJEDASA HA, TAYLOR P, MACMAHON TAM and AUSTIN N (2005) Water trading at the margin: The evolution of water markets in the Murray-Darling Basin. *Water Resour. Res.* **41** W07011. <https://doi.org/10.1029/2004wr003463>
- VECK GA and BILL MR (2000) Estimation of the residential price elasticity of demand for water by means of a contingent evaluation approach. WRC Report No. 790/1/00. Water Research Commission, Pretoria.
- WASOWSKI A (2001) Changing the American landscape. *J. Am. Wastewater Assoc.* **3** 40–44.
- WILLIS RM, STEWART RA, PANUWATWANICH K, WILLIAMS PR and HOLLINGSWORTH AL (2011) Quantifying the influence of environmental and water conservation attitudes on household end use water consumption. *J. Environ. Manage.* **92** (8) 1996–2009. <https://doi.org/10.1016/j.jenvman.2011.03.023>
- WRIGHT T and JACOBS HE (2016) Potable water use of residential consumers in the Cape Town metropolitan area with access to groundwater as supplementary household water source. *Water SA* **42** (1) 144–150. <https://doi.org/10.4314/wsa.v42i1.14>

Do rope and washer pumps provide safe water and satisfied users? A case study piloting new rural water supply technology in Rumphi District, Malawi

Jacob C Mkandawire¹, Mavuto Tembo², Muthi Nhlema³, Joel Luhanga⁴ and Rochelle H Holm^{5*}

¹Department of Agri-Sciences, Mzuzu University, P/Bag 201, Mzuzu 2, Malawi and District Water Office, Rumphi District Council, Rumphi, Malawi

²Centre of Excellence in Water and Sanitation, Mzuzu University, P/Bag 201, Mzuzu 2, Malawi

³BaseFlow, 2nd Floor, Galaxy House, Blantyre, Malawi

⁴Department of Forestry and Environmental Management, Mzuzu University, P/Bag 201, Mzuzu 2, Malawi

⁵Centre of Excellence in Water and Sanitation, Mzuzu University, P/Bag 201, Mzuzu 2, Malawi

ABSTRACT

Innovation is needed to develop rural water supply to support sub-Saharan Africa communities that are hard to reach. The purpose of this study was to critically review rope and washer pumps that have been installed on manually drilled boreholes in 48 communities as part of a pilot project in Rumphi District, Malawi, and which serve as a sustainable source of drinking water from both technical (water quality and functionality) and social (user satisfaction) perspectives. At each water source, an infrastructure checklist was used ($n = 48$); 10 users were interviewed ($n = 472$); and, if the pump had water, water quality samples were collected ($n = 24$). The results show that use of a professional driller does not guarantee a functioning rope and washer pump that produces safe water. Where the pumps were functional, most provided safe drinking water. However, only 8% (4/48) of pumps had good water quality, a flow rate of >20 L/min and a full consensus of positive satisfaction among users. Pumps are not necessarily working better or worse in more remote areas. A process of introducing and creating evaluative guidelines for new (approved) technologies for rural water supply has not been established in Malawi. Sub-Saharan African governments need to be open to innovative solutions while making sure that standards, including those for functionality, water quality, user satisfaction, private operators, and human capacity for local government regulators, are being followed to ensure safe water for rural communities.

Keywords: water, Malawi, rope and washer pump, groundwater, rural

INTRODUCTION

The Government of Malawi (2016) has approved only three handpump designs for rural water supply: (i) Afridev for deep wells at the community level, (ii) Malda for shallow water tables, and (iii) Climax, for deep water tables or institutional boreholes. The Government has been slow to respond to requests to introduce new technology, and there is no clear process for introducing or creating guidelines that evaluate new technologies for rural water supply in Malawi.

In Malawi, the operation and maintenance of rural water supply is decentralized and is covered under a community-based management model, i.e., in the hands of the user community. The user communities pay for the full costs of ongoing pump operation and maintenance, which include services from area mechanics and the purchase of spare parts in consultation with the local government and the private sector (Malawi Government, 2005; 2015). Chowns (2015) argues that, in regard to rural water supply in Malawi, this system of user management and financing of maintenance does not work well for rural communities when it is used to transfer the responsibilities of providing and maintaining public services from the local government to communities.

Recently, rope and washer pumps have been the subject of a relatively small number of studies (Fig. 1) in sub-Saharan Africa (Harvey and Drouin, 2006; Coloru et al., 2012; Butterworth et al., 2013; Holm et al., 2017). Water and public health professionals are trying to understand and manage the rural water supply for drinking water with new approaches

to supply small communities and individual families. Sutton (2017) notes that the least-served rural communities tend to be those with less than 150 people, and the lower the number of people a handpump serves, the higher per capita life cycle cost there is for the rural water supply.

When innovative rural water supply technologies are introduced in sub-Saharan Africa, they are rarely evaluated and the reasons for dis-adoption have not been documented. However, service providers and other stakeholders can learn from those reasons to make more informed decisions the next time they provide solutions for rural and low-income communities. This study builds from the earlier work of Holm et al. (2017), which examined 19 rope and washer pumps in the Rumphi District, Malawi, which are also included in our study. The purpose of this study is to reflect critically on rope and washer pumps that have been installed on manually drilled boreholes in 48 communities as part of a pilot project to provide sustainable sources of drinking water in Rumphi District, Malawi, with a focus on both technical (water quality and functionality) and social (user satisfaction) perspectives. Findings from this work will identify the issues for best practices in sub-Saharan Africa to encourage national guidelines for improving rural groundwater supply management based on the regulation of low-cost alternative technologies to ensure safe drinking water for rural communities.

METHODS

Forty-eight communities were purposely selected from Rumphi District in the Traditional Authorities, a subdivision of the district, of: Chisovya, Kachulu, Mwahenga, Mwalweni, Mwamlowe, and Mwankhunikira. These communities were

*Corresponding author, email: rochelle@rochelleholm.com

Received 4 July 2018; accepted in revised form 2 July 2019



Figure 1. Rope and washer pump, Rumph District

generally in hard-to-reach areas and had participated in the selection of the rope and washer pump for their community as the technology was introduced under a supported development partner project (Holm et al., 2017). Manually drilled wells were completed between approximately April and October 2016. The drillers reported that the depth ranged from 9.8 to 21 m below ground surface. The pumps were installed at each well within a few months of completion. The wells in this study were within the same geographical area and were presumed to be accessing the same aquifer.

A transformative mixed methods approach was used (Creswell, 2014). The field study was conducted during a 3-month period between December 2017 and February 2018. The first site visits were conducted in the hot season just before the first rains in December and the second visits were conducted during the wet months of January and February 2018. Each pump was visited a minimum of 2 times. At each water source, an infrastructure checklist was used, 10 users were purposely interviewed, and water quality samples were collected.

The infrastructure checklist included determining whether water was available on the day of the visit and observing the visual condition of the pump and the surrounding environment. For the user survey, there was a preference for interviewing the available female users of the study pumps; community chiefs and water point committee members were also included. All interviews were conducted by a local government representative.

The user survey included questions on recent breakdowns,

maintenance, and user satisfaction (yes/no) with the pump. Survey data were collected on an android system using Open Data Kit software (Open Source, University of Washington, Seattle, Washington, USA).

When water was available at the pump, samples were analysed for thermotolerant coliforms, turbidity, pH, total hardness, total alkalinity, free chlorine, and total chlorine. Samples for thermotolerant coliforms were collected in Whirl-Pak plastic bags (Nasco, Fort Atkinson, Wisconsin, USA) and transported to a centrally located field laboratory within Rumph District. All thermotolerant coliform analyses were performed in duplicate within 8 h. The Wagtech Potatest Membrane Filtration Unit (Palintest USA, Erlanger, Kentucky) was used at 44°C for 18 h for the determination of thermotolerant coliforms. One equipment blank was analysed daily to monitor for field laboratory contamination by using boiled and cooled water, and each equipment blank recorded 0 colony-forming units (cfu)/100 mL. An analysis of the water samples was also performed in duplicate onsite at the pump with the Wagtech Potatest Pocket pH Sensor (Palintest USA, Erlanger, Kentucky), Jackson turbidity tubes, and Hach Aquachek 5-in-1 Water Quality Test Strips (Hach Company, Loveland, Colorado) which were used to analyse the total hardness, total alkalinity, free chlorine, and total chlorine. The samples were not filtered.

In addition, the local government records on the well depth, the driller's name, and the pump committee chair gender were compiled. Wells that were classified as the most rural either required boats to access the community (no passable road during the field study) or were several hours of travel by road from a major trading centre.

The study covered two climatic seasons, and from the first to the second visit it was becoming wetter. Wells were classified for functionality based on the first site visit. Pumps were categorized as 'functional' if any amount of water was available at the time of the first site visit, even if the pump had problems. Non-functional wells were not pumping any water on the first visit. A subset of the functional wells were classified as sustainable. Sustainability with regard to this paper is defined as when the pumps had water during both the first visit before rains started and second visit during rains, safe drinking water based on thermotolerant coliforms (Malawi Bureau of Standards, 2005), all users surveyed indicated that they were satisfied with the pump, and that flow rate was at least 20 L/min.

Water quality, functionality, sustainability, and the infrastructure checklist data were analysed by using the R Project 3.3.2 statistical package (Vienna, Austria) at $p < 0.05$. The water quality was compared to national and World Health Organization (WHO) guidelines for drinking water (Malawi Bureau of Standards, 2005; WHO, 2017).

Ethical clearance for this study was obtained from the Government of Malawi, National Commission for Science and Technology (Protocol P.10/17/221). Written consent was obtained from user respondents. All research tools and data are available from the corresponding author.

RESULTS

Water quality

The water quality from rope and washer pumps was compared with Malawi specifications for the raw groundwater that is used for drinking and WHO drinking water guidelines (Table 1) (Malawi Bureau of Standards, 2005; WHO, 2017). There were

only 24 wells where water could be pumped at the time of water sampling during the first visit. One pump that could not be located is detailed further in the 'Study limitations' section below. Most pumps (22/24) had water that was suitable for drinking based on the local standards for thermotolerant coliforms (Malawi Bureau of Standards, 2005), yet less than half of the pumps (11/24) provided water that is suitable for drinking based on the more stringent WHO (2017) guidelines for thermotolerant coliforms. Many pumps had clear water (e.g., 15/24 pumps had turbidity at <5 Jackson turbidity units [JTU]).

Some wells were not necessarily sited properly, with 7/48 pumps within 30 m of a pit latrine. Whether the pit latrine was constructed after well installation or whether this was due to initial improper siting by the driller could not be determined.

Two pumps (2/24) had high thermotolerant coliform levels (> 200 cfu/100 mL), which indicated that the water was not suitable for drinking, yet at least some (3 users) of the interviewed users of each of these two pumps indicated that the water was used by their household for drinking. Additionally, while both pumps had cloudy water, one pump was also linked to the highest turbidity (500 JTU) in this study. Each of these two pumps during the first visit had a latrine that was observed to be located higher in elevation than the pump and with an unclean site (trash or animal waste surrounding the pump). But, only one pump was within 30 m of a pit latrine. Both of these pumps had a surface slab that was in good condition (no cracks in the apron) and a functioning soak-away pit for excess water to be drained into. Whether sources of thermotolerant coliform contamination were already there before well installation or a result of user activities around the pump after construction could not be determined.

Free and total chlorine results were both 0 mg/L for all pumps (24/24), which indicated that none of the pumps had recently been disinfected by using sodium hypochlorite. The pH ranged from 5.9 to 7.3, with only one pump outside of the national standard range (Malawi Bureau of Standards, 2005). The total alkalinity ranged from 40 to 180 mg/L as CaCO₃, while all samples for total hardness were below the 800 mg/L CaCO₃ national standard (Malawi Bureau of Standards, 2005).

Functionality

In addition to indicating a functional pump providing water, 'functionality' is also a gauge to ascertain whether the right holders who use the pump are also maintaining it. Functionality of rope and washer pumps depends on being: (i) installed properly; (ii) have an operation and maintenance system; and (iii) have a groundwater management strategy. Installation problems might include: improper alignment of the pipes and the rope, which might lead to faster wear and tear of the rope even when replaced with a new rope, the rope getting stuck in the rising main pipe, or the well not being drilled deep or straight enough. The operation and maintenance management system includes determining who is responsible and who pays for worn-out parts, which need to be affordable and available from an accessible trained pump repairer. Groundwater management strategies come down to aquifer characteristics and an appropriate number of users for a resource, because water availability is affected when the rate of abstraction is greater than the rate of groundwater recharge. During the dry season, this might lead to the water table moving below the depth of the well, which results in the lack of available water, even though the pump might not have any technical, operational or maintenance problems.

The functionality results of our study indicated that nearly half of the pumps (23/48) were non-functional (i.e., there was no water available) on the first site visit (Fig. 2). The study did not specifically observe any groundwater management or regional recharge concerns that are linked to non-functionality. While for 6 pumps, including some of the pumps that were non-functional for the longest time, non-functionality was linked to installation problems from the beginning. Yet, non-functionality was more likely to be linked to operation and maintenance. During the course of the study, several pumps (13/48) changed status, either changing their status from having water to not or vice versa. For those pumps that were non-functional on the first site visit, the issue was commonly attributed to the rope, such as wearing out due to friction and breaking quickly due to borehole verticality (16/17) or the rope stuck within the pipe (but still intact) (1/17). No cases of rope theft were identified. Where there is a need for a new rope, the

Table 1. Water quality results of rope and washer pumps and comparison with Malawi and World Health Organization drinking water guidelines

	Parameter				
	Thermotolerant coliform (colony forming units/100 mL)	Turbidity	Total hardness (mg/L as CaCO ₃)	Total alkalinity (mg/L as CaCO ₃)	pH
Minimum (n = 24)	0	<5 JTU	25	40	5.9
Mean (n = 24)	20	40 JTU	120	80	6.5
Median (n = 24)	1	<5 JTU	120	80	6.4
Maximum (n = 24)	>200 ^c	500 JTU	250	180	7.3
Malawi Standard ^a	50	25 NTU	800	-	6.0 to 9.5
WHO Standard ^b	0	1 NTU	-	-	-

^aMalawi Bureau of Standards (MBS), 2005

^bWorld Health Organization (WHO), 2017

^cResult was too numerous to count, upper detection limit of method is reported.

^dJTU = Jackson turbidity units. NTU = Nephelometric turbidity units. The two units are roughly equivalent (WHO Fact Sheet 2.33 – Turbidity measurement; http://www.who.int/water_sanitation_health/hygiene/emergencies/fs2_33.pdf)

- No established value

users would generally need to initiate the purchase of a new rope (a new rope was not observed to be kept on hand by the community), placement of washers and installation by a hired technician. The old rope would be kept by the committee. Ropes are only available to purchase at larger trading centres in the study area, not within each community. Whereas when the rope was stuck in the pipe this only required a minor adjustment by the user to put the rope in its normal position; however, until this was addressed it would result in no water flow. In Tanzania, Coloru et al. (2012) similarly found that the most common maintenance undertaken on the rope and washer pump was a replacement of the rope, as this component was most subject to wear and tear. Reasons for the non-functionality of the 23 rope

and washer pumps are as shown in Table 2.

Over half of the functional pumps (14/25) provided at least a 20 L/min flow rate, while 3 pumps were providing <10 L/min of water. Most pumps (47/48) had bolts and nuts that were visually observed to be in good condition. Of the 48 study pumps, only two were observed to have both problems with the surface slab condition (such as cracks in the apron) and a non-functioning soak away pit; both these pumps were non-functional on the first site visit.

Twelve individual professionally trained drillers from within the district each installed 1 to 13 pumps of the 48 pumps. Although the sample size is small, the following trends can still be considered: 5 of the drillers had all non-functional

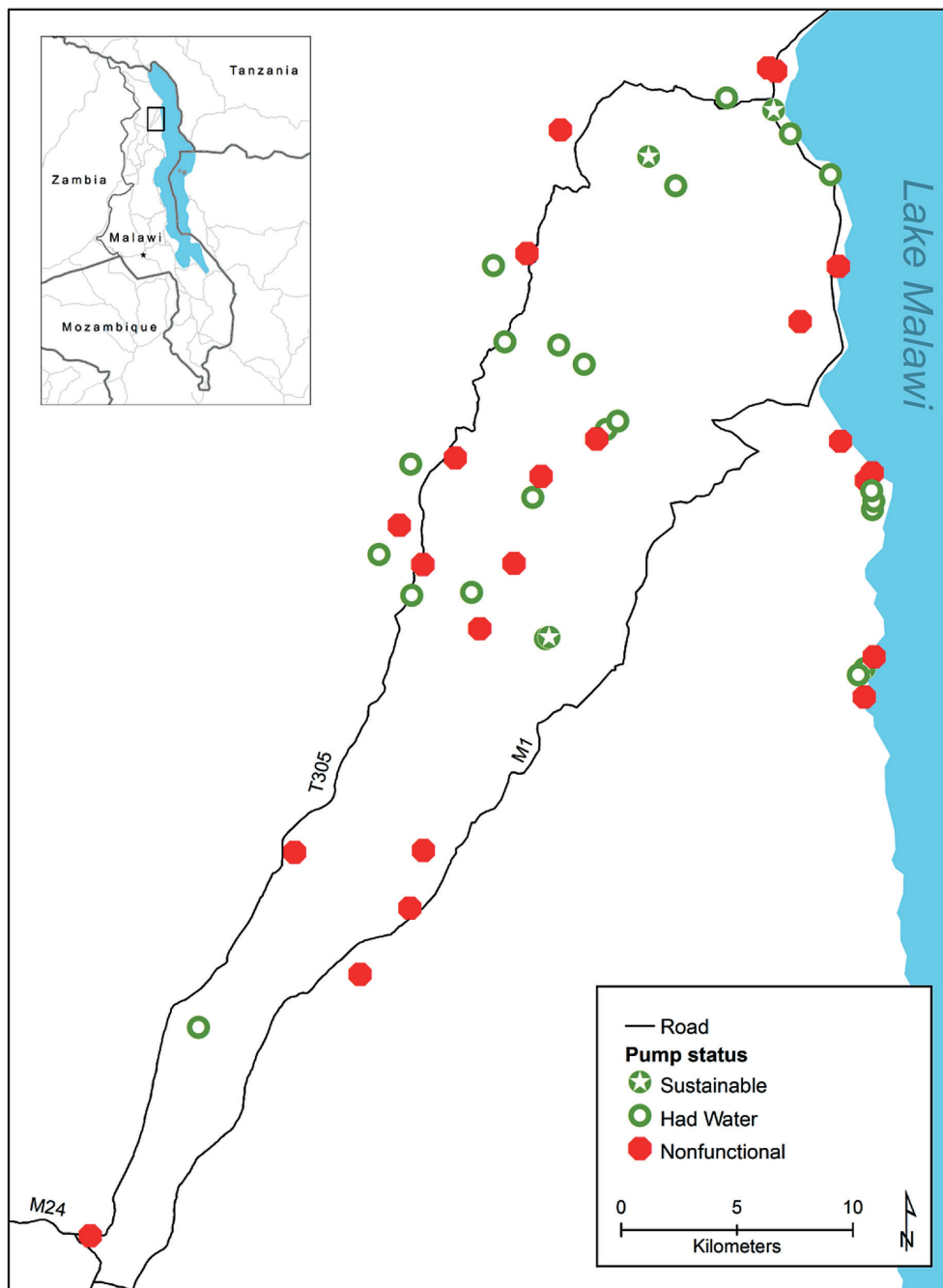


Figure 2. Map of the study area showing the pump location and status. The inset shows the study area within Rumphi District in Malawi (Africa).

Table 2. Probable reason(s) for each pump's non-functionality ($n = 23$)

Pump ID	Reported drilled depth (m)	Driller ID*	How long have the users reported that the well has been non-functional?	Probable reason(s) for each pump's non-functionality
P1	14	J	2 years	The well was not drilled deep enough; respondents indicated low water quantity. The pump has been used once since installation.
P2	18.4	H	1+ years	The well was not drilled deep enough; respondents indicated the well produces some water during rainy season only.
P3	11	F	1+ years	The rope is stuck in the rising main pipe.
P4	19.5	A	1+ years	The well was not drilled deep enough; respondents indicated low water quantity. The pump has not been used since installation.
P5	13.1	B	1 year	Respondents indicated that there is no trained committee and that the driller handed over a non-functional pump to the community.
P6	11.7	C	7 months	Worn rope, there is a need for a new rope.
P7	17.5	A	>4 months	Worn rope, there is a need for a new rope.
P8	12.6	C	4 months	Worn rope, there is a need for a new rope. As an alternative water source, the community is currently getting water at the school where there is an Afridev pump.
P9	11	C	4 months	Worn rope, there is a need for a new rope and washers.
P10	16	E	4 months	The well was not drilled deep enough; respondents indicated the well produces some water during rainy season only.
P11	9.8	C	3 months	There is a need for a replacing the worn rope and a new rising main pipe. The community has a poorly organized committee.
P12	13.5	C	2 months	The worn wheel needs to be replaced.
P13	10.9	C	1 month	Worn rope, there is a need for a new rope. The committee stated they do not have money to buy the rope.
P14	19.8	I	1 month	Worn rope, there is a need for a new rope. The community has a poorly organized committee and stated they do not have money to buy the rope.
P15	15.2	C	1 month	Worn rope, there is a need for a new rope.
P16	12	G	1 month	Worn rope, there is a need for a new rope. Otherwise the pump is in good condition.
P17	14.5	J	1 month	There is a need for a new rope.
P18	11.5	C	3 weeks	Worn rope, there is a need for a new rope which the committee stated will be done soon.
P19	12.8	C	2 weeks	Worn rope, there is a need for a new rope and set of washers. As an alternative water source, the community is currently getting water from the lake.
P20	15	G	2 weeks	Worn rope, there is a need for a new rope. Otherwise the pump is in good condition.
P21	12.5	G	2 weeks	Worn rope, there is a need for a new rope.
P22	21	D	2 days	Worn rope, there is a need for a new rope and washers.
P23	15.5	J	1 day	Worn rope, there is a need for a new rope and set of washers. Additionally, the well was possibly not drilled deep enough; the respondents indicated that the well produces some water and then stops.

* Driller names have been anonymized; there were 10 drillers for the 23 non-functional wells, and each was assigned a Driller ID of A to J

pumps, while 5 other drillers had a mix of functional and non-functional pumps, and only two of the drillers had each of the pumps they had installed as functional pumps. For the 5 drillers which had all non-functional pumps, each installed only 1 to 2 pumps and the probable reason(s) for non-functionality were, for 4 pumps, that it was not installed properly, whereas 2 pumps had an operation and maintenance system failure. The Kruskal-Wallis rank sum test showed that there were differences ($p = 0.010$) in the depth drilled that was reported by the driller; some drillers went deeper.

User feedback

User satisfaction may be linked to the quantity or quality of water, as well as ease of operation and operational and maintenance issues. Approximately 75% (352/472) of the respondents were satisfied with their pump and 57% (277/472)

of the users reported that after installation the rope and washer pumps had provided uninterrupted service when they needed it. However, variation in the water availability was reported by users within the same community at 32/48 pumps, which possibly indicated different usage patterns or other secondary sources of water other than the rope and washer pumps being studied. When those users without uninterrupted service were asked what alternative source(s) were used when the rope and washer pump was not used or not functioning, many (148/195) reported using an unprotected water source, such as a river, lake, or open shallow well, because these were the nearest available alternatives. When pumps were providing water (functional), the users were more likely to be satisfied than those with non-functional pumps, based on a Fisher's Exact Test ($p < 0.05$).

Rural water supply users may judge the quality of water based on subjective indicators, such as colour, smell, taste and hardness, rather than objective water quality results. Many (414/472) users

reported that they liked the taste of the water coming from the pump. But, if the rope and washer pump was functional, users were not always drinking water from it due to user preferences. Five respondents from two communities indicated that there was water available at their pump, but they were not drinking water from it because the water had a bad smell; however, the water was suitable for drinking based on the local standards for thermotolerant coliforms (Malawi Bureau of Standards, 2005).

For 37% (176/472) of the users, the pumps had been broken down for more than 1 day in the last 30 days (not including outages due to routine maintenance). This user feedback also triangulates well with the researcher field observations of pumps with frequently changing status of whether it had water or not. When a pump stopped providing water, users reported that it took 1 to 6 months to repair the pump. At one non-functional pump during the first site visit researchers probed further, which was possible because when the research team arrived at the site the trained driller who lived 0.5 km away also came to the community. Users reported the pump had been non-functional for 1 month, though before that the community reported that the water supply was good. During onsite discussions with the community, researchers, and the trained driller, it was determined that a section of the rope needed to be replaced and the handle needed oil to bring the pump back to a functional status. Materials for the repair were not available in the community, but were available at the nearest market town. The approximate cost was MK500 (0.71 USD) for materials and an additional MK500 (0.71 USD) for labour. This is a specific example where, despite the nearby trained driller and materials for repair available at the nearest market town, uninterrupted service could not be maintained as users are responsible for at least initiating repairs. The well was subsequently fixed, and observed to be functional by researchers during the second site visit.

It is an important criterion that users should know the management structure for their water point. Water point committees are a subset of 10 community users elected for decision-making and management roles related to pump operation and maintenance, while users that are not part of the committee would typically have limited roles in water source management. Water point committees were also responsible for cleaning the water point's surroundings and providing security against vandalism. Local government records indicate that each of the 48 pumps had a committee. Many (386/472) users reported that there was a functional water point committee for their rope and washer pump. Many users (298/472) also reported that it was this committee that was responsible for the maintenance of the water point, while 125/472 users reported that it was a hired technician who performed the maintenance and was paid by the committee with user community contributions. Of the pumps studied, 27/48 of these have a female committee chairperson. Women are encouraged by the local government to actively participate in decision making concerning rural water supply.

In addition to drinking water, other reported uses (186/472) included agricultural activity (gardens watered with buckets or livestock watering), brick making, and domestic uses (cooking, bathing and washing clothes). Only one community named another community within the geographic area, but outside the scope of the supported development partner project, that had a rope and washer pump installed at a small hotel offering accommodation for travellers and categorized it as being for commercial purposes.

Less than half of the users (181/472) reported that there

were cash resources that have been collected as user fees for the pump's operation and maintenance, and only 6% (30/472) knew how much a rope and washer pump costs. If given an opportunity to choose amongst the technologies that would work in their community, approximately one quarter (130/472) would still choose the rope and washer pump. This also compares to an approximately equal number of users (116/472) who indicated that, if given the opportunity to choose what might work well in their community, they would want a piped water supply.

Sustainability

Following the infrastructure checklist, user interviews, and water quality analysis, an assessment was made by researchers of the pumps' technical, operational or maintenance sustainability based on 4 checks: that it was providing water, water was safe for drinking based on thermotolerant coliforms (Malawi Bureau of Standards, 2005), all the users surveyed for that pump indicated that the individuals expressed satisfaction with the pump, and there was a flow rate of at least 20 L/min. The results suggest that rope and washer pumps were working as a sustainable rural water supply in 8% (4/48) of communities.

Of the 8 pumps classified as the most rural, only one of these qualified as being sustainable by our study criteria. To determine if there was a variation between the remoteness of the pump and the three pump status groups (is non-functional, had water, and is sustainable), Fisher's Exact Test showed no differences ($p = 0.86$). This indicates that the pumps are not necessarily working better or worse in more remote areas. To determine if there is a variation between the gender of the water committee chair and the three pump status groups (is non-functional, had water, and is sustainable), a Fisher's Exact Test showed no differences ($p = 0.56$). This means that a female committee leader did not necessarily lead to a more, or less, sustainable rope and washer pump. Also, the Kruskal-Wallis rank sum test further showed that deeper wells were not necessarily more likely to be sustainable ($p = 0.53$). The Fisher Exact Test relation between whether a specific driller was more likely to have a sustainable ranking showed no differences ($p = 0.54$).

In addition, the monitoring logistics of rural pumps for sustainability is expensive. This current study cost approximately 3 000 USD for the 48 pumps.

DISCUSSION

A rigorous examination of the elements of water quality, infrastructure, and user feedback might help to spread sustainable rural water supply practices. Where the rope and washer pumps were functional, for most the water quality met national guidelines. Yet, pumps were working sustainably as sources of rural water supply in only 4 of 48 study communities.

A professional driller did not guarantee a sustainable rural water supply. For at least some of the non-functional wells there is an indication that pumps may have been impacted by high failure rates soon after installation. In addition, some of the water sources may need to be drilled deeper. Yet, the aim to have water users and private operators completely independent of government regulation by using a low-cost technology was also not being completely met over the longer term, as observed in this study. It could be that the private operators, although having been trained and being geographically available to customers, may not have been profitable, able to be reached by users, or there may have been too many operators

within the study area. Thus, there may have been too much competition among private operators or poor communication with users within the start-up period, but this needs further investigation. The functionality of rope and washer pumps had great variability, but could be a cycle in which quality control is more difficult when local regulations are not in place for a new technology. However, the functionality of the rope and washer pumps in this study was even lower than the rates reported for Elephant pumps in rural Malawi, which are a similar covered rope and washer pump design and also not approved by the Malawi Government (Holm et al., 2015). More widely, the functionality of water points (inclusive of handpumps, piped water, and unimproved sources) in Malawi is reported to be 79% (Banks and Furey, 2016), in which case the pumps under study had a lower overall functionality than this average.

In the approximately 2 years from drilling completion to the time of our study, there was a slow scale-up of the low-cost technology for surrounding geographic areas for household or community utilization, for which private individuals may have paid the full cost to the trained drillers. Users who are happy with the rope and washer pump did not necessarily share this information outside their community, nor are they excited enough about rope and washer pumps to further install their own rope and washer pump at home. One of the factors that affects user satisfaction is the ease of the pumping motion by the user, which is a function of the depth of the well and maintenance, although this needs further research over time, both as the pump is used more and as it gets older. Although someone must take the lead to promote a new technology, the low adoption rate is likely a result of the non-functionality of rope and washer pumps in the study area, and that the communities that were targeted with the rope and washer pumps were small and isolated.

In view of the rural water supply model that is currently operating in Malawi and which depends on decentralized management (Malawi Government, 2005), one aspect that ensures that rural water supplies are sustainable is the ability of these communities to take care of their own supplies. The aspect of ownership ultimately rests with the community, which must be empowered and not wait for Government or outside agencies to repair and maintain supplies. There is a disconnect between pump users knowing that there is a water point committee and local government reporting that each of the 48 pumps had a committee. Our research is also consistent with findings of Chowns (2015), who found that the community management of rural water supply in Malawi does not work, because maintenance is not performed and water point committees are unable to manage funds. A specific example of this in our study is the high rate of pumps for which it was needed to replace or tighten the rope.

The wells under this study are, in general, strictly rural water supply systems. In some cases, the research team had to hike or take a boat to access the water sources, as neither a motorcycle nor a 4x4 vehicle could reach the sites. These hardest to reach areas are inaccessible to conventional drilling rigs. Of the 8 most rural pumps, only one was classified as sustainable, yet this is where rope and washer pumps with manual drilling have been marketed to have the most impact. Access to spare parts and access to information in these hardest to reach areas likely plays a role and needs further detailed study.

In Ethiopia, rope and washer pumps are being installed for joint irrigation and domestic use (Butterworth et al., 2013). The pumps in this study were used for both drinking and non-drinking water purposes, although possibly not as strongly as

in Ethiopia for multiple-use water services. In Ethiopia, many of the farmers using rope and washer pumps wanted to change their water lifting technologies (Nigussie et al., 2017). However, when the pumps in this study were providing water, users were generally satisfied.

The rising main on rope and water pumps is polyvinyl chloride (PVC); the quality of water is, therefore, site specific, and contamination may come from sources at, around, and near the water source and may not necessarily be related to the type of pump, assuming the pump is properly installed. In terms of safe water, the pumps in our study generally had better water quality than reported results for rope and washer pumps in other countries (Harvey and Drouin, 2006; Bennett et al., 2010; Coloru et al., 2012; Butterworth et al., 2013). However, there is limited drinking water quality data from Rumphi District (Holm et al., 2018) for the three Government of Malawi (2016) approved handpump designs. In the case where the pump is not being used for drinking water, the national and WHO guidelines used in our evaluation would not apply. This study has shown that, despite having a pump that is not fully enclosed, safe water can be provided effectively by rope and washer pumps.

Hutchings (2018) argues that management of rural water supply relies heavily on state supporting agencies. The Government of Malawi (2016) technical manual specifically outlines three approved handpump designs and leaves no room for the assessment criteria of alternative rural water supply technology. In addition, the community-based management training manual (Malawi Government, 2015) is equally specific and only covers Afridev pumps. In the case of the study pumps, the District Council was the regulator of this low-cost technology pilot project and reported to the national government. There is a need for a community-based management training manual that allows for alternative rural water supply technologies that may work in some areas, as well as supervision courses for regulators. In addition, there is an ongoing need to identifying causes of failure that stakeholders can target to improve user satisfaction, operation and maintenance.

Limitations

The study was limited to one district. Although it included all pumps that were installed as a cohort development project, this is still a small sample size. The water quality evaluation was conducted only once during the dry season, and one functional pump was skipped because it initially could not be found, although it was located later in the study. Borehole forensics, such as borehole verticality and aquifer hydraulic properties, were not conducted on non-functional wells. There may have been a conflict of interest, as the lead researcher (JM) at the time of the study was working for the Government of Malawi as the Rumphi District Water Officer managing the rural water supply. This research was not necessarily an assessment of rope and washer pumps for self-supply and, hence, did not cover financing by the user. The sustainability of groundwater resources for future use or year-round water supply was not studied, as each pump was only visited twice by researchers, though this set of rope and washer pumps had a well-documented history by Mzuzu University.

CONCLUSIONS

Most rope and washer pumps in the study, although not approved by the Government of Malawi, provided safe drinking water when working. The high non-functionality rate can be attributed

to community management and operational or maintenance problems. Only 8% (4/48) of pumps had water flowing at two researcher field visits, good water quality, a flow of > 20 L/min and full consensus of positive satisfaction among users. An approach to evaluate and introduce new technologies for rural water supplies that have not yet been approved by the national government has not been established in Malawi. However, there needs to be a balance so that practitioners may obtain data showing that an alternative rural water supply technology is working. Sub-Saharan African governments need to be open to innovative solutions while ensuring that standards, including standards for functionality, water quality, user satisfaction, private operators, and human capacity for local government regulators, are being followed to ensure safe water for rural communities.

ACKNOWLEDGEMENTS

This research has been supported as part of the Climate Justice Water Futures Programme funded by the Scottish Government and facilitated by University of Strathclyde in partnership with BASEflow through a mini-research grant. The authors appreciate the valuable feedback of Dr Kip McGilliard on a draft of this manuscript.

REFERENCES

- BANKS B and FUREY SG (2016) What's working, where, and for how long. A 2016 water point update to the RWSN (2009) statistics. In: *Proceedings of the 7th RWSN Forum*, 29 November 2016, Abidjan, Côte d'Ivoire.
- BENNETT HB, SHANTZ A, SHIN G, SAMPSON ML and MESCHKE JS (2010) Characterisation of the water quality from open and rope-pump shallow wells in rural Cambodia. *Water Sci. Technol.* **61** (2) 473–479. <https://doi.org/10.2166/wst.2010.817>
- BUTTERWORTH J, SUTTON S and MEKONTA L (2013) Self-supply as a complementary water services delivery model in Ethiopia. *Water Alternatives* **6** (3) 405–423.
- CHOWNS E (2015) Is community management an efficient and effective model of public service delivery? Lessons from the rural water supply sector in Malawi. *Public Admin. Dev.* **35** (4) 263–276. <https://doi.org/10.1002/pad.1737>
- COLORU B, MGAYA S and TAUBERT RP (2012) Appropriate technologies for rural water supply: A comparative study between "Rope pumps" and conventional piston pumps on water quality and other sustainability parameters. Fondazione ACRA-CCS, Njombe.
- CRESWELL JW (2014) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications, Inc., Los Angeles.
- GOVERNMENT OF MALAWI (2016) Technical Manual – Water Wells and Groundwater Monitoring Systems. Ministry of Agriculture, Irrigation and Water Development, Lilongwe.
- HARVEY PA and DROUIN T (2006) The case for the rope-pump in Africa: A comparative performance analysis. *J. Water Health* **04.4** 499–510. <https://doi.org/10.2166/wh.2006.0033>
- HOLM R, KAMANGIRA A, KASULO V, KAPONDA P, HARA E, CARNEY-FILMORE C and NHLEMA M (2017) The handpump choice is yours: A pilot study in Rumphi District, Malawi. *Waterlines* **36** (4) 358–366. <https://doi.org/10.3362/1756-3488.17-00006>
- HOLM R, STROUD R, MSILIMBA G and GWAYI S (2015) Functionality and water quality of Elephant pumps: Implications for sustainable drinking water supplies in rural Malawi. *Groundwater Sustainable Dev.* **1** 129–134. <https://doi.org/10.1016/j.gsd.2016.02.001>
- HOLM RH, KUNKEL G and NYIRENDA L (2018) A thought leadership piece: Where are the rural groundwater quality data for the assessment of health risks in northern Malawi? *Groundwater Sustainable Dev.* **7** 157–163. <https://doi.org/10.1016/j.gsd.2018.05.004>
- HUTCHINGS P (2018) Community management or coproduction? The role of state and citizens in rural water service delivery in India. *Water Alternatives* **11** (2) 357–374.
- MALAWI BUREAU OF STANDARDS (MBS) (2005) Malawi standard; Borehole and shallow well water quality– specification, MS 733:2005. Malawi Standards Board, Blantyre, Malawi.
- MALAWI GOVERNMENT (2005) National water policy. Malawi Government, Lilongwe.
- MALAWI GOVERNMENT (2015) Rural Water Supply Operation and Maintenance Series 1 - Community Based Management (O & M Refresher Course) Training Manual. Ministry of Agriculture, Irrigation and Water Development (MoAIWD). Malawi Government, Lilongwe.
- MALAWI GOVERNMENT (2016) Technical Manual - Water Wells and Groundwater Monitoring Systems. Ministry of Agriculture, Irrigation and Water Development. Malawi Government, Lilongwe.
- NIGUSSIE L, LEFORE N, SCHMITTER P and NICOL A (2017) Gender and water technologies: Water lifting for irrigation and multiple purposes in Ethiopia. International Water Management Institute, East Africa and Nile Basin Office, Addis Ababa.
- SUTTON S (2017) Trends in sub-Saharan rural water supply and the essential inclusion of self-supply to achieve 2030 SDG targets. *Waterlines* **36** (4) 339–357. <https://doi.org/10.3362/1756-3488.17-00013>
- WORLD HEALTH ORGANIZATION (WHO) (2017) Guidelines for Drinking-water Quality, Fourth Edition Incorporating the First Addendum. URL:<http://apps.who.int/iris/bitstream/handle/10665/254637/9789241549950-eng.pdf;jsessionid=054D53AD5E1E1CCD6C897ED3194E1EB80?sequence=1> (Accessed 22 April 2018).

Prevalence of bromide in groundwater in selected regions in South Africa

HL Lucht¹ and NH Casey¹

¹Department of Animal and Wildlife Sciences, University of Pretoria, South Africa

ABSTRACT

Many regions across South Africa are dependent on groundwater as the only water source for livestock watering and domestic use. This paper presents an analysis of 350 water samples from collated data of 5 reports published between 2001 and 2016 that show the vast range of 0–132.68 mg/L bromide (Br⁻) present in South African groundwater. It further highlights that Br⁻ may be a greater toxicity risk factor to livestock production and human health as an endocrine disrupting chemical (EDC) and through accumulation in organs than previously considered. Further validation is required of the physiological effects of Br⁻ for inclusion in water quality guidelines (WQG). Attention is drawn to the importance of site-specific water quality (WQ) monitoring and identification of vulnerable populations to enable adequate risk assessment and implementation of mitigating strategies to lower exposure risk in a specific area.

Keywords: bromine, halide, toxicity, water quality

INTRODUCTION

Many regions across South Africa are dependent on groundwater as the only water source for extensive and intensive livestock production, and wildlife in game farming and game reserves. In remote regions, domestic households might also be dependent on groundwater, as access to surface water or roof-harvested rainwater could be erratic, and therefore exposed to potentially hazardous elements. Initially health problems were reported in livestock in specific areas and fluorosis was identified as a major problem (Coetzee et al., 1997; 2000). Water samples were collected at various points from source and use-points for the purpose of compiling reports addressing the risk presented to livestock. Chemical analysis of water quality constituents (WQC) of the earliest water samples confirmed the presence of fluoride in excess of reported water quality guideline (WQG) safe levels. This information was used to determine an appropriate method of risk assessment and subsequent decision-making on fitness-for-use (FFU) of available water sources.

The initial focus of early research was to formulate and test solutions to most frequently identified palatability and toxicological water quality (WQ) problems of groundwater drawn from wells, springs and boreholes (Casey et al., 1998). Increased efforts to measure and assess the potential risk posed by geochemistry-related factors resulted from advances in identifying the role of inorganic constituents on the epidemiology of non-differential clinical symptoms commonly observed in livestock (Meyer et al., 2000).

As more information emerged, objectives of the research projects changed to include analysis of multiple elements in water sampled from various regions across South Africa. It emerged from those research projects that bromide (Br⁻) was present in many of the samples collected from areas already identified as at-risk of exposure to known hazardous elements. It further emerged that Br⁻ was present at potentially harmful concentrations for many of the selected sites.

There were no formal international or local WQG available for Br⁻ because it was initially merely acknowledged as a

ubiquitous micro-element of unknown essentiality and not considered problematic. Similarly, water was ignored as a nutrient. Traditionally the focal point of deficiency or toxicity research was limited to the contribution of micronutrients from premixes to nutrient composition of feed. As research progressed over time, a very different picture emerged and raised many questions. It emphasised that water should be given closer attention as a potential source of micro-elements in human and livestock nutrition (Casey and Meyer, 2001).

This paper presents collated results of data collected in selected regions in 6 provinces in South Africa and published in various reports between 2001 and 2016. The aims of the Water Research Commission (WRC) and other sponsored research were to determine FFU of water for livestock production. This was extended to include game such as ostrich production. The aim of this paper was to show the considerable range of Br⁻ concentrations present in groundwater in selected regions, and to draw attention to the reality that Br⁻ may have a much greater impact on animal production than previously considered.

MATERIALS AND METHODS

Data of groundwater Br⁻ concentrations were collated from water sample results (Table 1). The reports from which the data were collated were not all in the same format. Sampling for each report was done according to the objectives of the individual report, which resulted in an uneven dataset. There were no GPS co-ordinates available for points sampled for earlier reports and thus compilation of a distribution map for this paper was not possible.

Data were sorted according to sample source (Table 2) and locality across years and seasons to obtain information of overall Br⁻ concentrations present in groundwater used for livestock production and household use across selected regions. The majority of the samples were collected in areas where the human and livestock populations were dependent on groundwater as the only source of water for drinking and domestic use.

Initially potentially problematic physiologically significant trace elements were identified according to the specific requirements of each individual project prior to water sample collection. All water samples were collected for the

*Corresponding author, email: h.lucht@gmail.com

Received 28 June 2018; accepted in revised form 14 June 2019

completion of individual reports using the same method, and all samples were analysed using inductively coupled plasma atomic emission spectrometry (ICP-AES) techniques with full quantitative and semi-quantitative procedures by the Institute for Soil, Climate and Water at the Agricultural Research Council (ARC-ICSW), Pretoria. Elements present in the samples were classified as constituents of concern (COC) or potentially hazardous chemical constituents (PHCC) according to their presence at concentrations relative to local and international WQG standards. Not all elements were present in all water samples, and for the purpose of this paper, only Br⁻ concentrations were considered.

Where possible, water samples were collected at source from wells, springs and boreholes, from the surface of reservoirs, tanks or drinking troughs supplied by boreholes and at use-points such as from taps or water lines in poultry houses (Table 2). In WRC reports published before 2001, Br⁻ was not included as a COC, but as time progressed and the objectives of the projects expanded, it emerged from the various case studies that Br⁻ should be included in further testing. A total of 350 groundwater samples were ranked from highest to lowest Br⁻ concentrations across all years, sources, seasons and localities to determine the overall range of Br⁻ present in groundwater in a South African context. Water samples were grouped by collection source within province to determine an overall picture of how Br⁻ concentration may vary in different sample sources within a locality.

RESULTS AND DISCUSSION

Inconsistent sampling techniques, such as sample collection by different technicians and variations in sample collection depth, and inconsistent or seasonal use of boreholes in some areas could explain the occurrence of unavoidable sampling errors with effects on the accuracy of measurements.

It emerged from the collated data that the overall range of Br⁻ concentrations present in groundwater sampled across selected regions in South Africa was 0–132.68 mg/L.

Samples taken from dams, source, reservoirs and use-points included in the dataset were limited to those supplied by boreholes (Tables 2 and 3).

There were distinct differences in Br⁻ concentrations between sources within and across regions (Tables 2 and 3). The majority of water samples collected were for selected areas in the North West and Limpopo Provinces in accordance with the specific objectives of projects undertaken to address area-specific problems reported within those provinces.

Groundwater sampled from reservoirs showed higher Br⁻ concentration compared with groundwater sampled at source,

Table 1. WRC Reports used for data collation

Report	Author
WRC 857/1/01	Casey and Meyer, 2001
WRC 857/2/01	Casey, Meyer and Coetzee, 2001
032005/02/26	Meyer, 2005
WRC 1175/1/06	Casey and Meyer, 2006
WRC 2175/2/16	Korsten, Casey and Chidamba, 2016

Table 2. Br⁻ concentrations (mg/L) within source across all years, all seasons, all localities

Sample	n	Min	Max	μ	SE
Dam	2	0.79	0.84	0.81	0.04
Source	269	ND*	132.68	3.52	14.53
Reservoir	63	0.01	6.44	0.80	1.19
Use-point	16	0.01	0.38	0.17	0.12

*ND = not detected

with the exception of the 97 water samples collected at source in Limpopo Province (Table 3). Many factors, such as depth of sampling, sampling site, water flow rate, water usage rate, and pumping frequency could influence the accuracy of the measurement of a WQC in a definitive water sample of the source, and Br⁻ is no exception. However, since it is not possible with current technology to verify the measured concentration, it was assumed that each sample was an accurate representative sample and the measurement an accurate estimation of the concentration of a WQC at the time of sampling.

Water stored in open reservoirs or held in drinking troughs that are not subject to high stocking rates or frequent use are expected to contain higher concentration of Br⁻ than water sampled at source or at use-point. This is because exposure of open troughs and reservoirs to evaporation results in a concentrating effect on Br⁻ within that water body, which is true for all WQC. Groundwater pumped into open reservoirs exposed to UV radiation is subject to speciation of Br⁻ in the presence of oxygen to form bromate (BrO₃). The rate of conversion is dependent on pH and presence of other elements. This speciation of Br⁻ is potentially hazardous since BrO₃ is a known carcinogen (Jain et al., 1996; DeAngelo et al., 1998; Magazinovic et al., 2004; Bonacquisti, 2006; Moore and Chen, 2006). Similarly, water sampled from pipes exposed to sun may differ in composition to water sampled at source because the rate of elemental interactions within water accelerates with heating (Table 2).

The current South African WQG for livestock watering does not list Br⁻ as either a COC or PHCC (Casey and Meyer, 1996). It is common for products of endocrine disrupting chemical (EDC) metabolism to be more toxic than the parent

Table 3. Br⁻ concentrations (mg/L) in groundwater in selected regions clustered by province

Sample	n	min	max	μ	SE	Sample	n	Min	max	μ	SE
North West						Western Cape					
Source	141	0.00	2.14	0.31	0.37	Source	13	0.04	6.60	3.11	2.01
Reservoir	36	0.03	2.09	0.42	0.55	Reservoir	5	2.43	6.44	3.82	1.62
Use point	16	0.01	0.38	0.17	0.12	KwaZulu-Natal					
Limpopo						Source	5	0.01	0.23	0.12	0.09
Dam	2	0.79	0.84	0.81	0.04	Eastern Cape					
Source	97	0.01	132.68	8.87	23.29	Source	6	0.06	0.37	0.21	0.13
Reservoir	22	0.01	2.98	0.74	0.88	Eastern Cape					
Mpumalanga						Source	6	0.06	0.37	0.21	0.13
Source	6	0.05	0.25	0.14	0.09	Eastern Cape					

compound (Burger, 2005). An EDC is any naturally occurring or synthetic chemical that interferes with the structure or function of hormone receptor complexes, either in an antagonistic or synergistic way, to alter the correct function of an endocrine response within a target organ (Bornman et al., 2007). The USEPA (1997) expands the definition of an EDC to include that the exogenous substance causes adverse health effects in the intact organism, its progeny or (sub) populations. EDCs commonly monitored for hazards to human and animal health are usually lipophilic organic compounds with oestrogenic properties (Bornman et al., 2007). Naturally occurring Br⁻ is a hydrophilic inorganic element identified as an EDC in rats (Loeber et al., 1983) and chickens (Du Toit and Casey, 2012) and is expected to have the greatest direct disrupting effect on metabolism in vulnerable livestock and human populations.

The report by Casey and Meyer (2001) lists Br⁻ with a maximum permissible level (MPL) of 1–3 mg/L and a crisis level of 6 mg/L, with the recommended limit set at 1 mg/L due to risk of BrO₃ formation at that concentration. A subsequent report (Casey and Meyer, 2006) introduced 0.01 mg/L as a maximum contaminant level (MCL) to align it with USEPA (2005) guidelines. Faced with conflicting reports of what constituted a safe minimum concentration against which to compare results obtained from field samples, Casey (2016) accepted a minimum level of 0.01 mg/L as a point of departure for analysis and interpretation of the test results. This was in line with the accepted default maximum residue level (MRL) of 1 mg/kg used for most food additives not yet validated in terms of Regulation (EC) No. 396/2005 (European Parliament, 2005).

Traditional WQG propose generic safety levels of elements based on concentration-based estimates, which assume a linear relationship between the concentration of an element in source and its effects in vivo. Limitations of such a generic approach are that the accepted safety limits of elements in feed and water are seldom published in the same guidelines, and limited differentiation exists between different types of livestock or game species where applicable. Further limitations are that interactions between elements in the same source are ignored and the assumption that all groundwater sources in the same area are of equal quality.

The disadvantage of a concentration-based approach is that exposure risk is disregarded as being multifactorial when it is influenced by any factor that affects water intake rate or physiological state of an individual. Intake-based guidelines that are site-specific will better estimate exposure risk of a target population, thus allow for better mitigation of adverse effects.

All elements, whether essential or nonessential, can exert toxic effects when consumed in excess through water or feed, which includes minerals occurring in feed and water at trace levels otherwise regarded as incidental contaminants with no obvious important nutritional role (NRC, 2005). PHCC have adverse effects at relatively low levels, and magnitude of exposure risk depends on exposure period duration (Plant et al., 1996). Low-dose, long-term exposure to PHCC will most likely manifest in subclinical responses where toxicity is expressed as secondary induced deficiencies, making toxicity symptoms difficult to identify (Meyer and Casey, 2004). Similarly, EDCs exert their effect at very low exposure levels (Bornman et al., 2007).

The vast range in Br⁻ concentrations in water sampled from Limpopo Province compared with other provinces (Table 3) draws attention to the importance of site-specific analysis of groundwater sources when determining FFU of such sources and

the potential risk of vulnerable population exposure to hazardous chemical constituents. Site-specific risk assessment requires that geochemical factors on soil and plant concentrations be included in the total exposure risk estimation for a given area to enable formulation of contextual solutions (Meyer et al., 2000).

Some areas were chosen for sampling to determine the quality of alternative water sources in provinces that were not solely dependent on groundwater, in line with the research objectives to generate specific reports. This resulted in collection of relatively few groundwater samples from those provinces compared with areas where groundwater played a greater role (Table 3).

Many environmental health effects caused by nutritional element excess and deficiencies in South African agricultural systems have been documented, yet there are still health impacts of potentially harmful elements that are less known (Davies and Mundalamo, 2010). Heavy metals are known to be toxic due to their cumulative nature and cause increasing damage to brain, kidney and nervous system with extended exposure periods (Ezekwe et al., 2012). Similarly, Br⁻ has been shown to accumulate in liver, kidney and thyroid tissue (Du Toit and Casey, 2012; Mamabolo et al., 2009). Further fieldwork done by Meyer (2005) included tissue sampling and revealed evidence that Br⁻ had histopathological effects on thyroid and other tissues in commercial broilers reared in areas where Br⁻ concentrations in groundwater were high. It is known that Br⁻ has the ability to circulate freely and rapidly into the extracellular fluid and various tissues of the body except the central nervous system (Pavelka et al., 2000). This free movement throughout the body affords Br⁻ the opportunity to interfere with multiple biochemical processes. Although further validation is required, it appears that Br⁻ could be labelled an EDC, which is a concern for livestock farmers and people who might be exposed to Br⁻ in drinking water.

Identifying Br⁻ as a COC or PHCC in water sources of areas where no alternative water sources were available raised further questions about the best ways to define and identify vulnerable populations to determine FFU of these water sources. Risk assessment relies on the identification of vulnerable populations within an area, because water requirements will differ between groups within a population according to age and physiological state (Table 4). Vulnerable populations in livestock production include neonates, very young and actively growing animals, immunocompromised animals and pregnant and lactating females. Where multispecies water use is common, such as in game reserves with watering holes supplied by borehole water, interspecies differences in mineral tolerance must be considered in the FFU decision-making process due to the different species-specific metabolic requirements related to physiological state. Bornman et al. (2007) stated that EDCs can pose risks to reproductive function, immunity, thyroid function and neurodevelopment, dependent on the type of substance and its toxicodynamic and toxicokinetic mechanisms of action.

The geochemical character of groundwater depends on mineral chemistry of aquifer materials and biomediated ion exchange reactions (Ezekwe et al., 2012). Changes in environment such as ambient temperature, feed composition and water palatability influence water intake. Physiological differences between groups translate to differences in metabolism and assimilation rates of elements. Exposure risk depends on the per capita consumption of an element relative to body weight (Ezekwe et al., 2012) and this is clearly shown in Table 4. Immature and actively growing individuals are thus at greatest risk of developing toxicity symptoms from relatively

Table 4. Estimated intake of Br⁻ through water by humans

Persons	WQG	Br ⁻ in water (mg/L)		Water Intake	Br ⁻ /day by WI (mg)	
	mg/L	Max	μ	L/day*	Max	μ
Males: adults and adolescents	0.01	133	3	2.30	305	7
Children: both sexes 4–12 yr	0.01	133	3	0.55	73	2
Children: both sexes 0–3 yr	0.01	133	3	0.40	53	1
Women: pregnancy < 18 yr	0.01	133	3	2.30	305	7
Women: pregnancy 19–50 yr	0.01	133	3	2.30	305	7
Women: lactating < 18 yr	0.01	133	3	2.90	385	8
Women: lactating 19–50 yr	0.01	133	3	2.90	385	8

*Assumed for normal healthy people of moderate lifestyle at 95 % of the empirical distribution (EPA, 2004)

lower concentrations of COC or PHCC due to a combination of limited capacity of immature organs for adequate detoxification, and greater rates of assimilation of elements by tissues with high metabolic activity (Table 4). As a result, it is common practice to assign water sources of relatively poorer FFU scores to the least vulnerable groups within a population when alternative water sources are unavailable. In some cases, where practical, water treatment can improve its elemental quality sufficiently to make it safe for use.

The use of sentinel species is a useful tool to evaluate risk to vulnerable populations over time. Meyer (2015) used indigenous chicken breeds and commercial broiler chickens produced in a specific locale as a reference point for risk assessment of groundwater containing high concentrations of Br⁻ for the selected area. The additional collection of multiple tissue sample types from sentinel species, together with single water samples, allowed for better identification of chronic exposure risk to PHCC than water sampling alone, with liver samples reported to be the most appropriate tissue sample for assessment of Br⁻ exposure risk (Casey and Meyer, 2006). The most suitable choice of sentinel species in an area will depend on specific monitoring objectives and the practicality of tissue sample collection for testing. Future consistent sampling of the same sites over time will garner more information on the toxicity risk that Br⁻ in groundwater poses to populations in the area at different times of the year and in different situations. Monitoring specific groundwater sources could indicate which water usage patterns could effectively limit exposure of vulnerable populations to COC and PHCC.

CONCLUSION

The considerable range of concentrations of Br⁻ occurring in South African groundwater presented in this paper draws attention to the importance of monitoring site-specific WQ for FFU assessment for domestic and livestock use. It further highlights that Br⁻ may be a greater toxicity risk factor to livestock production and human health than previously considered. In order to be included in WQG, further validation is required on the physiological effects of Br⁻ and associated risk factors. Identification of vulnerable populations is paramount to the selection of the best solution to alleviate risks of exposure to Br⁻ in groundwater. Continued seasonal monitoring is recommended to identify potential risks linked to changes in WQ and to assist in the diagnosis of physiological anomalies.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the Water Research Commission of South Africa and Dr JA Meyer for access to the published report 032005/02/26, and the National Research

Foundation for funding through the incentive grant to NHC, reference: PR-IFR180205310035/UID96806.

REFERENCES

- BONACQUISTI TP (2006) A drinking water utility's perspective on bromide, bromate and ozonation. *Toxicology* **221** 145–148. <https://doi.org/10.1016/j.tox.2006.02.010>
- BORNMAN MS, VAN VUREN JH, BOUWMAN H, DE JAGER C, GENTHE B and BARNHOORN EJ (2007) The use of sentinel species to determine the endocrine disruptive activity in an urban nature reserve. WRC Report No. 1505/1/07. Water Research Commission, Pretoria. ISBN 978-1-77005-551-3.
- BURGER AEC (2005) WRC programme on endocrine disrupting compound (EDCs): Volume 1: Strategic research plan for endocrine disrupters in South African water systems. WRC Report No. KV 143/05. Water Research Commission, Pretoria. ISBN 1-77005-348-4.
- CASEY NH (2016) Inorganic chemical quality of water. In: Korsten L, Casey NH and Chidamba L (eds.) *Evaluation of the risks associated with the use of rooftop harvesting and groundwater for domestic use and livestock watering. Volume 2: Chemical quality of groundwater for potable use and livestock watering*. WRC Report No. 2175/2/16. Water Research Commission, Pretoria. ISBN 978-1-4312-0875-3. <https://doi.org/10.3133/wri834008>
- CASEY NH and MEYER JA (1996) In: Holmes S (ed.) *South African Water Quality Guidelines. Volume 5: Agricultural Use: Livestock Watering*. Department of Water Affairs and Forestry, Pretoria.
- CASEY NH and MEYER JA (2001) An extension to and the further refinement of a water quality guideline index system for livestock watering: Volume 1: Rural communal livestock production systems and wildlife production systems. WRC Report No. 857/1/01. Water Research Commission, Pretoria. ISBN 1-86845-713-3.
- CASEY NH and MEYER JA (2006) The application of risk assessment modelling in groundwater for humans and livestock in rural communal systems. WRC Report No. 1175/1/06. Water Research Commission, Pretoria. ISBN 1-77005-467-7.
- CASEY NH, MEYER JA and COETZEE CB (1998a) An investigation into the quality of water for livestock production with the emphasis on subterranean water and the development of a water quality guideline index system. Volume 1: Development and modelling. WRC Report No. 644/1/98. Water Research Commission, Pretoria. ISBN 1-86845-739-0.
- CASEY NH, MEYER JA and COETZEE CB (1998b) An investigation into the quality of water for livestock production with the emphasis on subterranean water and the development of a water quality guideline index system. Volume 2: Research results. WRC Report No. 644/2/98. Water Research Commission, Pretoria. ISBN 1-86845-380-4.
- CASEY NH, MEYER JA and COETZEE CB (1998c) An investigation into the quality of water for livestock production with the emphasis on subterranean water and the development of a water quality guideline index system. Volume 3: Appendix. WRC Report No. 644/3/98. Water Research Commission, Pretoria. ISBN 1-868457-381-2.
- CASEY NH, MEYER JA and COETZEE CB (2001) An extension to

- and further refinement of a water quality guideline index system for livestock watering. Volume 2: Poultry production systems and water quality for ostrich production. WRC Report No. 857/2/01. Water Research Commission, Pretoria. ISBN 1-86845-714-1.
- COETZEE CB, CASEY NH and MEYER JA (1997) Fluoride tolerance in layers. *Br. Poult. Sci.* **38** 597-602.
- COETZEE CB, CASEY NH and MEYER JA (2000) The effect of water-borne fluoride on the production of laying hens. *Water SA* **26** (1) 115-118.
- DAVIES TC and MUNDALAMO HR (2010) Environmental health impacts of dispersed mineralisation in South Africa. *J. Afr. Earth Sci.* **58** 652-666. <https://doi.org/10.1016/j.jafrearsci.2010.08.009>
- DEANGELO AB, GEORGE MH, KILBURN S, MOORE TM and WOLF DC (1998) Carcinogenicity of potassium bromate administered in the drinking water to male B6C3F1 mice and F344/N rats. *Toxicol. Pathol.* **26** 587-594. <https://doi.org/10.1177/019262339802600501>
- DU TOIT J and CASEY NH (2012) Iodine as an alleviator of bromine toxicity in thyroid, liver and kidney of broiler chickens. *Livest. Sci.* **144** 269-274. <https://doi.org/10.1016/j.livsci.2011.12.011>
- EUROPEAN PARLIAMENT (2005) Regulation (EC) No 396/2005 of the European Parliament and of the Council of 23 February 2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin and amending Council Directive 91/414/EEC. *Official Journal of the European Union*, 2005R0396 - EN - 26.10.2012 - 010.001. https://doi.org/10.1007/978-1-137-54482-7_24
- EZEKWE IC, ODU NN, CHIMA GN and OPIGO A (2012) Assessing regional groundwater quality and its health implications in the Lokpaukwu, Lekwesi and Ishiagu mining areas of southeastern Nigeria using factor analysis. *Environ. Earth Sci.* **67** 971-986. <https://doi.org/10.1007/s12665-012-1539-9>
- JAIN A, CHAURASIA A, SAHASRABUDDHEY B and VERMA KK (1996) Determination of bromide in complex matrices by pre-column derivatization linked to solid-phase extraction and high-performance liquid chromatography. *J. Chromatogr. A* **746** (1) 31-41. [https://doi.org/10.1016/0021-9673\(96\)00317-2](https://doi.org/10.1016/0021-9673(96)00317-2)
- KORSTEN L, CASEY NH AND CHIDAMBA L (2016) Evaluation of the risks associated with the use of rooftop harvesting and groundwater for domestic use and livestock watering: Volume 2: WRC Report No. 2175/2/16. Water Research Commission, Pretoria. ISBN 978-1-4312-0875-3.
- LOEBER JG, FRANKEN MAM and VAN LEEUWEN FXR (1983) Effect of sodium bromide on endocrine parameters in the rat as studied by immunocytochemistry and radioimmunoassay. *Fd. Chem. Toxic.* **21** (4) 391-404. [https://doi.org/10.1016/0278-6915\(83\)90093-5](https://doi.org/10.1016/0278-6915(83)90093-5)
- MAGAZINOVIC RS, NICHOLSON BC, MULCAHY DE and DAVID DE (2004) Bromide levels in natural waters: its relationship to both levels of chloride and total dissolved solids and the implications for water treatment. *Chemosphere* **57** 329-335. <https://doi.org/10.1016/j.chemosphere.2004.04.056>
- MAMABOLO MC, CASEY NH and MEYER JA (2009) Effects of total dissolved solids on the accumulation of Br, As and Pb from drinking water in tissues of selected organs in broilers. *S. Afr. J. Anim. Sci.* **39** (1) 169-172. <https://doi.org/10.4314/sajas.v39i1.61227>
- MEYER JA (2005) Project: Analyse borehole water for domestic use and livestock watering throughout the Republic of South Africa for a period of one year. Department of Agriculture Report No. 032005/02/26. Department of Agriculture, Pretoria.
- MEYER JA (2015) Animal health assessment. In: Dabrowski JM (ed.) Investigation of the contamination of water resources by agricultural chemicals and the impact on environmental health Volume 1: Risk assessment of agricultural chemicals to human and animal health. WRC Report No. 1956/1/15. Water Research Commission, Pretoria. ISBN 978-1-4312-0711-4.
- MEYER JA and CASEY NH (2004) Exposure assessment of potentially toxic trace elements in indigenous goats in the rural communal production systems of the northern region of South Africa. *S. Afr. J. Anim. Sci.* **34** (Suppl 1) 219-222.
- MEYER JA, CASEY NH and MYBURGH J (2000) The influence of geochemistry on health risks to animals and humans in geographically localised livestock production systems. *S. Afr. J. Anim. Sci.* **30** (Supplement 1) 82-84. <https://doi.org/10.4314/sajas.v30i4.3920>
- MOORE MM and CHEN T (2006) Mutagenicity of bromate: Implications for cancer risk assessment. *Toxicology* **221** 190-196. <https://doi.org/10.1016/j.tox.2005.12.018>
- NRC (2005) Bromine In: Klasing KC (ed.) *Mineral Tolerance of Animals: Second revised edition*. Committee on Minerals and Toxic Substances in Diets and Water for Animals, National Research Council, Washington, DC.
- PAVELKA S, BABICKÝ A, VOBECKÝ M, LENER J and ŠVANDOVÁ E (2000) Bromide kinetics and distribution in the rat. I. Biokinetics of ⁸²Br-Bromide. *Biol. Trace Elem. Res.* **76** 57-66. <https://doi.org/10.1385/bter:76:1:57>
- PLANT JA, BALDOCK JW and SMITH B (1996) The role of geochemistry in environmental and epidemiological studies in developing countries: a review. In: Appleton JD, Fuge R and McCall GJH (eds.) *Environmental Geochemistry and Health Geological Society Special Publication No. 113*. Geological Society, London. <https://doi.org/10.1144/gsl.sp.1996.113.01.02>
- USEPA (United States Environmental Protection Agency) (1997) Special report on environmental endocrine disruption: An effects assessment and analysis. Report No EPA/630/R-96/012. United States Environmental Protection Agency, Washington D.C. <https://doi.org/10.2172/569107>
- USEPA (United States Environmental Protection Agency) (2005) *Domestic Drinking Water Quality Criteria*. National Technical Information Service. United States Environmental Protection Agency, Washington D.C. <https://doi.org/10.2172/569107>

Organophosphorus flame retardants in surface and effluent water samples from the Vaal River catchment, South Africa: levels and risk to aquatic life

TB Chokwe^{1*} and SM Mporetji¹

¹Scientific Services, Rand Water, P.O. Box 3526, Vereeniging, 1930, South Africa

ABSTRACT

The occurrence and risk assessment of seven organophosphorus flame retardants (OPFRs) in surface water samples within the Vaal River catchment in South Africa were investigated. Wastewater treatment works effluents as the potential sources of OPFRs in surface water were also analysed. In surface water, tris-(chloro-propyl)-phosphate (TCPP) – the total of the three TCPP isomers studied, and tris-(2, 3 dibromo-propyl)-phosphate (TDBPP) were the most abundant OPFRs, with mean concentrations of 276 ng/L and 227 ng/L; respectively. In effluent water samples, the most abundant OPFR was TCPP with a mean concentration of 700 ng/L. A high detection frequency (> 80%) was observed for six of the seven OPFRs with tris-(1, 3- dichloro-propyl)-phosphate (TDCPP) detection frequency being the lowest at 17%. Assessment of risk to aquatic organisms using risk quotients based on measured environmental concentrations (MEC) and predicted no-effect concentrations (PNEC) ranged from no significant risk (for algae, daphnia and fish) to low potential for adverse effects (for algae and fish).

Keywords: organophosphorus flame retardants, water samples, concentrations, risk, Vaal River

INTRODUCTION

Organophosphorus compounds (OPs) are esters of phosphoric acid that are increasingly used as plasticizers and by the flame retardant industry to replace regulated brominated diphenyl ethers (pentaBDE, octaBDE and decaBDE) (Kim et al., 2013; Wei et al., 2015). OP production and application has thus been increasing in recent years (Morris et al., 2014) leading to their widespread occurrence in the environment, including air (Shoeib et al., 2014), dust (Abafe and Martincigh, 2014), river water (Cristale et al., 2013), drinking water (Li et al., 2014), biota (Gao et al., 2014), soil (Eguchi et al., 2013) and biosolids (Zeng et al., 2014a, 2014b). OPs have been used as additives in furniture, textile coating, upholstery, electronics, paints, polyvinyl chloride (PVC) plastics, polyurethane foams, lubricants and hydraulic fluids (Abafe and Martincigh, 2014; Wei et al., 2015; Zeng et al., 2014b). The chlorinated alkylphosphates such as tris (1-chloro-2-propyl) phosphate (TCPP), tri (3-chloropropyl) phosphate (3-TCPP), tris (1,3-dichloro-2-propyl) phosphate (TDCPP) and tris (2-chloroethyl) phosphate (TCEP) are used mainly in polyurethane foams, electronic equipment, textiles, plastic and building materials, while the non-derivatized organophosphates such as triphenyl phosphates (TPP) are mainly used as lubricants and to regulate pore sizes such as in concrete (Marklund et al., 2005; EFRA, 2006; Van der Veen and De Boer, 2012; Kim et al., 2013; Abafe and Martincigh, 2014). The presence of both chlorine and phosphorus is advantageous for the optimum non-flammability, working in both the solid and gaseous phases.

Ongoing toxicological studies have shown several toxic effects of these compounds, such as the potential for ecological and human health concerns of neurotoxin and carcinogenic nature (Abafe and Martincigh, 2014; Cristale et al., 2013). The extent and magnitude of OPs occurrence in the environment, combined with striking structural similarity to toxic organophosphorus pesticides, has led to public concern over

risks posed by these substances (Oliveri et al., 2015; Wei et al., 2015). TCEP has been found to have teratogenic and haemolytic effects and has carcinogenic potential in rats and mice (Beth-Hubner, 1999; Sato et al., 1997). TPP was found to show adverse biological effect (Lin, 2009) and can cause contact dermatitis in humans and is a potent inhibitor of human carboxyl esterase (WHO, 1991). One study found that, although developmental exposure to TDCPP did not cause behavioural effects or alter monoamine levels in larval zebrafish, females exposed to TDCPP through to adulthood showed depressed levels of both dopamine and serotonin later in life (Wang et al., 2015). Correspondingly, in vivo work has shown that TDCPP can affect cellular function in the PC12 cell line and increase its differentiation into both cholinergic and dopaminergic cell types (Dishaw et al., 2011).

The environmental concerns associated with the use of these compounds as flame retardants is that they have the potential to be released into surface water bodies either directly or via industrial and wastewater treatment works (WWTWs) discharges (Meyer and Bester, 2004), and from atmospheric depositions (Möller et al., 2011). In this regard they have been found in water at levels of ng to µg/L (Cristale et al., 2013; Wang et al., 2011). The levels and distribution of contaminants in the surface and effluent water samples could reflect the pollution status in the environment. Thus, in the present study, 15 representative surface water samples located in the Vaal River catchment and 6 WWTWs effluents were selected as target sites to gain insight regarding levels and risk assessment of OPs in the aquatic environment.

MATERIALS AND METHODS

Chemicals and reagents

Approximately, 1.2 mL each of tris (2-chloroethyl) phosphate (TCEP), tris (1-chloropropyl) phosphate (1-TCPP), tris (3-chloropropyl) phosphate (3-TCPP), tris (1-chloro-2-propyl) phosphate (TCPP), tris (1, 3-dichloro-2-propyl) phosphate (TDCPP), triphenyl phosphate (TPP), tris (2, 3 dibromopropyl) phosphate (TDBPP) of analytical grade were purchased from

*Corresponding author, email: tlouchokwe6@gmail.com

Received 4 April 2018; accepted in revised form 2 July 2019

AccuStandard; USA. The solvents acetone, dichloromethane (DCM), methanol (MeOH), ethyl acetate (EtOAc), methyl *tertiary*-butyl ether (MtBE) and hexane used in the study were of GC grade purchased from Merck, Johannesburg, South Africa, and were used without further purification. Strata X (33 and 100 μm), C₁₈ and styrene divinyl benzene (SDVB) cartridges were purchased from Separations, Johannesburg, South Africa. PestiCarb cartridges were purchased from Chemetrix, Johannesburg, South Africa. Helium as He 5.5 pure was purchased from Air Products, Vereeniging, South Africa.

Sample collection

Water samples were collected from the Vaal River catchment within the Free State (Sites 1-4), Gauteng (Sites 5-7) and Mpumalanga (Sites 8-14) Provinces of South Africa, as shown in Figure 1. The selected sampling sites were classified as hotspots for pollution because of activities within the catchment area (Chokwe et al., 2017). Some of the key activities within the catchment include flows from wastewater treatment works, stock farming, and irrigation agriculture, as well as an overall increase in human population within the sub-catchment area. Water samples (of 2.5 L) were collected from each of the 21 sampling sites during October – December 2017.

Extraction of analytes from simulated water samples

Solid-phase extraction (SPE) was used as the isolation technique throughout the experiment. Five different types of SPE cartridges (Strata X 33 μm , Strata X 100 μm , PestiCarb, C18 and SDVB) were tested for the pre-concentration of halogenated organophosphorus

and un-halogenated organophosphorus compounds from water samples. Satisfactory results were obtained with SDVB cartridges with elution by 3×2 mL EtOAc: MtBE (4:1). Before use, the SPE cartridge was conditioned with 12 mL of hexane, 12 mL of DCM, 12 mL of MeOH and 12 mL of MilliQ water. About 500 mL of MilliQ water preserved with 5 mL MeOH was spiked with 100 μL of organic OPFRs standard mixture and extracted at a flow rate of approximately 10 mL/min. After passing the sample through the cartridge, the cartridge was dried under vacuum for 1 h. The compounds were eluted with 3×2 mL of mixture of (4: 1) EtOAc: MtBE. The eluates were evaporated to dryness under a gentle stream of nitrogen at 40°C. Parathion D10 (200 ng/L) as an internal standard was added and the volume made up to 200 μL with ethyl acetate (EtOAc). 2 μL of the extracts and internal standards was injected into the GC-MS.

Extraction of analytes from water samples

The optimized procedure obtained using the simulated water samples was used to isolate the targeted compounds from surface and effluent water samples within the Vaal River catchment. During the analysis of both types of water samples, 500 mL of surface water and 500 mL of effluent water samples were spiked with 100 μL of OPFRs standard mixture in triplicate to check the robustness of the developed method.

Instrumentation and GC-MS conditions

An Agilent 6890 GC equipped with 5975 mass selective detector (MSD) was used for GC-MS analysis. The GC was equipped with an Agilent 7683 Series Injector autosampler. The injection

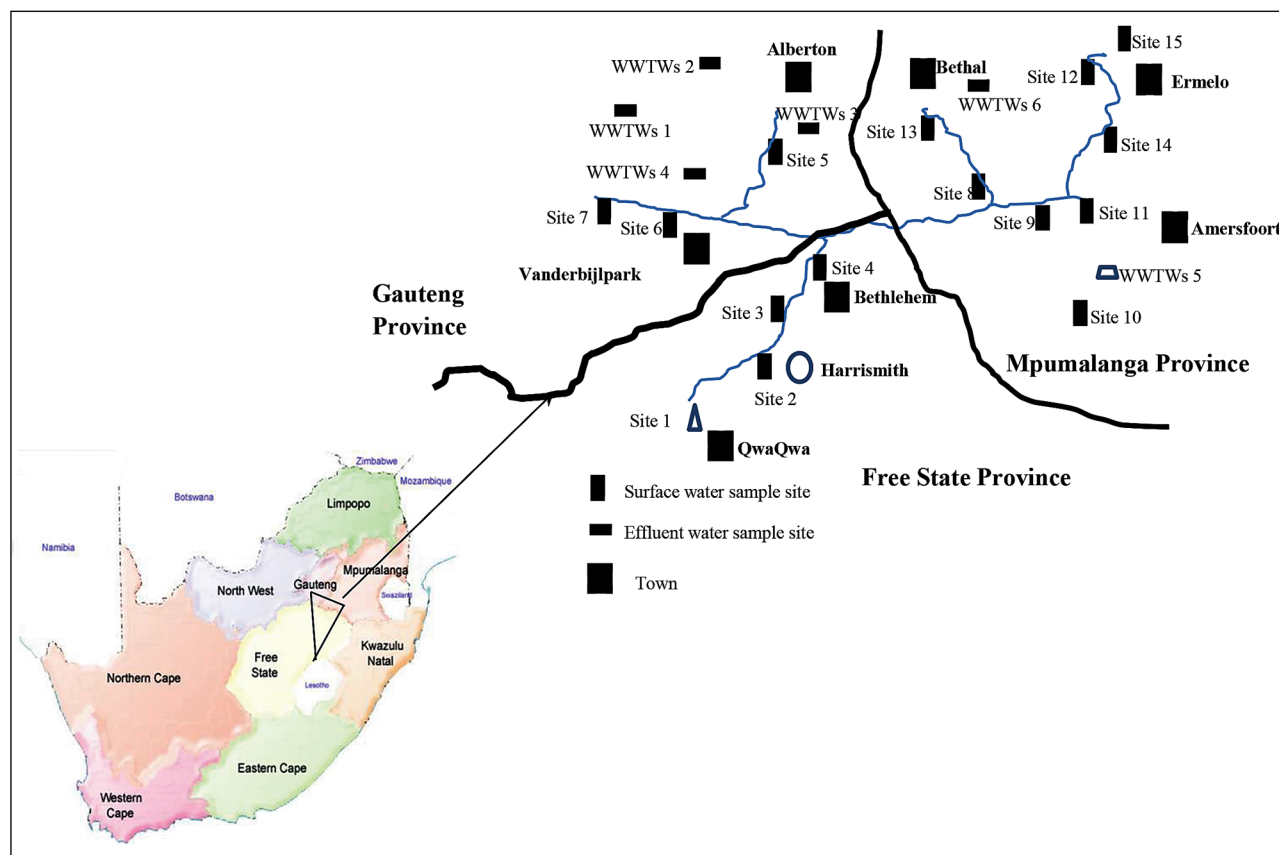


Figure 1. Map of South Africa (bottom) with expanded sampling sites (top)

port was fitted with a SGE split/splitless liner; single taper with quartz wool 4 mm ID deactivated inlet liner (Restek, for Agilent GCs). The GC separation was conducted on a Restek RTx-1614, capillary column (film thickness 0.10 µm, 15 m × 0.25 mm I.D), (Chromspec cc South Africa)). The injections were made in splitless mode with the injector temperature set at 280°C. The injection volume was 2 µL. The GC temperature programme conditions were as follows: initial temperature 90°C, heated to 200°C by a temperature ramp of 20°C/min then 230°C by a temperature ramp of 3°C/min then finally heated to 300°C (held for 4 min) by a temperature ramp of 60°C/min. Helium was used as carrier gas at a flow rate of 1.3 mL/min and a constant linear velocity of 60 cm/s. For the MS, the ion source and transfer line temperatures of 150 and 300°C, respectively, and ionization energy of 70 eV were used. The monitored m/z ions for each compound were obtained in full scan mode within the range 200–800 AMU with data acquired with ChemStation software (Agilent Technologies, Santa Clara, CA, USA).

Quality assurance

Since certified reference materials were not available, a spiking method was used to check the recoveries of analytes in the water samples. Recoveries were obtained by spiking 500 mL water sample with OPFRs standard mixture (200 ng/L (3-TCPP and TPP), 400 ng/L (TCEP, 1-TCPP and TCPP) and 1 000 ng/L (TDCPP and TDBPP)). The spiked water sample was taken through the same extraction procedure prior to GC-MS analysis. TCEP, 1-TCPP and TCPP were detected above their respective LOQs in the blank samples (i.e. MilliQ water). The percentage recoveries in the spiked MilliQ water samples after background corrections ranged from 74.60 ± 1.18 (TCPP) – 108.67 ± 9.19 (TDBPP). Prior to the analysis of water samples, 100 ng/L Chlorpyrifos D10 was added into each sample as surrogate standard. The recovery of Chlorpyrifos D10 from all the samples ranged from 65–97%. Several other quality assurance measures were routinely checked in this study and included analysing blanks in between samples, analysing a quality control sample after every 7 samples as well as analysing samples in triplicates.

The quantification of OPFRs was accomplished using internal standard method by relating the responses of OPFRs to the response of Parathion D10. The response factors were determined from the slope of a plot of the ratios of peak areas against the concentration levels. The values of plots were obtained from a 5-point analysis of the OPFRs standard solutions in the concentration range of 80–400 ng/L (3-TCPP and TPP); 160–800 ng/L (TCEP, 1-TCPP and TCPP) and 400–2 000 ng/L (TDCPP and TDBPP). Limit of detection (LOD) was defined as a signal/noise ratio of 3 while the limit of quantification was defined as signal/noise ratio of 10. For

the compounds detected in blank samples, (TCEP, 1-TCPP and TDCPP), the LOD was calculated as mean blanks values plus 3 times the standard deviation. The following descriptive statistics: regression, sum, mean, median, minimum and maximum were calculated with Microsoft Excel 2010. Samples below LOD were treated as zero throughout the statistical analysis. The concentrations of analytes were expressed as ng/L and were not recovery corrected. All glassware was cleaned with laboratory wash solution, rinsed with distilled water, acetone and hexane.

Risk assessment towards aquatic organisms

In this study, risk assessment using risk quotients (RQs) for non-target organisms, as described by Cristale et al. (2013), were adopted. The RQs were calculated as the quotient of measured environmental concentration and the predicted no effect concentration (PNEC) using Eq. 1. PNEC was estimated as a quotient of toxicological relevance concentration and a security factor (*f*) using Eq. 2.

$$RQ = MEC / PNEC \quad (1)$$

$$PNEC = L(E)C_{50} / f \quad (2)$$

where RQ is the risk quotient, MEC is the measured environmental concentration, PNEC is the predicted no effect concentration, LC_{50} is the lethal concentration required to kill 50%, EC_{50} is the concentration of a toxicant which induces a response halfway between the baseline and maximum after a specified exposure time, and *f* is the security factor.

For RQ calculations, the lowest $L(E)C_{50}$ for fish, *Daphnia* and algae associated with TCEP, TCPP, TDCPP, TPP and TDBPP and a factor of 1 000 were used (EC, 2003; Cristale et al., 2013; Shi et al., 2016). Table 1 presents the acute toxicity used for risk assessment.

For data interpretation, the maximum probable risk for ecological effect from contaminated water as prescribed (Wentsel et al., 1996; Cristale et al., 2013) was followed. $RQ < 1$ indicates no significant risk; RQ between 1 and 10 indicates a small potential for adverse effects, RQ s between 10 and 100 indicate a significant potential for adverse effects while RQ s > 100 indicate that potential adverse effects should be expected.

RESULTS AND DISCUSSION

Method performance and validation

The recoveries of the spiked MilliQ water ranged from 74–108%, for the surface water sample after background concentration correction ranged from 61–92%, and the

Table 1. Acute toxicity (LC_{50} or EC_{50}) used for risk assessment for fish, *Daphnia* and algae

Compound	Fish			Algae			<i>Daphnia</i>		
	$L(E)C_{50}$ (mg/L)	Species	Ref	$L(E)C_{50}$ (mg/L)	Species	Ref	$L(E)C_{50}$ (mg/L)	Species	Ref
TCEP	90	<i>Carassius auratus</i>	Verbruggen et al., 2005	51	<i>Scenedesmus subspicatus</i>	Verbruggen et al., 2005	330	<i>Daphnia magna</i>	Verbruggen et al., 2005
TCPP	30	<i>Poecilia reticulata</i>	Verbruggen et al., 2005	45	<i>Scenedesmus subspicatus</i>	Verbruggen et al., 2005	91	<i>Daphnia magna</i>	Verbruggen et al., 2005
TDCPP	5.1	<i>Carassius auratus</i>	Verbruggen et al., 2005	39	<i>Pseudokirchneriella subcapitata</i>	Verbruggen et al., 2005	4.2	<i>Daphnia magna</i>	Verbruggen et al., 2005
TDBPP	0.516	<i>Salmo gairdneri</i>	NICNAS, 2005	0.545	<i>Scenedesmus abundans</i>	NICNAS, 2005	4.568	<i>Daphnia magna</i>	NICNAS, 2005

recoveries from the spiked WWTWs effluent ranged from 57–81% with the RSD from all the water samples below 20%. Except for the TDBPP results, the recoveries from all the samples were satisfactory, with the highest recoveries obtained from both MilliQ and surface water samples and a slightly lower recovery obtained from the WWTW effluent samples. The recovery of TDBPP was the highest from MilliQ water (108%) followed by surface water at 61%, with recovery from the WWTW effluent sample at 57%. The recovery results are summarised in Table 2.

In order to evaluate the experimentally found optimum extraction conditions, MilliQ water was spiked at low and higher concentrations with the OPFRs standard mixture. At low concentration, the percentage recovery ranged from 66.6 ± 13.4 (TPP) – 128.7 ± 1.1 (TDBPP) while at higher concentration spike the percentage recoveries ranged from 60.8 ± 13.3 (TCPP) – 67.2 ± 17.3 (TDBPP) as shown in Table 3. Although the recoveries were lower as compared to the recoveries from lower concentration spikes, the results may indicate that at higher loading of the analytes there may be saturation on the cartridge. However, as the concentrations obtained in the current study were much lower than the concentration at higher loading, the results may be accepted. The RSD of the analytes from the high spiking concentration were all below 20%, indicating a good repeatability of the extraction procedure. The calibration curves were linear ($r^2 > 0.975$) across the concentration range as shown in Table 3. LOD and LOQ ranged from 4.2–24.0 ng/L and 13.9–80.3 ng/L, respectively.

Concentrations of OPFRs in surface water samples

OPFRs were detected at all the sampling sites with detection frequency (%DF) decreasing as follows: TDBPP (93.3%) > TDCPP (73.3%) > TCPP (66.7%) > 1-TCPP (60%) > TCEP

(46.7%) > 3-TCPP (40%) > TPP (0%). Table 4 presents the contamination levels of OPFRs within the Vaal River. Low detection of TPP in surface water samples may be attributed to the fact that it has a higher $\log K_{ow}$ (i.e., 4.59) (Reemtsma et al., 2008) indicating its affinity to organic carbon. Also, a study by Marklund et al. (2005) indicated that TPP has an elimination rate of around 60% during wastewater treatment.

The mean concentrations of OPFRs within the Vaal River followed the following pattern: TDBPP (228 ng/L) > TCPP (149 ng/L) > TDCPP (116 ng/L) > 1-TCPP (102 ng/L) > TCEP (31 ng/L) > 3-TCPP (25 ng/L). The sum concentrations of OPFRs (Σ_7 OPFRs) showed a 100% detection frequency from all the sampling sites with the concentrations ranging from 90–1 424 ng/L (mean 650 ng/L). Site 15 and Site 12 were the most polluted sampling sites within the Vaal River with concentrations of 1 424 and 1 360 ng/L, respectively. These sampling sites are most influenced by industries present in the area. Another site that exhibited high concentrations of OPFRs was Site 13 which receives the effluent discharges from the Bethal area. The least polluted site was Site 10 with only TDCPP detected. The Σ_7 OPFRs from Site 4, Site 6, Site 8 and Site 11 were 267, 314, 306 and 330 ng/L respectively. High variation in the concentrations of OPFRs at each site may be due to different industries in the vicinity together with limited sampling sites, especially in the Free State and Gauteng Provinces.

In comparison to the concentrations of OPFRs in surface water, the mean concentrations of TCEP (31 ng/L) were lower than the mean concentration obtained in Spain (85 ng/L) while the mean concentration of TDCPP (116 ng/L) obtained in our study was higher than the mean concentration (70 ng/L) obtained in surface water from Spain (García-López et al., 2010). Martínez-Carballo et al. (2007) reported concentration ranges of 13–130 ng/L and < LOQ–19 ng/L for TCEP and TDCPP; respectively, in surface water from Austrian rivers.

Table 2. Recoveries of OPFRs in MilliQ water, surface water and WWTWs effluent water samples ($n = 3$)

Compound	Spiked MilliQ water		Surface water spiked		WWTW effluent spiked	
	Recovery (%)	RSD (%)	Recovery (%)	RSD (%)	Recovery (%)	RSD (%)
TCEP	89.9	1.9	82.2	6.9	70.1	6.5
3-TCPP	77.1	1.7	75.9	5.9	81.7	10.2
1-TCPP	76.8	1.6	81.2	5.6	73.3	6.0
TCPP	74.6	1.2	92.0	14.5	85.6	12.2
Chlorpyrifos D10	84.8	0.9	68.6	4.6	68.2	13.8
TDCPP	81.6	4.9	67.9	9.7	64.0	6.0
TPP	77.0	6.5	81.2	5.4	68.3	11.9
TDBPP	108.7	9.2	61.8	5.4	57.2	4.8

Table 3. Recoveries and RSD of OPFRs in MilliQ water ($n = 3$) and some of the validation parameters

Compound	Spiked MilliQ water				Validation parameter			
	Level 1		Level 2		Calibration range (ng/L)	Linearity (r^2)	LOD (ng/L)	LOQ (ng/L)
	Recovery (%)	RSD (%)	Recovery (%)	RSD (%)				
TCEP	87.2	0.8	64.4	9.1	160 - 800	0.991	7.0	23.5
3-TCPP	66.9	6.7	61.2	9.7	80 - 400	0.999	4.2	13.9
1-TCPP	73.4	12.4	61.9	13.5	160 - 800	0.999	9.1	31.1
TCPP	77.5	2.5	60.8	13.3	160 - 800	0.979	10.6	35.4
Chlorpyrifos D10	67.6	3.9	73.9	13.1	40 - 200	N/A	N/A	N/A
TDCPP	92.1	1.5	61.6	5.7	400 -2000	0.994	13.8	45.9
TPP	66.6	13.4	62.3	3.9	80 - 400	0.997	6.2	20.4
TDBPP	128.7	1.4	67.2	17.3	400 - 2000	0.975	24.0	80.3

N/A: not applicable

Level 1: 100 ng/L (3-TCPP and TPP), 200 ng/L (TCEP, 1-TCPP and TCPP) and 500 ng/L (TDCPP and TDBPP) were spiked into 500 mL of MilliQ water
Level 2: 340 ng/L (3-TCPP and TPP), 680 ng/L (TCEP, 1-TCPP and TCPP) and 1700 ng/L (TDCPP and TDBPP) were spiked into 500 mL of MilliQ water

Table 4. Mean concentrations of OPFRs within the Vaal River catchment

Site	TCEP (ng/L)	3-TCPP (ng/L)	1-TCPP (ng/L)	TCPP (ng/L)	TDCPP (ng/L)	TPP (ng/L)	TDBPP (ng/L)	Σ_7 OPFRs (ng/L)
Site 1	nd	16.0	20.2	96.1	116.2	nd	270.4	518.9
Site 2	37.4	nd	19.9	0.010	54.7	nd	663.9	775.6
Site 3	nd	61.0	265.8	289.1	nd	nd	134.7	750.6
Site 4	nd	nd	nd	21.7	47.6	nd	198.1	267.4
Site 5	70.0	110.0	110.0	198.2	130.0	nd	170.0	788.2
Site 6	40.0	nd	0.0	nd	130.0	nd	143.6	313.6
Site 7	60.0	60.0	40.0	120.0	140.0	nd	189.3	609.3
Site 8	nd	nd	nd	nd	93.6	nd	212.5	306.1
Site 9	nd	nd	70.3	334.1	nd	nd	277.3	681.7
Site 10	nd	nd	nd	nd	90.0	nd	nd	90.0
Site 11	nd	nd	nd	50.0	100.0	nd	180.0	330.0
Site 12	235	34.3	34.8	165.3	624.4	nd	266.7	1360.6
Site 13	20.5	91.4	398.0	389.1	nd	nd	116.1	1015.1
Site 14	nd	nd	nd	nd	212.2	nd	301.7	513.9
Site 15	nd	nd	568.3	566.2	0.0	nd	289.3	1423.8
%DF	46.7	40	60	66.7	73.3	nd	93.3	100
Range	nd – 235	nd – 110	nd – 568	nd – 566	nd – 624	nd	nd – 664	90 – 1424
Mean	31	25	102	149	116	nd	228	650
Median	0.0	0.0	20	96	94	nd	198	609

Bold: results that are discussed in the report; DF: detection frequency; nd: not detected

Higher concentration, ranging from 113–26 050 ng/L for TCEP, were reported in surface water from the River Aire in the UK (Cristale et al., 2013). In the same study, a TCEP range of 119–316 ng/L was observed, which was similar to the results obtained in our study. Concentration ranges of 38–3 700 ng/L were reported from Songhua River, China (Wang et al., 2011). Shi et al. (2016) reported ranges of < LOD–2 072 ng/L (mean 291 ng/L) and < LOD–855 (mean 46.3 ng/L) for TCEP and TDCPP in surface water from Beijing respectively. The TDCPP results were similar to the results obtained in our study while the TCEP values were almost 10 times higher than the results obtained in this study.

Concentrations of OPFRs in WWTWs effluents

The detection frequency of OPFRs in final effluents of WWTWs were as follows: 1-TCPP/ TCEP (100%) > TCEP/ 3-TCPP/ TDBPP (83.3%) > TDCPP (16.7%) > TPP (0%). The mean concentrations (997 ng/L) of OPFRs in effluent samples were higher than the mean concentrations (650 ng/L) obtained in

surface water, indicating that the presence of OPFRs in surface water may be attributed to effluents from WWTWs. Table 5 presents the concentrations of OPFRs in effluents samples from WWTWs.

The summed concentrations of OPFRs ranged from 345–2073 ng/L in effluent samples. The highest Σ_7 OPFRs concentrations were detected in WWTW6 which treats both the domestic and industrial wastewaters from the Bethal area in Mpumalanga. It was followed by WWTW3 at 1 168 ng/L which also treats both domestic and industrial wastewaters in the Gauteng region. In all the effluents analysed, 1-TCPP and TCEP were the most abundant OPFRs with mean concentrations of 300 ng/L and 364 ng/L, respectively. The higher abundance of TCEP in comparison to TDCPP may be attributed to industrial replacement of TCEP by TDCPP used in flexible foam in Europe (EC, 2003; Quednow and Püttmann, 2009). These results from our study supported the finding by Andresen et al. (2004) and Martínez-Carballo et al. (2007) indicating WWTWs effluents as a source of OPFRs in surface water.

Table 5. Concentrations of OPFRs in effluent water samples

Site	TCEP (ng/L)	3-TCPP (ng/L)	1-TCPP (ng/L)	TCPP (ng/L)	TDCPP (ng/L)	TPP (ng/L)	TDBPP (ng/L)	Σ_7 OPFRs (ng/L)
WWTW1	120.6	0.0	85.9	420.2	0.0	0.0	290.5	917.2
WWTW2	79.2	20.5	42.6	202.7	0.0	0.0	0.0	345.1
WWTW3	15.8	77.9	316.4	280.9	0.0	0.0	476.8	1 167.8
WWTW4	0.0	27.4	379.9	318.3	0.0	0.0	182	907.8
WWTW5	110.0	20.0	60.0	110.0	120.0	0.0	150.0	570.0
WWTW6	19.7	61.6	916	849.6	0.0	0.0	226.3	2 073.1
%DF	83.3	83.3	100	100	16.7	0.00	83.3	100
Min	<LOQ	<LOQ	43	110	<LOQ	<LOQ	<LOQ	345
Max	121	78	916	850	120	0.00	477	2 073
Mean	58	35	300	364	20	0.00	221	997
Median	49	24	201	300	0.00	0.00	204	912

Bold: results that are discussed in the report; DF: detection frequency; LOQ: limit of quantification

In comparison with results for effluents from other parts of the world, higher mean concentrations of TCPP in WWTWs effluents from Sweden and Germany (4 400 ng/L and 3 000 ng/L, respectively) were reported (Marklund et al., 2005; Meyer and Bester, 2004). TCEP concentration (350 ng/L) was reported by Marklund et al. (2005) in effluents from Sweden. From Spanish wastewater samples, a similar concentration range for TCPP (290–680 ng/L) was reported (Garcla-López et al., 2010). The TCPP concentration results from our study were comparable to the TCPP concentration (290–1 400 ng/L) reported from Austrian WWTP effluents (Martínez-Carballo et al., 2007). TCEP, TCPP, TDCPP and TDBPP were detected at concentrations of 133 ng/L, 440 ng/L, 227 ng/L and not detected; respectively, in WWTWs effluents from Canada (Woudneh et al., 2015).

Risk to aquatic life

In this study, the risk assessment for aquatic organisms was estimated for the detected OPFRs, including, TCEP, TCPP (total of 3-TCPP, 1-TCPP and TCPP), TDCPP and TDBPP in surface water samples. No significant risk ($RQ < 1$) was observed within the Vaal River catchment associated to TCEP, TCPP, TDCPP and TDBPP for any of the three organisms. RQs for TCEP, TCPP, TDCPP and TDBPP varied from 0.0006–0.418 for algae; from 0.00009–0.0499 for *Daphnia*, and from 0.0003–0.442 for fish, respectively. However, to estimate the joint effect of OPFRs within the Vaal River catchment the sum of RQs derived from each site of the detected compound were used (Cristale et al., 2013). Table 6 presents the summed RQs of the OPFRs for the chosen three aquatic organisms, following the recommendation of the Technical Guidance Document on Risk Assessment (EC, 2003) that requires at least three trophic levels from the assessed environment.

The risk assessment from Site 2 showed a low potential for adverse effects ($1.0 \leq RQ \leq 10$) for fish and algae with RQ values of 1.298 and 1.221; respectively. The rest of the sampling sites showed no significant risk ($RQ \leq 1.0$) from OPFRs for algae, *Daphnia* and fish. The summed RQ values ranged from 0.248–1.221, 0.0363–0.2106 and 0.0176–1.298 for algae, *Daphnia* and fish; respectively. Though risk assessment was low for OPFRs, more monitoring studies should be undertaken

Table 6. Summed RQ results for OPFRs from each sampling site within the Vaal River catchment for algae, *Daphnia* and fish

Site	Summed RQs		
	Algae	<i>Daphnia</i>	Fish
Site 1	0.527	0.0884	0.625
Site 2	1.221	0.1586	1.294
Site 3	0.261	0.0363	0.285
Site 4	0.248	0.0549	0.394
Site 5	0.326	0.0730	0.372
Site 6	0.267	0.0625	0.304
Site 7	0.357	0.0770	0.402
Site 9	0.518	0.0651	0.551
Site 10	0.002	0.0214	0.0176
Site 11	0.334	0.0244	0.370
Site 12	0.515	0.2106	0.540
Site 13	0.233	0.0351	0.254
Site 14	0.559	0.1165	0.626
Site 8	0.392	0.0688	0.430
Site 15	0.558	0.0758	0.599

as studies have shown that most of the halogenated OPFRs are not removed during wastewater treatment (Reemtsma et al., 2008) and similar processes (biodegradation and sorption) affect their concentrations and fate in the dissolved phase. In addition, recent studies have reported the presence of OPFRs in biota (Kim et al., 2011; Sundkvist et al., 2010) indicating that OPFRs can be accumulated in biota. Thus further studies on bio-accumulation and bio-magnification of OPFRs in biota from aquatic environments for comprehensive risk assessment is of utmost importance.

CONCLUSIONS

In the present study, the concentrations and risk assessment of OPFRs in surface and effluent water samples within the Vaal River catchment were reported. Six of the seven OPFRs were detected in the surface water samples with TCPP (all of the three studied isomers) and TDBPP detected at higher concentrations followed by TDCPP and TCEP. TPP was not detected in all samples. The same pattern in concentration was observed in effluent samples, albeit at high concentrations. Risk assessment based on acute toxicity data on three aquatic organisms (i.e. algae, *Daphnia* and fish) suggested no significant risk in most of the sampling sites. However, the joint effect of OPFRs derived from the sum of RQs from one site in our study indicated a low potential for adverse effect on algae and fish. Taking into account the high levels of these pollutants in WWTW effluents, long-term exposure and bioaccumulation of these OPFRs and other emerging flame retardants in the aquatic environment indicates that further studies are needed to define the environmental risk produced by these pollutants.

ACKNOWLEDGEMENTS

The authors would like to thank Rand Water – Analytical Services for providing a technical environment for the project. The authors also declare there is no conflict of interest.

REFERENCES

- ABAFE OA and MARTINCIGH BS (2014) Organophosphorus flame retardants and plasticizers in dust from automobile, homes, offices and university classrooms: implications for personal exposure via inadvertent dust ingestion. *Organohalogen Compounds* **76** 277–280.
- ANDRESEN JA, GRUNDMANN A and BESTER K (2004) Organophosphorus flame retardants and plasticizers in surface waters. *Sci. Total Environ.* **332** 155–166. <https://doi.org/10.1016/j.scitotenv.2004.04.021>
- BETH-HUBNER M (1999) Toxicological evaluation and classification of the genotoxic, carcinogenic, repro-toxic and sensitizing potential of tris-(2-chloroethyl)-phosphate. *Int. Arch. Occup. Environ. Health* **72** M17–M23. <https://doi.org/10.1007/pl00014210>
- CHOKWE TB, OKONKWO JO and NWAMADI MS (2017) Occurrence and distribution of tetrabromobisphenol A and its derivative in river sediments from Vaal River catchment, South Africa. *Emerging Contam.* **3** 121–126. <https://doi.org/10.1016/j.emcon.2017.11.001>
- CRISTALE J, KATSOYIANNIS A, SWEETMAN AJ, JONES KC and LACORTE S (2013) Occurrence and risk assessment of organophosphorus and brominated flame retardants in the river Aire (UK). *Environ. Pollut.* **179** 194–200.
- DISHAW LV, POWERS CM, RYDE IT, ROBERTS SC, SEIDLER FJ, SLOTKIN TA and STAPLETON HM (2011) Is the PentaBDE replacement, tris (1,3-dichloro-2-propyl) phosphate (TDCIPP), a developmental neurotoxicant? Studies in PC12 cells. *Toxicol. Appl. Pharmacol.* **256** (3) 281–289. <https://doi.org/10.1016/j.taap.2011.01.005>

- EGUCHI A, ISOBE T, RAMU K, TUE NM, SUDARYANTO A, DEVANATHAN G, VIET PH, TANA RS, TAKAHASHI S, SUBRAMANIAN A and TANABE S (2013) Soil contamination by brominated flame retardants in open waste dumping sites in Asian developing countries. *Chemosphere* **90** 2365–2375. <https://doi.org/10.1016/j.chemosphere.2012.10.027>
- EUROPEAN COMMISSION (EC) (2003) Technical Guidance Document on Risk Assessment in Support of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for Existing Substances, Directive 98/8/EC of the European Parliament and of the Council Concerning the Placing of Biocidal Products on the Market. Part II. <https://doi.org/10.1002/352760197x.app1>
- EUROPEAN FLAME RETARDANTS ASSOCIATION (EFRA) (2006) Phosphorous - Alkyl Halogenated Phosphate Esters. URL: http://www.flameretardants.eu/DocShareNoFrame/docs/3/FDEFAGLBNLDPLEBJDILIMBHJPIB146H7BOHKKK4N1KBO/EFRA/docs/DLS/2HalogenatedPhosphateEstersFactSheetAB-1_00.pdf (Accessed September 2017).
- GAO Z, DENG Y, YUAN W, HE H, YANG S and SUN C (2014) Determination of organophosphorus flame retardants in fish by pressurized liquid extraction using aqueous solutions and solid phase microextraction coupled with gas chromatograph-flame photometric detector. *J. Chromatogr. A* **1366** 31–37. <https://doi.org/10.1016/j.chroma.2014.09.028>
- GARCÍA-LÓPEZ M, RODRÍGUEZ I and CELA R (2010) Mixed-mode solid-phase extraction followed by liquid chromatography-tandem mass spectrometry for the determination of tri- and di-substituted organophosphorus species in water samples. *J. Chromatogr. A* **1217** 1476–1484. <https://doi.org/10.1016/j.chroma.2009.12.067>
- KEMMLEIN S, HAHN O and JANN O (2003) Emissions of organophosphate and brominated flame retardants from selected consumer products and building materials. *Atmos. Environ.* **37** 5485–5493. <https://doi.org/10.1016/j.atmosenv.2003.09.025>
- KIM JW, ISOBE T, CHANG KH, AMANO A, MANEGA RH, ZAMORA PB, SIRINGAN FP and TANABE S (2011) Levels and distribution of organophosphorus flame retardants and plasticizers in fishes from Manila Bay, the Philippines. *Environ. Pollut.* **159** 3653–3659. <https://doi.org/10.1016/j.envpol.2011.07.020>
- KIM JW, ISOBE T, SUDARYANTO A, MALARVANNAN G, CHANG KH, MUTO M, PRUDENTE M and TANABE S (2013) Organophosphorus flame retardants in house dust from the Philippines: occurrence and assessment of human exposure. *Environ. Sci. Pollut. Res.* **20** (2) 812–822. <https://doi.org/10.1007/s11356-012-1237-x>
- LI J, YU N, ZHANG B, JIN L, LI M, HU M, ZHANG X, WEI S and YU H (2014) Occurrence of organophosphorus flame retardants in drinking water from China. *Water Res.* **54** 53–61. <https://doi.org/10.1016/j.watres.2014.01.031>
- LIN K (2009) Joint acute toxicity of tributyl phosphate and triphenyl phosphate to *Daphnia magna*. *Environ. Lett.* **7** (4) 309–312. <https://doi.org/10.1007/s10311-008-0170-1>
- MARKLUND A, ANDERSSON B and HAGLUND P (2005) Organophosphorus flame retardants and plasticizers in Swedish sewage treatment plants. *Environ. Sci. Technol.* **39** 7423–7429. <https://doi.org/10.1021/es051013l>
- MARTÍNEZ-CARBALLO E, GONZÁLEZ-BARREIRO C, SITKA A, SCHARF S and GANS O (2007) Determination of selected organophosphate esters in the aquatic environment of Austria. *Sci. Total Environ.* **388** 290–299. <https://doi.org/10.1016/j.scitotenv.2007.08.005>
- MEYER J and BESTER K (2004) Organophosphate flame retardants and plasticizers in wastewater treatment plants. *J. Environ. Monit.* **6** 599–605.
- MORRIS PJ, MEDINA-CLEGHORN D, HESLIN A, KING SM, ORR J, MULVIHILL MM, KRAUSS RM and NOMURA DK (2014) Organophosphorus flame retardants inhibit specific liver carboxylesterases and cause serum hypertriglyceridemia. *Chem. Biol.* **9** 1097–1103. <https://doi.org/10.1021/cb500014r>
- MÖLLER A, XIE Z, CABA A, STURM R and EBINGHAUS R (2011) Organophosphorus flame retardants and plasticizers in the atmosphere of the North Sea. *Environ. Pollut.* **159** (12) 3660–3665. <https://doi.org/10.1016/j.envpol.2011.07.022>
- NICNAS (National Industrial Chemicals Notification and Assessment Scheme) (2005). Priority Existing Chemical Assessment Report No. 27. URL: https://www.nicnas.gov.au/___data/assets/word_doc/0020/34832/PEC27-TBBP.docx (Accessed September 2017).
- OLIVERI AN, BAILEY JM and LEVINA ED (2015) Developmental exposure to organophosphate flame retardants causes behavioral effects in larval and adult Zebrafish. *Neurotoxicol. Teratology* **52** 220–227. <https://doi.org/10.1016/j.ntt.2015.08.008>
- QUEDNOW K and PÜTTMANN W (2009) Temporal concentration changes of DEET, TCEP, terbutryn, and nonylphenols in freshwater streams of Hesse, Germany: possible influence of mandatory regulations and voluntary environmental agreements. *Environ. Sci. Pollut. Res.* **16** 630–640. <https://doi.org/10.1007/s11356-009-0169-6>
- REEMTSMA T, GARCÍA-LÓPEZ M, RODRÍGUEZ I, QUINTANA JB and RODIL R (2008) Organophosphorus flame retardants and plasticizers in water and air I. Occurrence and fate. *Trends Anal. Chem.* **27** 727–737. <https://doi.org/10.1016/j.trac.2008.07.002>
- SATO T, WATANABE K, NAGASE H, KITO H, NIKAWA M, and YOSHIOKA Y (1997) Investigation of the hemolytic effects of various organophosphoric acid triesters (OPEs) and their structure-activity relationship. *Toxicol. Environ. Chem.* **29** 277–287. <https://doi.org/10.1080/0277249709358444>
- SHI Y, GAO L, LI W, WANG Y, LIU J and CAI Y (2016) Occurrence, distribution and seasonal variation of organophosphate flame retardants and plasticizers in urban surface water in Beijing, China. *Environ. Pollut.* **209** 1–10. <https://doi.org/10.1016/j.envpol.2015.11.008>
- SHOEIB M, AHRENS L, JANTUNEN L and HARNER T (2014) Concentrations in air of organobromine, organochlorine and organophosphate flame retardants in Toronto, Canada. *Atmos. Environ.* **99** 140–147. <https://doi.org/10.1016/j.atmosenv.2014.09.040>
- SUNDKVIST AM, OLOFSSON U and HAGLUND P (2010) Organophosphorus flame retardants and plasticizers in marine and fresh water biota and in human milk. *J. Environ. Monit.* **12** 943–951. <https://doi.org/10.1039/b921910b>
- VAN DER VEEN I and DE BOER J (2012) Phosphorus flame retardants: properties, production, environmental occurrence, toxicity and analysis. *Chemosphere* **88** 1119–1153. <https://doi.org/10.1016/j.chemosphere.2012.03.067>
- VERBRUGGEN EMJ, RILA JP, TRAAS TP, POSTHUMA-DOODEMAN, CJAM and POSTHUMUS R (2005) Environmental risk limits for several phosphate esters, with possible application as flame retardant. RIVM Report 601501024/2005. National Institute for Public Health and the Environment, Bilthoven.
- WANG Q, LAM JCW, MAN YC, LAI NLS, KWOK KY, GUO YY, LAM PK and ZHOU B (2015) Bioconcentration, metabolism and neurotoxicity of the organophosphorus flame retardant 1,3-dichloro 2-propyl phosphate (TDCPP) to Zebrafish. *Aquat. Toxicol.* **158** 108–115. <https://doi.org/10.1016/j.aquatox.2014.11.001>
- WANG XW, LIU JF and YIN YG (2011) Development of an ultra-high-performance liquid chromatography-tandem mass spectrometry method for high throughput determination of organophosphorus flame retardants in environmental water. *J. Chromatogr. A* **1218** 6705–6711. <https://doi.org/10.1016/j.chroma.2011.07.067>
- WEI GL, LI DQ, ZHUO MN, LIAO YS, XIE ZY, GUO TL, LI JJ, ZHANG SY and LIANG ZQ (2015) Organophosphorus flame retardants and plasticizers: sources, occurrence, toxicity and human exposure. *Environ. Pollut.* **196** 29–46. <https://doi.org/10.1016/j.envpol.2014.09.012>
- WENTSEL RS, LAPOINT TW, SIMINI M, CHECKAIL RT, LUDWIG D and BREWER, L (1996) Tri-service Procedural Guidelines for Ecological Risk Assessment. US Army Edgewood Research, Development, and Engineering Center, Aberdeen Proving Ground, MD. <https://doi.org/10.21236/ada427056>
- WHO (World Health Organization) (1991) Environmental Health Criteria 111. Triphenyl phosphate. WHO, Geneva.
- WOUNDNEH MB, BENSKIN JP, WANG G, GRACE R, HAMILTON MC and COSGROVE JR (2015) Quantitative determination of 13 organophosphorus flame retardants and plasticizers in a wastewater treatment system by high performance liquid

- chromatography tandem mass spectrometry. *J. Chromatogr. A* **1400** 149–155. <https://doi.org/10.1016/j.chroma.2015.04.026>
- ZENG L, YANG R, ZHANG Q, ZHANG H, XIAO K, ZHANG H, WANG Y, LAM PK and JIANG G (2014a) Current levels and composition profiles of emerging halogenated flame retardants and dehalogenated products in sewage sludge from municipal wastewater treatment in China. *Environ. Sci. Technol.* **48** 12586–12594. <https://doi.org/10.1021/es503510q>
- ZENG X, HE L, CAO S, MA S, YU Z, GUI H, SHENG G and FU J (2014b) Occurrence and distribution of organophosphorus flame retardants/ plasticizers in wastewater treatment plant sludge from the Pearl River Delta, China. *Environ. Toxicol. Chem.* **33** 1720–1725. <https://doi.org/10.1002/etc.2604>
-

Evaluating the effectiveness of freshwater fishes as bio-indicators for urban impacts in the Crocodile (West) catchment, South Africa

Jonathan C Levin¹, Darragh J Woodford^{1, 2*} and Gavin C Snow¹

¹School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Private Bag 3, Wits 2050, Johannesburg, South Africa

²South African Institute for Aquatic Biodiversity, Private Bag 1015, Grahamstown 6140, South Africa

Urbanisation in South Africa has resulted in the degradation of aquatic ecosystems across a rural-to-urban gradient; impacting the availability of clean water. Biological organisms, including fish assemblages, have been used as indicators of environmental change, as part of monitoring programmes designed to protect and improve aquatic ecosystem conditions. However, the effectiveness of individual freshwater fish species as bio-indicators for urban impacts has not yet been evaluated. This study investigated the occurrence of freshwater fish species across three urban gradients within the upper Crocodile River sub-management area as potential bio-indicators. Having collected presence and absence data, five native fish species were determined to be widespread. Their effectiveness as bio-indicators for six environmental drivers, identified through principle component analysis, was assessed using species stressor-response curves derived from logistic regression analysis. Of the five species, the largescale yellowfish (*Labeobarbus marequensis*) and stargazer catfish (*Amphilius uranoscopus*) showed potential to be effective bio-indicators for urban impacts on aquatic water quality and instream habitat. These taxa, as effective urban bio-indicators, have the potential to improve the efficiency of urban river health assessments through reducing data gathering and staff training requirements.

Keywords: aquatic ecosystems, integrity, Fish Response Assessment Index (FRAI), species occurrence, stressor-response curves

INTRODUCTION

South Africa's aquatic ecosystems provide numerous essential ecosystem services, including the provision of water to rural communities, and ensuring water and food security, thus supporting socio-economic development (Karr and Chu, 2000; Ollis et al., 2006). Through the years, various anthropogenic factors, such as urbanisation (Wepener et al., 2015), have degraded aquatic ecosystems to a point where their ability to provide crucial ecosystem services has become compromised (Deksissa et al., 2003). Urban impacts on freshwater ecosystems include alteration of instream habitats through canalisation, sedimentation and the loss of riparian vegetation (Paul and Meyer, 2001) and physico-chemical change through polluted runoff and the inflow of sewage from wastewater treatment works (Nyenje et al., 2010). These physical and chemical alterations can have severe impacts on aquatic biodiversity, ecological function, and the usability of the water as a resource to human communities downstream (Paul and Meyer, 2001; Jackson et al., 2016).

To ensure water management areas remain fit to supply water for present and future domestic, agricultural and recreational needs (du Plessis et al., 2014), it is vital to manage and protect aquatic ecosystems surrounding populated areas (Roux, 1999). Monitoring aquatic ecosystems is one of the tools used in their protection, as it allows for environmental degradation to be detected and measured. Ecological monitoring relies on a combination of abiotic and biotic environmental factors to assess the relationship between urbanisation and its impact on aquatic ecosystems (Roux, 1999). Biomonitoring makes use of organisms known as bio-indicators that live within, and respond accordingly to, environments experiencing external and internal stresses

(Li et al., 2010). In South Africa, many bio-indicators, including macrophytes, macroinvertebrates, diatoms, bacteria and fish communities have been used to assess the integrity of aquatic ecosystems (Wepener, 2008; Fourie et al., 2014). Fish, being relatively long-lived and occurring at various trophic levels (Armon and Hänninen, 2015), are frequently used as bio-indicators, as they can respond to a multitude of physical, chemical and biological processes within the aquatic environment, across wide temporal scales (Kleynhans, 1999).

In South Africa, the Fish Response Assessment Index (FRAI) uses the freshwater fish assemblage found in rivers as a measure to assess the ecological state of the country's aquatic ecosystems (Kleynhans, 2007; Avenant, 2010). The index was developed for the then Department of Water Affairs and evolved from a previous index known as the Fish Assemblage Integrity Index (Kleynhans, 1999). The FRAI is based upon expert opinion regarding the environmental preferences and intolerances of reference fish assemblages across South Africa, to certain sets of environmental drivers (Kleynhans, 2007). The index aims to measure how fish assemblages respond, through their occurrence, to instream aquatic habitat modifications; due to instream environmental drivers shaping the river reach (Kleynhans, 2007). These drivers include hydrology and geomorphology as well as the physico-chemical water quality properties (Avenant, 2010).

While the FRAI is designed to be a general index of fish community response to environmental change, it has not specifically been used to measure the response of fish to the impacts of urbanisation. Understanding the specific impacts of urbanisation-derived biophysical stressors on aquatic indicators is required in South Africa, as urbanisation is a leading cause of current land-cover change (Jewitt et al., 2015) and is thus a leading driver of ongoing aquatic ecosystem degradation. Understanding the specific responses of conspicuous, easy-to-identify indicator taxa like fish to urban environmental stressors could greatly improve

*Corresponding author, email: darragh.woodford@wits.ac.za

Received 13 July 2018; accepted in revised form 27 June 2019

our ability to monitor the impact of expanding urban landscapes on aquatic ecosystems. Wenger et al. (2008) demonstrated that urbanisation gradients drive fish species occurrence, owing to the susceptibility of some species to anthropogenic impacts (e.g. physico-chemical). There is, however, no literature currently available indicating the effectiveness of individual freshwater fish species as indicators for the ecological integrity (instream habitat and water quality) of aquatic ecosystems along rural–urban gradients, particularly in South Africa. An appropriately chosen set of individual indicator species could be as useful a monitoring tool as community-based indices of environmental change, provided they show a strong and consistent positive affinity with undisturbed habitats to which they are endemic (Carignan and Villard, 2002).

This study aimed to assess the usefulness of individual freshwater fish species as indicators for aquatic ecosystem modifications associated with urbanisation. The study targeted species with an expected wide natural distribution within a partially urbanised catchment, to ensure a biologically defensible indicator-stressor relationship across the rural-to-urban gradient (Carignan and Villard, 2002).

METHODS

Study area

The study sites for this investigation were selected from rivers within the upper Crocodile (West) River catchment, South

Africa. This catchment is a sub-management area within the larger Crocodile (West) Marico Water Management Area (Fig. 1). The southern section of the sub-management area is characteristically urban, with large residential and industrial land uses across the City of Johannesburg Municipality (DWAf, 2008). Moving towards Midrand and the City of Tshwane (Pretoria) the two latter land uses continue, while urban sprawls become increasingly common (DWAf, 2008). The remaining land uses within the sub-management area are comprised of peri-urban small-holdings, agricultural and mining activities.

River study sites were selected based upon a gradient of urbanisation within the upper Crocodile River sub-management area. The urban gradient, comprised of urban, peri-urban and rural categories, was classified based upon the percentage of urban land cover in the rivers' upstream reaches. Seven rivers were selected for the research, with a total of eight river study sites (Fig. 1).

Characterising the urban gradient

A 2014 national land-cover map (Geoterrimage, 2014) was analysed using ArcGIS 10.3 (ESRI, 2015) to identify land cover upstream of each river site. The surrounding land cover was classified into six classes: water, vegetation, bare non-vegetated, agriculture, urban, and mining. The urban category was created by combining all urban land uses (urban schools, urban residential, etc.) into one 'urban' land cover class. This class was then used to place each river site along the urban gradient, as

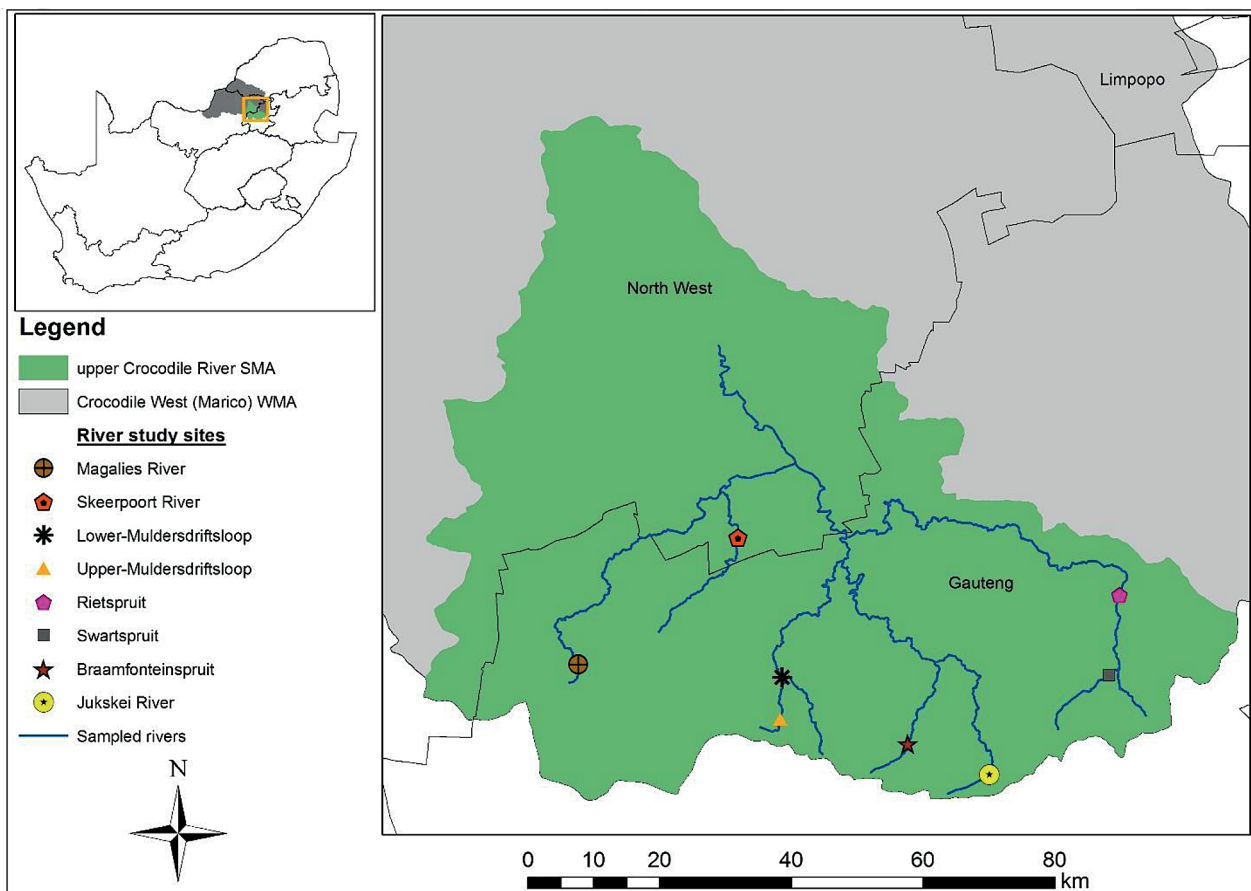


Figure 1. Locations of the eight study sites within the upper Crocodile sub-management area (SMA), which falls within the greater Crocodile West (Marico) Water Management Area (WMA), South Africa.

either rural, peri-urban or urban (Ding et al., 2015). To obtain the upstream land cover classes surrounding each river, a 1 km buffer was placed around them. To assign each river study site an urban category, a standardised rule was formulated, following McEwan and Joy (2009). The rule classified sites with less than 10%, between 10% and 60%, and greater than 60% upstream urban land covers to be rural, peri-urban and urban, respectively. The eight study sites were comprised of two rural reference sites, four peri-urban sites and two urban sites.

Physico-chemical parameters

Temperature (°C), dissolved oxygen (mg/L and %), conductivity (µS/cm), and pH (pH units) were recorded using a YSI Professional Plus Multi-parameter meter. To obtain average results for each study site, physico-chemical results were automatically recorded every 30 s for 5 min, 3 times within 30 min, each in a different aquatic meso-habitat (where available). Velocity (m/s) and discharge (m³/s) were calculated once at each study site using an auto-calibrated SonTek/YSI FlowTracker discharge probe (Xylem, San Diego). Turbidity was measured using a HI98703-01 Hanna Turbidity Meter, in nephelometric turbidity units (NTU) (DWAF, 1996). To analyse nutrient concentrations within the aquatic ecosystem, water was collected on site and filtered through 0.45 µm syringe filters. The filtrate was then placed into sealable sterile plastic bags, then a portable ice chest, and later moved into a freezer. The samples were defrosted and immediately analysed spectrophotometrically according to the methods outlined in Bate and Heelas (1975) for nitrate-N + nitrite-N, and Parsons et al. (1984) for ammonium-N and orthophosphate-P. Nutrient concentrations were expressed in mg/L.

Assessing prevalence of velocity/depth classes

To measure the availability of velocity/depth classes, which are the principal habitat availability signifiers on which fish habitat preferences in the FRAI are based (Kleynhans, 2007), the Rapid Habitat Assessment Method (RHAM) was performed at each study site (following Kleynhans and Louw, 2009). Cross-sectional transects were placed across all available instream meso-habitats, where available, within a 100 m river reach following Peck et al. (2006). The sampled meso-habitats included riffles, runs, pools and pool-runs (Kleynhans and Louw, 2009). The field data were then analysed using the RHAM algorithmic macro in Excel (Kleynhans and Louw, 2009), to indicate the dominant velocity-depth class/classes within each of the meso-habitats sampled, per site.

Sampling freshwater fishes

Each river site was sampled using the Fish Response Assessment Index (FRAI) sampling protocol as outlined in Kleynhans (2007). Each site was sampled once, during low flow conditions (March-April 2017). Moving upstream within the study site, a SAMUS 720MS backpack electro-fisher was used to sample each available meso-habitat for 10 min (Kleynhans, 2007). To prevent avoidance of the electrical current, a 1 x 5 m seine net was used as a downstream block net on each meso-habitat. Each fish present was then processed according to the individual habitat in which it was found, and released into the stream.

Table 1. Percentage prevalence of key velocity-depth classes across surveyed meso-habitats at sampling sites derived using the Rapid Habitat Assessment Method

Site	Velocity-Depth Class			
	Slow Shallow (%)	Slow Deep (%)	Fast Shallow (%)	Fast Deep (%)
Braamfonteinspruit	100	25	25	25
Jukskei	100	0	75	75
Upper-Muldersdriftsloop	100	25	50	50
Lower-Muldersdriftsloop	100	75	0	50
Magalies	100	0	50	25
Skeerpoort	100	0	33	0
Swartspruit	100	0	50	100
Rietspruit	100	0	66	100

The prevalence of velocity-depth classes across meso-habitats determined from RHAM (Table 1) and class preferences of each sampled species (following predictions of FRAI Version 1) enabled an expected species prevalence (across sampled meso-habitats) to be assigned to each site. For example, a species with a known preference for slow-shallow habitat would be given an expected frequency-of-occurrence (FROC) score of 5 if slow-shallow prevalence was 100%, or a FROC score of 1 if the same class prevalence was 25%. An expected species pool for each site was generated using a pristine headwater fish community dataset from the adjacent Groot Marico catchment (Kimberg et al., 2014) within the Crocodile (West) and Marico Water Management Area (WMA3), and checked for missing expected species using the FROC dataset for each site's respective sub-quaternary catchment (Kleynhans et al., 2007). Expected prevalence scores per species per site were assigned based on availability of velocity-depth classes, and observed fish species prevalence data were then entered into the FRAI Version 1 spreadsheet, which used metric groups of individual fish species preferences and tolerances to numerically assess divergence from the expected fish assemblage composition at a given site (Kleynhans, 2007). Automated preferences and weightings were used for all algorithmic settings. The output of the FRAI model was a Present Ecological Status (PES) percentage for each fish assemblage. Using the PES, an ecological category was derived. Ecological categories (ranging from A – pristine to F – critically modified) are indicative of the aquatic ecosystem's ecological integrity (Kleynhans, 2007). To determine whether there was any significant difference in the integrity of the rivers across the three land cover categories, a non-parametric Kruskal-Wallis test was performed on the PES percentages within each category.

Evaluating individual fish species as suitable indicator species

To assess individual fish species as possible indicators, we determined their response to environmental stressors along the urban gradient. To identify the key environmental stressors which best characterise the urban gradient, principle component analysis (PCA) was performed in R (R Core Team, 2017). The PCA identifies the variables that explain the largest portion of the dataset's variance (Shetty et al., 2015). To accurately determine which variables had a significant association with each principle component (PC), the Dimdesc

function in the FactoMineR package was used (Le et al., 2008).

Having identified the key stressors describing the urban gradient and its impacts, the distributional responses of fish species to these stressors were assessed. Logistic regression analysis was performed on the naturally widespread fish species, relative to each key stressor. The regression output was individual species stressor-response curves (sensu Meador et al., 2005), indicating the species' threshold of survivability relative to each key environmental driver (Market et al., 2003).

Effective bio-indicator species display a negative probability of occurrence relationship with metrics of increasing environmental degradation, which can be modelled using logistic regression (Oliveira and Cortes, 2006). Consequently, we defined an ideal indicator species to be one that displays a negative binomial pattern of occurrence across the rural-urban gradient, representing a species stressor-response curve. To indicate how well the species tracks the stressor, an R^2 value derived from the fit of the probabilistic curve relative to the fish species occurrence, was used as a goodness-of-fit indicator. An R^2 value of 1 demonstrated a perfect fit to a binomial stressor-response curve, thus indicating a good bio-indicator response of the fish species to that particular stressor. These data were analysed using R (R Core Team, 2017) and later plotted in SigmaPlot (Systat Software, 2017).

RESULTS

Land cover analysis

Following the land cover analysis upstream of each study site, the sites on the Magalies and Skeerpoort rivers were classified as rural, owing to having urban land covers of 0.05% and 0.62%, respectively. The study sites on the Upper-Muldersdriftsloop and the Swartspruit were determined to have urban land covers of 58.38% and 52.46%, respectively. The study sites on the Lower-Muldersdriftsloop and Rietspruit had urban land covers of 55.48% and 24.58%, respectively. All

four sites are thus characterised as peri-urban. The study sites on the Braamfonteinspruit and Jukskei rivers were determined to be urban, owing to having urban land covers of 75.49% and 73.94%, respectively.

Physico-chemical and nutrient stressors

Field sampling took place between 20 April and 8 May 2017. Among the physico-chemical parameters measured, turbidity increased on average across the gradient (Table 2), with the urban gradient having the highest (16.96 NTU), similarly so with electrical conductivity (519.11 μ S/cm). This is opposed to dissolved oxygen (7.42 mg/L to 5.49 mg/L) and pH (7.39 to 6.45), which on average both decreased as gradients became increasingly urban. Among the river sites, the urban Jukskei River study site had the highest electrical conductivity (624.73 μ S/cm), lowest dissolved oxygen (5.49 mg/L and 47.17%) and pH (6.39).

The concentrations of the three macro-nutrients investigated, on average, progressively increased as river sites became progressively more urban. Regarding nitrate + nitrite, the peri-urban sites (3.14 mg/L), particularly the Rietspruit (6.03 mg/L) study site, had the highest average concentration. The most noticeable macro-nutrient increase was pertaining to ammonium, which on average increased from 0.02 mg/L to 5.25 mg/L along the urban gradient (Table 2), with the Jukskei River study site having the highest concentration (9.01 mg/L). Orthophosphate, in contrast, increased the least on average along the gradient.

Fish Response Assessment Index (FRAI)

A total of 9 fish species were sampled, which matched the expected species pool. Three of the nine species, including the shortspine suckermouth (*Chiloglanis pretoriae* van der Horst, 1931), the straightfin barb (*Enteromius paludinosus* Peters, 1852) and the banded tilapia (*Tilapia sparrmanii* A. Smith, 1840) were not present in upland rivers (Table 3). Additionally,

Table 2. The average (\pm SD) physico-chemical variables measured for each urban gradient class.

Gradient Class	Physico-chemical drivers								
	Conductivity (μ S/cm)	Dissolved oxygen (mg/L)	Dissolved oxygen (%)	pH	Temperature ($^{\circ}$ C)	Turbidity (NTU)	Nitrate + nitrite (mg N/L)	Ammonium (mg N/L)	Ortho-phosphate (mg P/L)
Rural	296.22 \pm 35.72	7.42 \pm 0.37	75.76 \pm 4.94	7.39 \pm 0.33	16.60 \pm 0.29	2.53 \pm 1.14	0.17 \pm 0.14	0.02 \pm 0.03	0.01 \pm 0.00
Peri-urban	341.22 \pm 172.67	8.14 \pm 0.28	79.85 \pm 3.47	7.37 \pm 0.33	14.44 \pm 1.07	10.54 \pm 5.40	3.14 \pm 2.53	0.22 \pm 0.80	0.13 \pm 0.11
Urban	519.11 \pm 149.38	6.58 \pm 1.53	55.48 \pm 11.75	6.45 \pm 0.08	16.02 \pm 2.69	16.96 \pm 2.43	1.71 \pm 1.10	5.25 \pm 7.66	0.58 \pm 0.48

Table 3: The presence (1) and absence (0) of the nine native freshwater fish species sampled across the river profile and urban gradient where the species do and do not (NA) occur naturally.

Species Code	Scientific name	Common name	Fish presence (1) and absence (0)					
			Upland rivers			Lowland rivers		
			Rural	Peri-urban	Urban	Rural	Peri-urban	Urban
LMAR	<i>Labeobarbus marequensis</i>	Largescale yellowfish	1	0	0	1	0	0
PPHI	<i>Pseudocrenilabrus philander</i>	Southern mouthbrooder	1	1	0	1	1	0
EMOT	<i>Enteromius motebensis</i>	Marico barb	1	1	0	1	1	0
AURA	<i>Amphilius uranoscopus</i>	Stargazer catfish	1	0	0	1	0	0
LPOL	<i>Labeobarbus polylepis</i>	Smallscale yellowfish	1	0	0	1	1	0
CGAR	<i>Clarias gariepinus</i>	Sharptooth catfish	NA	NA (1)	NA (1)	1	0	0
CPRE	<i>Chiloglanis pretoriae</i>	Rock catlet	NA	NA	NA	1	0	0
EPAU	<i>Enteromius paludinosus</i>	Straightfin barb	NA	NA	NA	1	0	0
TSPA	<i>Tilapia sparrmanii</i>	Banded tilapia	NA	NA	NA	1	0	0

sharp-toothed catfish (*Clarias gariepinus* Burchell, 1822) was only found in upland streams when directly downstream of an impoundment. As a result, five of the nine fish species sampled were determined to be widespread (occurring in upland and lowland streams). The designation of species as upland, lowland or widespread was verified using community data from Kimberg et al. (2014), which associated fishes from within WMA3 with natural habitat features along the elevation gradient.

Of the widespread species, the largescale yellowfish (*Labeobarbus marequensis* A. Smith, 1841) and stargazer catfish (*Amphilius uranoscopus* Pfeffer, 1889) both only occurred in rural upland and lowland rivers, while the southern mouthbrooder (*Pseudocrenilabrus philander* Weber, 1897) and Marico barb (*Enteromius motebensis* Steindachner, 1894) occurred in the rural and peri-urban sites of upland and lowland rivers. The smallscale yellowfish (*Labeobarbus polylepis* Boulenger, 1907) was found in peri-urban and rural lowland rivers, as well as rural upland rivers (Table 3).

A PES, indicative of the current condition of the river, was derived for each of the eight study sites using FRAI (Table 4). The rural Magalies and Skeerpoort study sites were determined to have the highest average PES of 85.41%, deviating 14.60% from reference conditions, with an Ecological Category of A/B. As the percentage land cover became progressively more urban, the PES of the study sites declined. The peri-urban study sites had an average PES of 36.9%, deviating 63.1% from reference conditions, with an Ecological Category of E. Additionally, urban sites had an average PES of 22.91%, deviating 77.09% from reference conditions, and an Ecological Category of E. The average PESs for the three urban categories did not differ significantly ($\chi^2 = 5.13$, $df = 2$, $p = 0.08$), although this may have been an artefact of insufficient site replication across the three urban gradient classes.

Table 4. The Ecological Categories as determined by the Fish Response Assessment Index (FRAI) for each of the river study sites.

River	Urban land cover (%)	Urban gradient	FRAI class (Ecological Category)	PES
Magalies River	0.05	Rural	A/B	88.19
Skeerpoort River	0.62	Rural	B	82.62
Lower-Muldersdriftsloop	55.48	Peri-urban	D	48.24
Swartspruit	52.46	Peri-urban	D	44.72
Upper-Muldersdriftsloop	58.38	Peri-urban	E	30.66
Rietspruit	24.58	Peri-urban	E	23.98
Braamfonteinspruit	75.49	Urban	E	25.81
Jukskei River	73.94	Urban	F	20

Table 6. Three of the five native widespread fish species are potential indicators, owing to demonstrating perfect binomial response curves ($R^2 = 1$) for most of the environmental drivers, particularly those drivers in Principle Component 1, which best describes the urban gradient and its impacts. There were six degrees of freedom throughout. Standard Errors (SE) of each binomial response curve are provided.

Species	Urban land cover			Ammonium			Turbidity		
	Coefficient	SE	R^2 (McFadden)	Coefficient	SE	R^2 (McFadden)	Coefficient	SE	R^2 (McFadden)
LMAR	-1.961	4 216.51	1	-13.04	2 835.02	1	-23.77	24 778.17	1
PPHI	-14.27	9 207.25	1	-1.39	1 507.14	1	-0.11	0.11	0.13
EMOT	-0.04	0.03	0.15	-0.04	0.03	0.42	-0.097	0.1	0.1
AURA	-0.96	4 216.51	1	-13.04	2 835.02	1	-23.77	24 778.18	1
BPOL	-0.06	0.04	0.35	-0.34	0.37	0.46	-21.06	25 831.06	1

Stressor principle component analysis

The cumulative percentage of PCs 1 to 3 accounted for 89% of recorded variance and were thus retained for further analyses (following Mishra, 2010). Urban land cover, turbidity and ammonium were significantly correlated with PC 1 (Fig. 2, Table 5). Temperature and dissolved oxygen were correlated with PC 2 (Fig. 2, Table 5), and had an auto-correlated relationship, ($r = -0.68$, $df = 6$, $p = 0.07$). Conductivity was the only variable significantly correlated with PC 3 (Table 5). Subsequently, six key environmental drivers were determined for use in the logistic regression analysis as they contribute the greatest ($\Sigma = 89\%$) to the dataset's variance.

The study sites were distinguishably clustered into their respective rural-urban classifications along PC 1, which accounts for most of the dataset's variance (41.2%). The three environmental drivers which were significantly correlated with PC 1 were thus considered the key stressors best describing the urban gradient and its impacts (Table 5).

Logistic regression analysis

Logistic regression analysis was performed on each of the five widespread native fish species relative to the six key environmental drivers determined during the PCA. Through the plotting of species response curves, using a probabilistic model, it was determined that each of the five species had at least one perfect species stressor-response curve, as represented by $R^2 = 1$. Of the five species, three demonstrated perfect species stressor-response curves for more than two of the environmental drivers (Table 6) and thus, are the potential indicator candidates. These three species only displayed perfect binomial responses to the environmental drivers in PC 1 (Table 6).

Among the three potential indicator species, the southern

Table 5. The six key environmental drivers that were significantly correlated to the first three principle components

Variable	Correlation coefficients (component loadings)	p-value
Principle Component 1		
Percent urban	0.92	<0.01
Ammonium	0.89	<0.01
Turbidity	0.86	0.01
Principle Component 2		
Dissolved oxygen (mg/L)	0.94	<0.01
Temperature	-0.86	0.01
Principal Component 3		
Conductivity	0.8	0.02

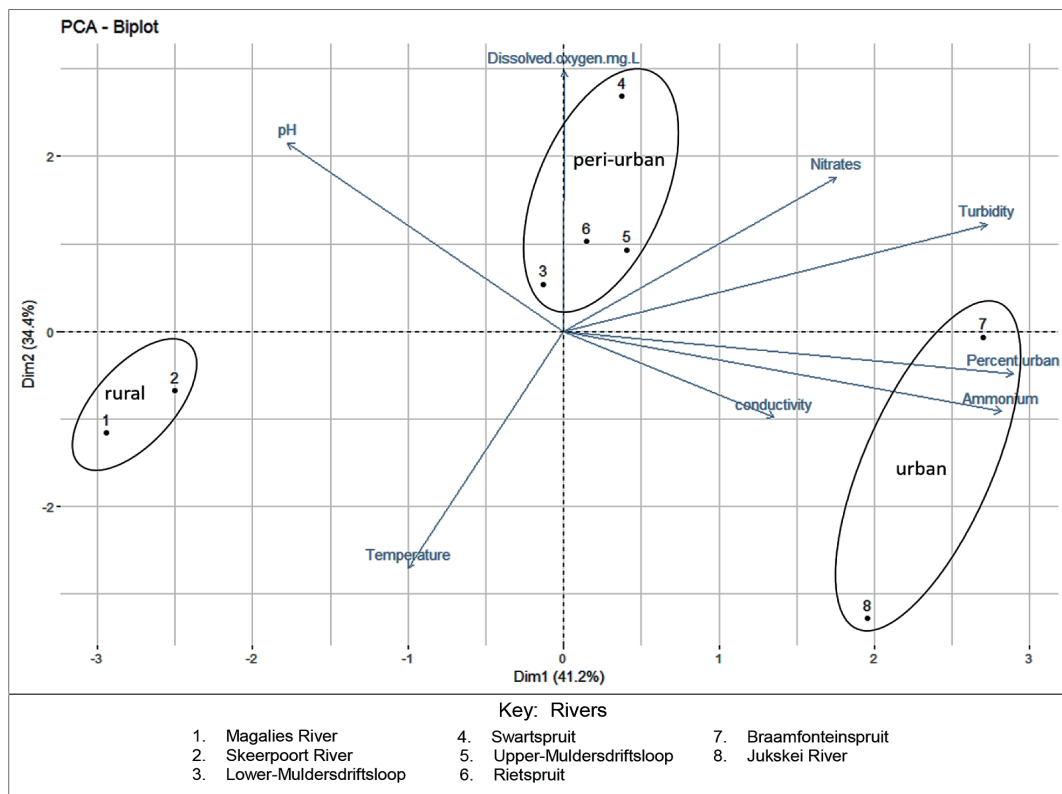


Figure 2: A Principle Component Analysis biplot showing the distinguishable grouping of the study sites into their respective urban category (gradient) relative to the strength of the different environmental drivers within each of the two principle components

mouthbrooder demonstrated perfect species response curves relative to stressors 'urban land cover' and 'ammonium'. The species' apparent survivability threshold was reached when urban land cover exceeded 58% and ammonium exceeded 0.45 mg/L (Fig. 3). The largescale yellowfish and the stargazer catfish both displayed perfect, identical species stressor-response curves relative to urban land cover, ammonium and turbidity. The apparent survivability thresholds for both species were 13% for urban land cover, 0.08 mg/L for ammonium and 4.00 NTU's for turbidity (Fig. 3). Although the southern mouthbrooder responded well to two of the six drivers, the largescale yellowfish and the stargazer catfish responded to all three drivers best describing the urban gradient (PC1), and thus can be considered better potential indicator species for urban impacts.

DISCUSSION

Investigating the response of freshwater fishes at species level has the potential to provide significant information on species survivability thresholds, relative to particular anthropogenic factors (e.g. pollution and habitat modification) in urban river catchments (Wenger et al., 2008). In this study, we detected strong presence/absence responses in four native fish species to key environmental stressors directly associated with the rural-to-urban gradient. This indicates a potential for these species to perform as indicators of urban impacts on stream ecosystem integrity.

Characteristics of the rural-urban gradient

As urban land cover increased, there were noticeable changes in the average physico-chemical water quality parameters within the rivers. This can be attributed to the greater number

of domestic and industrial establishments along the gradient, which have the potential to discharge pollution either directly into the rivers (Ansara-Ross et al., 2008) or through stormwater drains, which eventually exit into rivers (Koehn et al., 2011).

On average, turbidity increased from 2.53 NTU in the rural study sites, to 10.54 NTU and 16.96 NTU in peri-urban and urban sites, respectively. Such elevated turbidity levels in both peri-urban and urban rivers may be attributed to the greater number of impervious surfaces and associated stormwater drains as regions become increasingly urban (Mwangi, 2014). Increases in turbidity may function as an indicator for catchment hardening due to drains and impermeable surfaces becoming more predominant across the urban gradient (Koehn et al., 2011). Turbidity was one of the stressors best describing the urban gradient in this study and its impacts are a recognised environmental driver influencing freshwater fish survival (Kjelland et al., 2015). Higher suspended solids are also associated with increased benthic siltation, potentially limiting the availability of instream habitats for habitat specialists (Gorman and Karr, 1978).

Nutrients, including nitrogen and phosphorus, were other major stressors associated with increased urbanisation. Stormwater drains and impermeable surfaces have a significant role in the nutrient loading of peri-urban and urban rivers (Ward and Winter, 2016). Drains and their associated urban surfaces facilitate the transportation of urban non-point-source nutrient run-off, including detergents, fertilisers and animal waste, into aquatic ecosystems (Carey et al., 2013). Both phosphorus and nitrogen are crucial water quality parameters as elevated concentrations of either may cause excessive toxic and non-toxic algal growth. Such growth and the subsequent bacterial decay of accumulated detritus may result in hypoxic

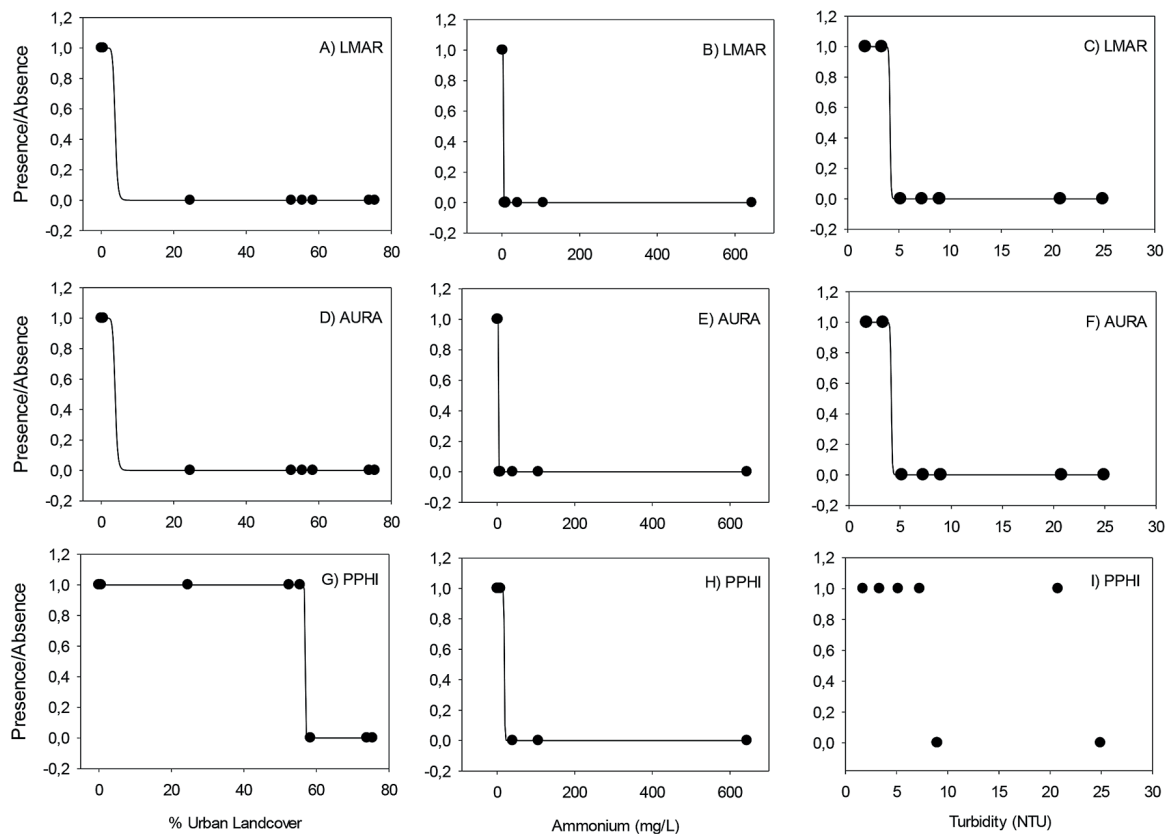


Figure 3. The perfect binomial response curves for the three potential indicator species, being A-C) the largescale yellowfish (LMAR), D-F) stargazer catfish (AURA) and G-I) southern mouthboorder (PPHI) relative to the three environmental drivers (urban land cover, ammonium and turbidity) found to be significant correlators with principle component one. Plotted trend lines indicate logistic regression where $R^2 = 1$.

or anoxic water conditions (Carey et al., 2013), depleting fish communities (Smith et al., 1999). Along the gradient, both the peri-urban (0.13 mg/L) and urban (0.58 mg/L) sites exceeded the orthophosphate limit (<0.025 mg/L) used to prevent harmful eutrophic conditions (Dallas and Day, 2004). A similar trend was evident for average nitrogen concentrations, where the peri-urban sites (3.14 mg/L) exceeded the nitrogen threshold (<2.5 mg/L) set to limit harmful eutrophic conditions (Dallas and Day, 2004). Among the sites, the peri-urban Rietspruit study site had the highest nitrate concentration (6.03 mg/L); exceeding the national prescribed nutrient threshold (DWAF, 1996). The most likely source of elevated nitrogen in this tributary is the East Rand Water Care Association (ERWAT), a wastewater treatment works (WWTW) located in Kempton Park, Gauteng, which continuously discharges treated effluent into the river (N Thlaku, ERWAT, pers. comm., 2017). Such facilities are a main point source for nitrogen in many rivers across South Africa (Dabrowski and De Klerk, 2013).

Ammonium was another key stressor associated with urbanisation. In urban areas, blocked and ruptured sewage mains are a common occurrence, resulting in untreated sewage entering nearby rivers (Matowanyika, 2010). Associated with instream sewage contamination are elevated ammonium concentrations; thus ammonium is an effective indicator for such contamination (Passell et al., 2007; Constable et al., 2015). Such sewage input may explain the increase in ammonium concentrations across the urban gradient, resulting in average peri-urban (0.22 mg/L) and urban concentrations (5.25 mg/L) exceeding the acute toxicity threshold (<0.1 mg/L) (DWAF,

1996). Among the sites, the Jukskei River had the highest ammonium concentration (9.01 mg/L) and was the only site to exceed the South African Target Water Quality Range for toxic ammonia (> 0.007 mg/L); set to prevent physiological harm to fish species (DWAF, 1996; Passell et al., 2007). It is likely that inflow of untreated sewage into the Jukskei River during the sampling period may account for the site's hypoxic conditions, which is a known by-product of high quantities of organic waste (Dallas and Day, 2004).

Urban land cover percentage, another strong correlate with the urban gradient in the PCA, is representative of urban expansion. While it is a key driver impacting nearby water quality, urban expansion also impacts the integrity of riparian vegetation. As urbanisation progresses, riparian zones become increasingly degraded (Meador and Goldstein, 2003) and urban land cover can thus be an indicator for riparian integrity. Such degradation has profound impacts on aquatic ecosystem integrity through altering water temperature, and instream food availability through decreasing the input of allochthonous material and increasing river sedimentation; all of which can impact fish assemblage diversity (Ryan, 1991; Meador and Goldstein, 2003; Reid et al., 2008).

Fish community response to urbanisation

To assess whether fish species responded to the distinguishable physico-chemical water quality conditions along the urban gradient, fish assemblage structure was first assessed using a standardised fish sampling procedure, the Fish Response

Assessment Index (FRAI). Having performed FRAI, the Ecological Category, and associated description of current ecological condition of the water resource, was derived for each study site (Kleynhans and Louw, 2007).

From the Ecological Categories, it is evident that as the land cover surrounding the rivers became increasingly urban, the ecological integrity of the rivers subsequently decreased as represented by declining fish diversity (Meador et al., 2005). This is observed in the change from near-natural conditions (Category: A/B) in rural aquatic ecosystems such as the Magalies and Skeerpoort rivers, to seriously modified conditions (Category E) in the urban Braamfonteinspruit and Jukskei rivers. This negative fish diversity response is supported by Roux and Selepe (2015), who found that fish assemblages in South Africa respond negatively to urbanisation, resulting in a river with a lower Ecological Category (Kleynhans, 2007). This is owing to the fish community responding to poorer water and habitat quality conditions derived from surrounding anthropogenic activities (Avenant, 2010; Shetty et al., 2015). Although there was no significant difference in the average PES among urbanisation classes, FRAI was nonetheless effective in detecting a decline in ecological integrity across the rural-to-urban gradient.

Potential bio-indicator species

Of the nine fish species recorded, only five were determined to be naturally widespread within the study area. The sharptooth catfish, which is naturally associated with lowland rivers in the western Limpopo basin (Kimberg et al., 2015), was also present in two upland study sites. Both sites, however, are located downstream of urban impoundments, which are likely demographic sources for this species in these small streams (Weyl et al., 2016). Due to the possibility that this species was historically stocked for angling in these waterbodies (Weyl et al., 2016), it was excluded from the list of naturally widespread species for the purpose of this study.

Logistic regression analysis performed on each of the five widespread species, relative to the six environmental drivers, produced probabilistic curves representative of species stressor-response curves. These curves were used to analyse the effectiveness of individual fish species as bio-indicators for environmental drivers along the urban gradient. All five of the widespread fish species responded with a minimum of one perfect binomial response curve ($R^2 = 1$). Three of the five species, the southern mouthbrooder, largescale yellowfish and stargazer catfish, were the best indicator candidates as they responded with perfect binomial response curves to most of the environmental drivers, including urban land cover, ammonium and turbidity.

Of the three potential indicator species, the southern mouthbrooder appeared the most robust to the impacts of urbanisation. It displayed higher occurrence thresholds for urban land cover (58%) and ammonium (0.45 mgN/L), compared to those of the largescale yellowfish and stargazer catfish (13% and 0.08 mgN/L). Furthermore, while southern mouthbrooder only responded to ammonium (an indicator for water pollution) and urban land cover (indicating riparian degradation), the largescale yellowfish and stargazer catfish also responded to turbidity (indicating suspended sediment). Owing to this apparent sensitivity to all three urban stressors, they are the two best potential indicator species for urban impacts.

These observed binomial responses can be attributed to the species' environmental preferences. For instance, the largescale

yellowfish predominately feeds on plant and algal matter, and insects (Skelton, 2001; Fouché, 2009). This may account for the species' response to urban land cover, which is an indicator of riparian zone degradation that limits the allochthonous input of plant matter and insects needed to maintain fish assemblage diversity (Harris, 1995). The largescale yellowfish also responded to turbidity; however, owing to it being a habitat generalist (Fouché, 2009), sedimentation does not necessarily impact its habitats (Berkman and Rabeni, 1987). Rather, its response may be attributed to sediment-laden water inhibiting respiration through clogging gills and/or reducing hunting effectiveness through visual impairment (Ryan, 1991; Fouche, 2009). The susceptibility of largescale yellowfish to pollution can also be inferred from its perfect binomial response to ammonium (at low concentrations).

The binomial response of stargazer catfish to urban land cover, an indicator of riparian degradation, strongly suggests that the species prefers habitat where riparian vegetation is present. This species is a habitat specialist, occupying rocky biotopes in fast-flowing waters and is thought to lay eggs in the interstitial spaces (Skelton, 2001). However, as the instream sediment load increases, these interstitial spaces collect sediment (Ryan, 1991), subsequently removing potential spawning habitat. Increased siltation may also impact this species' feeding ability, as it feeds on small organisms off of rocky surfaces (Skelton, 2001), which are susceptible to being disturbed by increased sediment (Berkman and Rabeni, 1987).

CONCLUSION

Fish have become well established in South Africa as an assemblage-level indicator of ecosystem health, and in this study confirmed that rivers decline in their eco-status classification across a rural-to-urban gradient. In addition, two widespread fish species within the upper Crocodile River sub-management area were identified as potential species-level bio-indicators of stressors associated with urbanisation.

The largescale yellowfish and the stargazer catfish have the potential to indicate modifications in both water quality and instream habitat. This is due to both species having responded with negative logistic stressor-response curves to variables indicating pollution (ammonium), catchment hardening (turbidity) as well as riparian degradation (urban land cover). To confirm the apparent effectiveness of these species and others as bio-indicators for urban impacts, it is recommended that further sampling is conducted to improve replication of this dataset and to further develop research on other fish species-stressor relations.

Individual fish species as bio-indicators for water quality and habitat modifications provide an opportunity for tiered river health assessments optimised for time and practitioner capability. For instance, assessing the presence of indicator species in their known habitats could act as a first-tier assessment for practitioners who do not have the time, resources or training to perform established procedures such as RHAM or FRAI. If an indicator species disappears from a site, it could trigger the requirement to have trained practitioners perform the latter assessments to establish whether the ecological status of the river has degraded due to increased urbanisation. Thus for long-term monitoring by agencies with human capacity limitations, less time could be spent on training individuals to identify only key indicator species for such first-tier monitoring. We recommend that more research be directed towards identifying such species and verifying their

susceptibility to key anthropogenic impacts like urbanisation, as well as others such as mining. In the case of other forms of human impact, we recommend that the key physico-chemical and habitat stressors associated with that impact be identified and then directly compared to the distributions of potential indicator species.

ACKNOWLEDGEMENTS

This research was funded by National Research Foundation incentive funding (Grant 103581) and a Wits University Science Faculty research grant to DJW. The authors thank the Gauteng Provincial Government (Permit 0191) and North West Department of Rural, Environmental and Agricultural Development (Permit HQ-04/09/17-352) for fieldwork authorisation, and Rietvlei Nature Reserve, Walter Sisulu Botanical Gardens, Riverstone Lodge, Highlander Trout Lodge, Serengeti Golf Estate and Happy Acres Environmental Education Centre for access to river sites. Sampling was authorised by the University of the Witwatersrand Animal Research Ethics Committee (clearance certificate 2017/03/23/C). The authors thank Refilwe Chilo and Prenisha Govender for their assistance in the field, Gina Walsh for FRAI and RHAM application support, and Jason Marshall for statistical analysis support.

REFERENCES

- ANSARA-ROSS TM, WEPENER V, VAN DEN BRINK PJ and ROSS MJ (2008) Probabilistic risk assessment of the environmental impacts of pesticides in the Crocodile (west) Marico catchment, North-West Province. *Water SA* **34** (5) 637–646.
- ARMON RH and HÄNNINEN O (2015) *Environmental Indicators*. Springer, New York. 643 pp.
- AVENANT MF (2010) Challenges in using fish communities for assessing the ecological integrity of non-perennial rivers. *Water SA* **36** (4) 397–405. <https://doi.org/10.4314/wsa.v36i4.58407>
- BATE GC and HEELAS BV (1975) Studies on the nitrate nutrition of two indigenous Rhodesian grasses. *J. Appl. Ecol.* **12** 941–952. <https://doi.org/10.2307/2402100>
- BERKMAN HE and RABENI CF (1987) Effect of siltation on stream fish communities. *Environ. Biol. Fishes* **18** (4) 285–294. <https://doi.org/10.1007/BF00004881>
- CAREY RO, HOCHMUTH GJ, MARTINEZ CJ, BOYER TH, DUKES MD, TOOR GS and CISAR JL (2013) Evaluating nutrient impacts in urban watersheds: Challenges and research opportunities. *Environ. Pollut.* **173** 138–149. <https://doi.org/10.1016/j.envpol.2012.10.004>
- CARIGNAN V and VILLARD MA (2002) Selecting indicator species to monitor ecological integrity: a review. *Environ. Monit. Assess.* **78** (1) 45–61. <https://doi.org/10.1023/A:1016136723584>
- DABROWSKI JM and DE KLERK LP (2013) An assessment of the impact of different land use activities on water quality in the upper Olifants River catchment. *Water SA* **39** (2) 231–244. <https://doi.org/10.4314/wsa.v39i2.6>
- DALLAS HF and DAY JA (2004) The effect of water quality variables on aquatic ecosystems: a review. WRC Report No. TT 224/04. Water Research Commission, Pretoria.
- DEKSISSA T, ASHTON PJ and VANROLLEGHEM PA (2003) Control options for river water quality improvement: A case study of TDS and inorganic nitrogen in the Crocodile River (South Africa). *Water SA* **29** (2) 209–218. <https://doi.org/10.4314/wsa.v29i2.4858>
- DING J, JIANG Y, FU L, LIU Q, PENG Q and KANG M (2015) Impacts of land use on surface water quality in a Subtropical River Basin: a case study of the Dongjiang River Basin, Southeastern China. *Water* **7** (8) 4427–4445. <https://doi.org/10.3390/w7084427>
- DU PLESSIS A, HARMSE T and AHMED F (2014) Quantifying and predicting the water quality associated with land cover change: a case study of the Blesbok Spruit Catchment, South Africa. *Water* **6** (10) 2946–2968. <https://doi.org/10.3390/w6102946>
- DWAF (Department of Water Affairs and Forestry, South Africa) (1996) South African Water Quality Guidelines Volume 7 Aquatic Ecosystems: Second Edition. DWAF, Pretoria.
- DWAF (Department of Water Affairs and Forestry, South Africa) (2008) The development of a reconciliation strategy for the Crocodile (West) water supply system. Department of Water Affairs and Forestry, DWAF Report No. P WMA 03/000/00/3408. DWAF, Pretoria.
- ESRI (2015) ArcGIS Desktop: Release 10.3. Redlands, CA: Environmental Systems Research Institute.
- FOURIE HE, THIRION C and WELDON CW (2014) Do SASS5 scores vary with season in the South African highveld? A case study on the Skeerpoort River, North West province, South Africa. *Afr. J. Aquat. Sci.* **39** (4) 369–376. <https://doi.org/10.2989/16085914.2014.978256>
- FOUCHÉ PSO (2009) Aspects of the ecology and biology of the Lowveld Largescale Yellowfish (*Labeobarbus marequensis*, Smith, 1843) in the Luvuvhu River, Limpopo River System, South Africa. Doctorate of Philosophy in Zoology, University of Limpopo, South Africa.
- GORMAN OT and KARR JR (1978) Habitat structure and stream fish communities. *Ecology* **59** (3) 507–515. <https://doi.org/10.2307/1936581>
- GEOTERRAIMAGE (2015) 2013 – 2014 South Africa National Land-Cover Dataset, Geoterraimage. URL: http://bgis.sanbi.org/DEA_Landcover/project.asp (Accessed 2 March 2017).
- HARRIS JH (1995) The use of fish in ecological assessments. *Austral Ecol.* **20** (1) 65–80. <https://doi.org/10.1111/j.1442-9993.1995.tb00523.x>
- JACKSON MC, WOODFORD DJ and WEYL OLF (2016) Linking key environmental stressors with the delivery of provisioning ecosystem services in the freshwaters of southern Africa. *Geogr. Environ.* **3** (2) 1–12. <https://doi.org/10.1002/geo2.26>
- JEWITT D, GOODMAN PS, ERASMUS BFN, O'CONNOR TG and WITKOWSKI ETF (2015) Systematic land-cover change in KwaZulu-Natal, South Africa: Implications for biodiversity. *S. Afr. J. Sci.* **111** (9) 1–9. <https://doi.org/10.17159/sajs.2015/20150019>
- KARR JR (1981) Assessment of biotic integrity using fish communities. *Fisheries* **6** (6) 21–27. [https://doi.org/10.1577/1548-8446\(1981\)006<0021:AOBIUF>2.0.CO;2](https://doi.org/10.1577/1548-8446(1981)006<0021:AOBIUF>2.0.CO;2)
- KARR JR and CHU EW (2000) Sustaining living rivers. *Hydrobiologia* **422** 1–14. <https://doi.org/10.1023/A:1017097611303>
- KIMBERG PK, WOODFORD DJ, ROUX H and WEYL OLF (2014) Species-specific impact of introduced largemouth bass *Micropterus salmoides* in the Groot Marico Freshwater Ecosystem Priority Area, South Africa. *Afr. J. Aquat. Sci.* **39** 451–458. <https://doi.org/10.2989/16085914.2014.976169>
- KJELLAND ME, WOODLEY CM, SWANNACK TM and SMITH DL (2015) A review of the potential effects of suspended sediment on fishes: potential dredging-related physiological, behavioral, and transgenerational implications. *Environ. Syst. Decis.* **35** (3) 334–350. <https://doi.org/10.1007/s10669-015-9557-2>
- KLEYNHANS CJ (1999) The development of a fish index to assess the biological integrity of South African rivers. *Water SA* **25** (3) 265–278.
- KLEYNHANS CJ (2007) Module D: Fish Response Assessment Index in River EcoClassification: Manual for EcoStatus Determination (version 2). Joint Water Research Commission and Department of Water Affairs and Forestry report. WRC Report No. TT 330/08. Water Research Commission, Pretoria.
- KLEYNHANS CJ and LOUW M (2007) River EcoClassification: Manual for EcoStatus Determination (Version 2) Module A: EcoClassification and EcoStatus Determination. WRC Report No. TT 329/08. Water Research Commission, Pretoria.
- KLEYNHANS CJ and LOUW M (2009) The Rapid Habitat Assessment Method Manual. Water for Africa. Department of Water Affairs and Forestry Report No. RDM/ Nat/00/CON/0707. DWAF, Pretoria.
- KLEYNHANS CJ, LOUW MD and MOOLMAN J (2007) Reference Frequency of Occurrence of Fish Species in South Africa.

<https://doi.org/10.17159/wsa/2019.v45.i3.6745>

Available at <https://www.watersa.net>

ISSN 1816-7950 (Online) = Water SA Vol. 45 No. 3 July 2019

Published under a [Creative Commons Attribution 4.0 International Licence \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

- Report produced for the Department of Water Affairs and Forestry (Resource Quality Services) and the Water Research Commission. URL: <http://www.dwaf.gov.za/iwqs/rhp/eco/FROC/ReportFinalFROC.pdf>
- KOEHN K, BRYE KR and SCARLAT C (2011) Quantification of stormwater runoff using a combined GIS and curve number approach: a case study for an urban watershed in the Ozark Highlands, USA. *Urban Water J.* **8** (4) 255–265. <https://doi.org/10.1080/1573062X.2011.595802>
- LE S, JOSSE J and HUSSON F (2008) FactoMineR: An R Package for Multivariate Analysis. *J. Stat. Softw.* **25** (1) 1–18. URL <https://cran.r-project.org/web/packages/FactoMineR/index.html>
- LI L, ZHENG B and LIU L (2010) Biomonitoring and bioindicators used for river ecosystems: definitions, approaches and trends. *Procedia Environ. Sci.* **2** 1510–1524. <https://doi.org/10.1016/j.proenv.2010.10.164>
- MATOWANYIKA W (2010) Impact of Alexandra Township on the water quality of the Jukskei River. MSc thesis, University of the Witwatersrand, Johannesburg, South Africa.
- MCEWAN AJ and JOY MK (2009) Differences in the distributions of freshwater fishes and decapod crustaceans in urban and forested streams in Auckland, New Zealand. *New Zeal. J. Mar. Freshw. Res.* **43** 1115–1120. <https://doi.org/10.1080/00288330.2009.9626534>
- MEADOR MR and GOLDSTEIN RM (2003) Assessing water quality at large geographic scales: relations among land use, water physicochemistry, riparian condition, and fish community structure. *Environ. Manage.* **31** (4) 504–517. <https://doi.org/10.1007/s00267-002-2805-5>
- MEADOR MR, COLES JF and ZAPPIA H (2005) Fish assemblage responses to urban intensity gradients in contrasting metropolitan areas: Birmingham, Alabama and Boston, Massachusetts. *Am. Fish. Soc. Symp.* **47** 409–423.
- MISHRA A (2010) Assessment of water quality using principal component analysis: a case study of the river Ganges. *J. Water Chem. Technol.* **32** (4) 227–234. <https://doi.org/10.3103/S1063455X10040077>
- MWANGI FN (2014) Land use practices and their impact on the water quality of the upper Kuils River (Western Cape Province, South Africa). PhD thesis, University of Western Cape, South Africa.
- NYENJE PM, FOPPEN JW, UHLENBROOK S, KULABAKO R and MUWANGA A (2010) Eutrophication and nutrient release in urban areas of sub-Saharan Africa—a review. *Sci. Total Environ.* **408** (3) 447–455. <https://doi.org/10.1016/j.scitotenv.2009.10.020>
- OLLIS DJ, DALLAS HF, ESLER KJ and BOUCHER C (2006) Bioassessment of the ecological integrity of river ecosystems using aquatic macroinvertebrates: an overview with a focus on South Africa. *Afr. J. Aquat. Sci.* **31** (2) 205–227. <https://doi.org/10.2989/16085910609503892>
- OLIVEIRA SV and CORTES RMV (2006) Combining logistic models with multivariate methods for the rapid biological assessment of rivers using macroinvertebrates. *Environ. Monit. Assess.* **112** 93–113. <https://doi.org/10.1007/s10661-006-0766-5>
- PARSONS TR, MAITA Y and LALLI CM (1984) *A Manual of Chemical and Biological Methods for Seawater Analysis*. Pergamon Press, New York. 173 pp.
- PASSELL HD, DAHM CN and BEDRICK EJ (2007) Ammonia modeling for assessing potential toxicity to fish species in the Rio Grande, 1989–2002. *Ecol. Appl.* **17** (7) 2087–2099. <https://doi.org/10.1890/06-1293.1>
- PAUL MJ and MEYER JL (2001) Streams in the urban landscape. *Annu. Rev. Ecol. Syst.* **32** (1) 333–365. <https://doi.org/10.1146/annurev.ecolsys.32.081501.114040>
- PECK DV, AVERILL DK, HERLIHY AT, HUGHES RM, KAUFMANN PR, KLEMM DJ, LAZORCHAK JM, MCCORMICK FH, PETERSON SA, CAPPAERT MR and co-authors (2006) Environmental monitoring and assessment program – surface waters western pilot study: field operations manual for non-wadeable rivers and streams. EPA/620/R-06/003, 2006. United States Environmental Protection Agency, Washington DC.
- R CORE TEAM (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- REID DJ, LAKE PS, QUINN GP and REICH P (2008) Association of reduced riparian vegetation cover in agricultural landscapes with coarse detritus dynamics in lowland streams. *Mar. Freshw. Res.* **59** (11) 998–1014. <https://doi.org/10.1071/MF08012>
- ROUX DJ (1999) Design of a national programme for monitoring and assessing the health of aquatic ecosystems, with specific reference to the South African river health programme. *Environ. Sci. For.* **96** 13–32.
- ROUX F and SELEPE M (2015) EcoStatus of the Komati River Catchment, Inkomati River System. Mpumalanga Tourism and Parks Agency, South Africa.
- RYAN PA (1991) Environmental effects of sediment on New Zealand streams: a review. *New Zeal. J. Mar. Freshw. Res.* **25** (2) 207–221. <https://doi.org/10.1080/00288330.1991.9516472>
- SKELTON PH (2001) *A Complete Guide to the Freshwater Fishes of Southern Africa*. Struik, Cape Town.
- SHETTY A, VENKATESHWARLU M and MURALIDHARAN M (2015) Effect of water quality on the composition of fish communities in three coastal rivers of Karnataka, India. *Int. J. Aquat. Biol.* **3** (1) 42–51. <https://doi.org/10.22034/ijab.v3i1.46>
- SMITH VH, TILMAN GD and NEKOLA JC (1999) Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environ. Pollut.* **100** (1) 179–196. [https://doi.org/10.1016/S0269-7491\(99\)00091-3](https://doi.org/10.1016/S0269-7491(99)00091-3)
- THLAKU N (2017) Personal communication, 07 July 2017. Ms Neo Thlaku, Corporate Social Investment and Stakeholder Relations, Ekurhuleni Water Care Company (ERWAT), P.O. Box 13106, Norkeem Park, 1631, South Africa.
- WARD EW and WINTER K (2016) Missing the link: urban stormwater quality and resident behaviour. *Water SA* **42** (4) 571–576. <https://doi.org/10.4314/wsa.v42i4.07>
- WENGER SJ, PETERSON JT, FREEMAN MC, FREEMAN BJ and HOMANS DD (2008) Stream fish occurrence in response to impervious cover, historic land use, and hydrogeomorphic factors. *Can. J. Fish. Aquat. Sci.* **65** (7) 1250–1264. <https://doi.org/10.1139/F08-046>
- WEPENER V (2008) Application of active biomonitoring within an integrated water resources management framework in South Africa. *S. Afr. J. Sci.* **104** 367–373.
- WEPENER V, DLAMINI P, O'BRIEN GC and MALHERBE W (2015) Development of a relative risk assessment framework to assess multiple stressors in the Klip River system. WRC Report No. 22404/1/15. Water Research Commission, Pretoria.
- WEYL OLF, DAGA VS, ELLENDER BR and VITULE JRS (2016) A review of *Clarias gariepinus* invasions in Brazil and South Africa. *J. Fish Biol.* **89** 386–402. <https://doi.org/10.1111/jfb.12958>

Water quality in non-perennial rivers

JA Day^{1*}, HL Malan¹, E Malijani¹ and AP Abegunde¹

Institute for Water Studies, University of the Western Cape, Robert Sobukwe Avenue, 7530 Bellville, South Africa

ABSTRACT

More than half the river-lengths of rivers in southern Africa dry up occasionally or – more commonly – seasonally. Here we review the literature on water quality (WQ) in non-perennial rivers (N-PRs), with emphasis on river management and southern African systems. Hydrological regimes cover a spectrum from relatively predictable and unvarying in perennial rivers, to unpredictable and highly variable in non-perennial rivers, which are complex, continually shifting mosaics of flowing water, standing-water pools and terrestrial habitats. N-PRs are uncommonly difficult to manage because they represent a limited source of water that is renewed unpredictably and is competed for by local people as well as being required by wildlife. Groundwater, and therefore its chemical and physical features, contributes significantly to base flow and to the maintenance of pools remaining in the bed when the river is not flowing. Water chemistry reflects catchment geology except in polluted systems. Salinity varies temporally, and spatially over three dimensions, and is the variable controlling the composition of the biotas of many N-PRs. Hydrological regimes are seldom predictable with any certainty; WQ varies naturally over time and space; groundwater often determines the WQ of surface water, especially in pools; and WQ in non-perennial rivers and pools may be affected by activities far upstream in the catchment. As yet we have no more than a sketchy understanding of the extent to which data on any one system can be applied to any other. Until we have a better understanding of these systems, the following basic principles should guide the management of WQ in N-PRs: (i) Rivers need to be assessed on a case-by-case basis. (ii) Understanding of the groundwater regime, including its chemistry, is crucial. (iii) Effluents need to be controlled by conservative effluent standards set for both ground and surface waters. (iv) Flows may need to be augmented at certain times of the year.

Keywords: non-perennial, rivers, water quality, review, water management, southern Africa

INTRODUCTION

Many rivers cease to flow from time to time and some even dry up completely for a few days, a few months, or a few years at a stretch. All of these can be said to be non-perennial rivers (N-PRs), in that they do not behave like ‘normal’ rivers that flow all the time. Various terms are used to describe the degree of non-perenniality of rivers. Datry et al. (2014a), for instance, classify as **intermittent** all rivers in which flow is temporary, ephemeral, seasonal or episodic. These terms are – confusingly – used differently by different authors. In this paper we use the terminology of Seaman et al. (2010), who developed a conceptual framework for non-perenniality (Fig. 1). In Fig. 1 dashed lines indicate that the boundaries are not fixed: for instance, some apparently perennial rivers may cease to flow in extreme drought conditions (Rossouw et al., 2005). A proportion of N-PRs are seasonal, ceasing to flow and even drying up for some months every year; these are ‘ephemeral’ or ‘semi-permanent’. (Note that the term ‘ephemeral’ is used – more correctly – in studies of temporary wetlands to refer to those in which water occurs only fleetingly after rain – see, for instance, Keddy, 2010.)

Beyond cessation of flow, a river becomes a landscape of isolated longitudinal pools (Fazi et al., 2013), a phenomenon that has profound consequences for its biota and also for humans who use its waters. Uys and O’Keeffe (1997) present a conceptual framework suggesting ways in which predictability/variability of flow, connectivity of aquatic habitats, natural disturbance, and biotic/abiotic controls change with

hydrological state. Both biotic and abiotic forces influence the species present and thus affect community structuring. As will be seen later in this review, the hydrological state, flow variability, aquatic connectivity and, to some extent, vegetation and faunal community structure, can all influence the concentration of chemical constituents and the values of physical variables in both the water column and the groundwater. These factors therefore need to be kept in mind when interpreting water quality (WQ).

Non-perennial rivers are complex, continually shifting mosaics of flowing-water (lotic), standing-water (lentic) and terrestrial habitats (Datry et al., 2014a). Such systems tend to be ecologically fragile (Rossouw et al., 2005) and yet they are often exceptionally important to people living in the vicinity. Particularly in arid areas, they support vegetation and associated animals and represent scarce resources in a dry terrestrial landscape (Jacobson, 1997). According to Rossouw et al. (2005), nearly half of the lengths of South African rivers are non-perennial and globally such rivers are far more common (Datry et al., 2014a) than is generally recognised. Furthermore, the number and geographical extent of N-PRs is increasing world-wide as a result of climate change and increased water abstraction for human use (Von Schiller et al., 2011; Jaeger et al., 2014; Kibria, 2016).

Historically, ecologists have tended to avoid working in N-PRs (Larned et al., 2010; Datry et al., 2014a) at least partly because their variable hydrology makes them difficult to study. Rossouw et al. (2005), Seaman et al. (2010) and Seaman et al. (2016 a; b) were pioneers in the field in South Africa, providing useful insights into the hydrology, geohydrology, geomorphology, water quality, ecology, and people–ecosystem interactions of N-PRs. More recently, international attention

*Corresponding author, email: jenny.day@uct.ac.za

Received 31 August 2018; accepted in revised form 25 June 2019

has increasingly been focused on these systems, as evidenced by the European MIRAGE (Mediterranean Intermittent Rivers Management: Prat et al., 2014; www.mirage-project.eu) and IRBAS (Intermittent River Biodiversity Analysis and Synthesis: www.irbas.fr) projects. The current paper is part of a multi-disciplinary project aimed at improving our understanding of the relationships between river flow, groundwater, and ecosystem characteristics, in order to inform decision-making related to the ecological and social consequences of modifying the hydrology of N-PRs. Here we review water quality in these non-perennial systems, discussing the major drivers influencing the concentrations of chemical constituents and the magnitudes of physical variables in these systems, as well as the ways in which they differ from perennial systems, and in which key WQ variables interact within them. Finally, some principles for managing WQ in N-PRs are presented. In this review, we make use of the literature pertaining to WQ in wetlands as well as N-PRs, since during periods of low flow N-PRs are effectively

linear systems of wetlands with varying degrees of hydrological connectivity. More comprehensive reviews of WQ in aquatic ecosystems from a South African perspective are to be found in DWAF (1996); Davies and Day (1998); Malan and Day (2002; 2012) and Day and Dallas (2011).

Effects of interactions between water quality variables on living organisms

Individual WQ variables can influence each other, acting synergistically (e.g. the effect of two or more chemicals together having a greater effect than either alone), or antagonistically (two or more agents in combination having a lesser effect than each alone). The major interactions that can occur between WQ parameters, and the resultant effects on a benthic community, are summarized in Fig. 2. These interactions are key to understanding WQ in aquatic ecosystems and are discussed individually below.

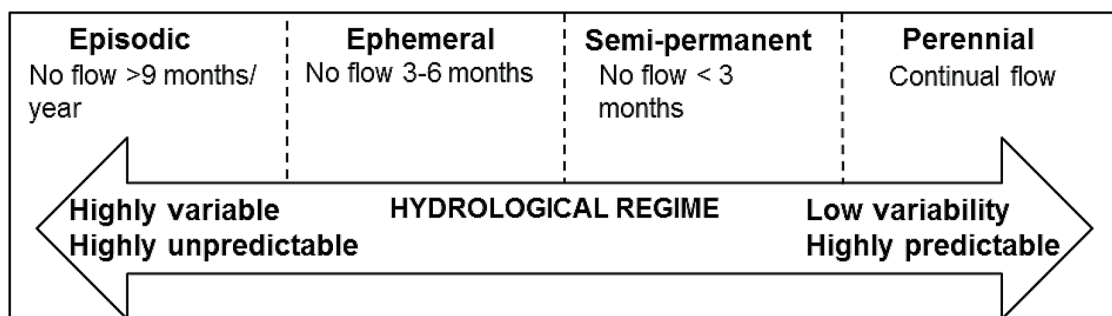


Figure 1. Conceptual framework of the hydrological continuum from perennial to non-perennial rivers (Seaman et al., 2010)

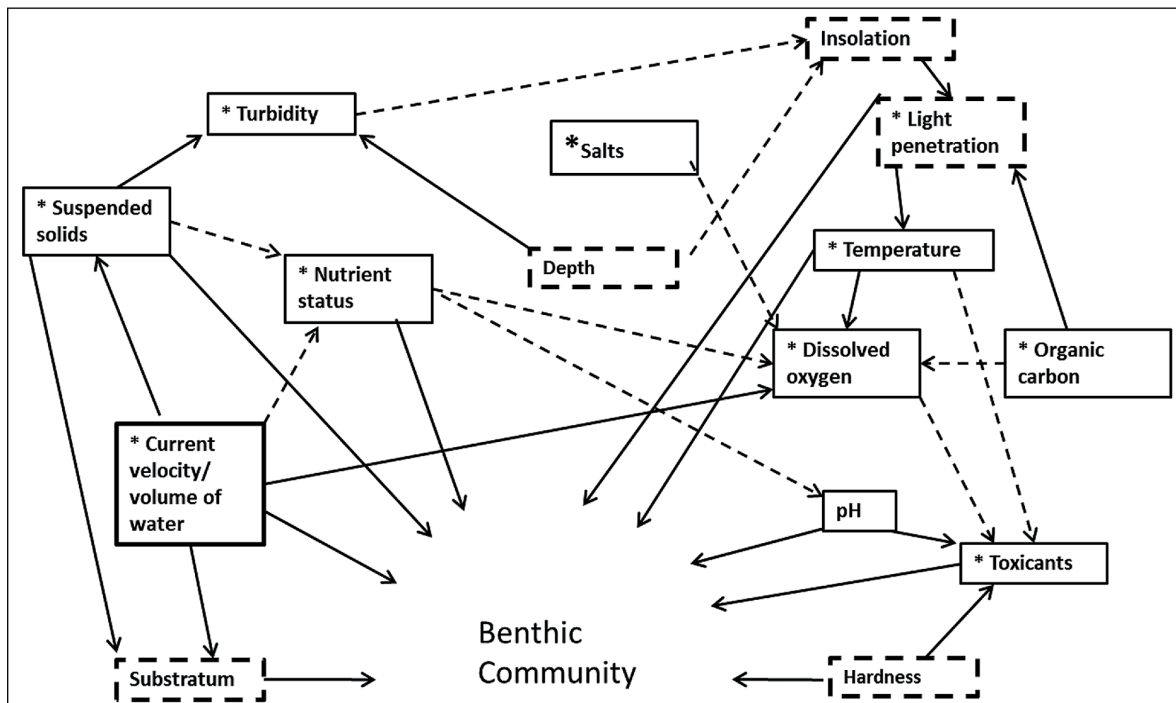


Figure 2. Interactions between water quality variables and other drivers of ecosystem functioning in rivers and wetlands. Water quality variables discussed in this review are enclosed in solid text boxes; those of peripheral importance are in dotted boxes. Direct effects are indicated by solid arrows and indirect effects by stippled arrows. Asterisks indicate aspects of particular relevance to N-PRs. Adapted from Hawkes (1979) and Dallas and Day (2004).

Surface and groundwater in non-perennial rivers

In N-PRs the single most important factor influencing water column concentrations of chemical constituents, and the magnitudes of physical variables, is the amount of water in the system (Rossouw et al., 2005). Larned et al. (2010) note that as pools become disconnected because of decreasing discharge they undergo abrupt thermal and chemical changes. The amount of water present determines the hydrological state of the river – whether it is in a state of flood, a small trickle, a longitudinal system of isolated pools, or dry river bed (with or without) hyporheic (under-bed) flow (Datry et al., 2016). The amount of water is likely to influence, amongst other factors, the concentrations of solutes (salts, toxins), nutrients (Von Schiller et al., 2011) and water temperature. As flow reduces, decreased turbulence will affect the entrainment of air within the water, and thus the concentration of DO (dissolved oxygen), and the entrainment of sediments from the bottom, and thus sediment loads – and turbidity – in the water (Mosley, 2015).

Although WQ in N-PRs is affected by the quantity and degree of movement of river water, the source of that water is also significant. Numerous studies (e.g. Rossouw et al., 2005; Seaman et al., 2010; Woelfle-Erskine et al., 2017) have emphasised the great importance of groundwater in non-perennial rivers, often contributing significantly to base flow (Rossouw et al., 2005). (It should be noted that, conversely, dryland rivers may also recharge aquifers: Rau et al., 2017). Where a connection between groundwater and surface water exists in the dry season, long-lasting pools may develop as a result of groundwater discharge (Seaman et al., 2013); shallow groundwater, rather than surface flow, is in fact the main source of 'new' water during the dry season in this type of system.

Understanding the link between ground and surface water is crucial for assessing the movement of water, and WQ, in N-PRs (Rau et al., 2017). Practically, it is fortunate that ground and surface waters in the same river often have different physical and chemical properties (e.g. Woelfle-Erskine et al. (2017), allowing natural components of WQ to be used for tracing the links between them. Various approaches, including isotopes and geochemistry, have been used for identifying and quantifying ground water–surface water interactions (Deodhar et al., 2014). The most commonly used isotopes in hydrology are Oxygen-18 (18O) and deuterium (2H), their isotopic compositions and ratios in water varying due to fractionation of isotopes during evaporation and condensation (Deodhar et al., 2014). Rau et al. (2017) have recently developed an approach that uses temperatures at different depths to establish relationships between ground- and surface water, based on the thermal diffusivity of sediments, which causes temperatures to vary with depth. From groundwater and surface water levels and the thermal signatures obtained, the different hydrological components can be identified.

Sub-surface flow in the hyporheic zone of many gravel- and sand-bed N-PRs represents a substantial proportion of the total discharge in the river, and exchange of nutrients and contaminants between hyporheic and surface water significantly affect biogeochemical processes occurring in these systems (Larsen et al., 2014). Not only may groundwater differ chemically and physically from surface water, but the variations in geohydrology and geochemistry over even a relatively small area can be complex. For example, Woelfle-Erskine et al. (2017) investigated two streams in California for several years over a period of increasing drought. Using electrical conductivity

(EC) and tracers to estimate the source of base flow, these authors identified several aquifers in the study region, and showed that base flow originated from heterogeneous sources.

One of the most significant principles regarding ecosystem functioning in N-PRs is the critical importance of pools (Bernardo and Alves, 1999; Rossouw et al., 2005; Seaman et al., 2010). From observations on the Seekoei River in the Orange River catchment in South Africa (Seaman et al., 2010), the amount of water in pools during periods of no-flow seemed to depend on the balance between groundwater discharge from springs and evaporation from the surface of the pool; evaporation would vary with climatic conditions and time of year. Various conceptual flow models (e.g. Watson and Burnet, 1993; Ward and Robinson, 2000) are available in the literature to describe groundwater/surface water interactions, including a useful model developed for the Mokolo River (Seaman et al. 2013), which describes hydrological processes involved specifically in N-PRs.

In relation to the above discussion, Seaman et al. (2010) found that not only the location, timing and persistence of pools along the length of a river were highly unpredictable, but that the WQ in such pools was also variable. Even in the same geomorphic reach, pools could vary in conductivity, for example, probably related to differences in the sources and quantities of groundwater entering the system. Such variability makes it difficult to predict the nature of water chemistry along the length of the river. Choy et al. (2002), who investigated four endorheic rivers in the central arid zone of Australia, also found that water chemistry varied both spatially and temporally in complex ways.

Pools represent refugia for aquatic organisms during periods when there is no flowing water. Parsons and Wentzel (2007) cite the findings of Brown et al. (2003) on the Doring River, an N-PR in the south-western Cape, South Africa. Brown and her colleagues showed that groundwater played a critical role in maintaining pools that provided refugia for animals, including fish, during periods of no flow. Interestingly, it was noted that indigenous fish used only those pools fed by fresh groundwater sources, while alien fish were found in all pools.

A consequence of the variability of WQ in N-PRs is that the organisms living in them are exposed to significant stresses, primarily desiccation and chemical variations in terms of salinity, ionic proportions, and concentrations of nutrients (Rossouw and Vos, 2008).

WATER QUALITY IN N-PRS DURING THE DRYING PHASE

In a recent key paper, Mosley (2015) summarizes the literature on drought and its impact on WQ separately for rivers and streams, and for lakes and reservoirs (Fig. 3). This paper has some useful observations for inferring WQ in N-PRs during drying or no-flow periods, although the results need to be interpreted with care because the findings for rivers and streams do not necessarily pertain to the non-flowing state. In addition, lakes and reservoirs are usually much bigger and deeper than pools, so different processes might be at play; arid areas seldom have permanent freshwater lakes of any size so there are few opportunities to compare permanent lakes and pools in N-PRs. Nevertheless, there is some useful information in Mosley's review and Fig. 3 is used as a template as we discuss predicted changes in WQ in N-PRs during the drying-down phase of the hydrological cycle. Note that the literature is sparse, and only the few papers and reports already mentioned deal with South African conditions.

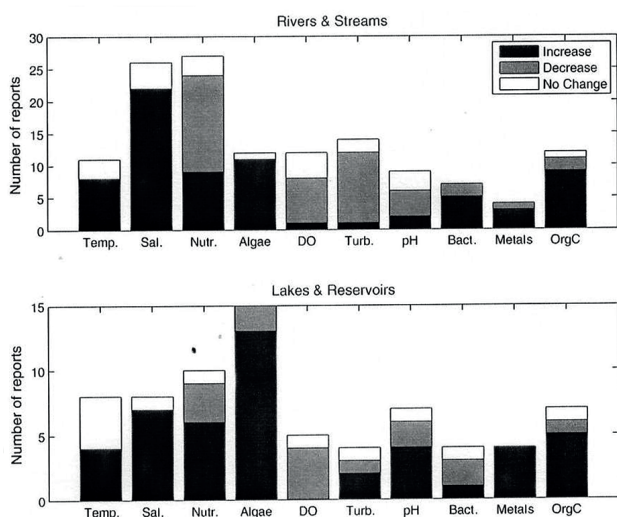


Figure 3. Summary of the effect (increase, decrease, no change) of drought on water quality variables for temperature (Temp.), salinity (Sal.), nutrients (Nutr.), algae, dissolved oxygen (DO), turbidity (Turb.), pH, bacteria (Bact.), metals and organic carbon (OrgC) in various kinds of aquatic ecosystem. Note that for most variables an increase in concentration is usually considered to be detrimental, except for pH and DO, where a decrease is usually considered detrimental. (From Mosley, 2015).

Temperature

General comments

Natural fluctuations in water temperature occur daily and seasonally, and indigenous organisms are adapted to such variations (Dallas, 2008). Organisms living in areas within their thermal ranges may not be affected by small changes in temperature, but the additional stress of even small increases in temperature may be critical when organisms are near the limit of their tolerance range (Sweeting, 1994). Furthermore, an increase in temperature enhances the toxic effects of some pollutants, including metals. A very important co-parameter is oxygen: oxygen solubility is depressed as temperature increases, leading to oxygen stress in susceptible organisms. Temperature has also been shown to affect the transport of fine sediments (e.g. Leliavsky, 1955). Below impoundments, for example, loads of suspended sediments may vary seasonally as a result of the higher viscosity of cold than of warm water (Leliavsky, 1955). Removal of riparian vegetation, and associated shading, may also result in temperature changes and, of course, in increased light penetration (Dallas, 2008).

Temperature in N-PRs

Water temperature generally increases in rivers when they are not flowing, especially if accompanied by an increase in air temperature (Lillebø et al., 2007). Larned et al. (2010) also note that as pools become disconnected by decreasing flow, diurnal temperature fluctuations increase in amplitude due to the decrease in thermal mass. According to Mosley (2015), thermal stratification in lakes may become more pronounced under conditions of drought; thermal stratification in deep pools of rivers is also likely to occur (Malan and Day, 2002) and may result in decreased DO levels with concomitant changes in the chemical environment, as discussed further on.

Salinity, electrical conductivity and total dissolved solids

General comments

The concentration of total dissolved solids (TDS) is the mass of all compounds dissolved in a sample of water, whereas electrical conductivity (EC) is a measure of the ionised substances within the water and represents a subset (usually a very large proportion) of the dissolved components. Salinity, a phenomenon used most commonly by marine chemists, also refers to the quantity of dissolved material, and is usually measured as EC. Note that it is often accepted practice to use the term 'salinity' for inland waters with high salt concentrations. An important point to note is that N-PRs are usually found in arid areas where ambient air temperatures are relatively high. As a consequence, evaporation rates are also elevated, increasing the salinity of standing water and damp soils (Rossouw et al., 2005; Gunkel et al., 2015).

Salinity in N-PRs

In isolated pools, salt concentrations are likely to increase through the process of evapo-concentration of solutes during the drying process (Rossouw et al., 2005; Woelfle-Erskine et al., 2017). A review of the literature by Mosley (2015) showed an increase in salinity to be the most likely outcome of drought in both rivers and lakes. Such increases are usually due to evaporation, but sometimes also to reduced dilution of saline groundwater, and in some cases to the inflow of effluents. In cases where the groundwater is not saline, or is fresher than surface water, salinities tend not to increase. Salinity was found to be particularly pronounced in highly seasonal, shallow systems near the point of complete dry-down.

In many arid environments salinity is elevated as a result of irrigation, clear-felling of trees, sewage return-flows, and river regulation (Walker, 2006). In this way naturally high salinities are exacerbated by human activities, sometimes leading to a loss of terrestrial and aquatic biodiversity, and of productive agricultural land. Bailey et al. (2006) contend that salt is the most under-recognised pollutant in the arid and semi-arid regions of the world, largely because salinization is seen as a problem only for agriculture. These authors provide a detailed review of 'salinization as an ecological perturbation to rivers, streams and wetlands of arid and semi-arid regions'.

Nutrients

General comments

Nutrients (often called 'limiting nutrients') are those elements required for plant growth; those of particular importance in aquatic ecosystems are nitrogen (N) and phosphorus (P). Eutrophication, or excessive accumulation of nutrients in a water body, is a serious problem worldwide, because it can result in rampant plant and algal growth (Ridge et al., 1995; Human et al., 2018). Many studies (e.g. Bouwman et al., 2017) have shown that the ratio of inorganic nitrogen to inorganic phosphorus is important for the normal functioning of aquatic ecosystems. Organic matter in unimpacted streams typically has an inorganic N:P ratio ranging between 25:1 and 40:1, whilst most impacted streams have ratios of <10:1.

Nutrients in N-PRS

Of all the variables considered in Fig. 3, nutrients respond with greatest diversity to drought in both lotic and lentic ecosystems. According to Mosley (2015), citing a range of authors, **lowered** concentrations of dissolved and total nutrients (N and P) during the drying phase of the hydrological cycle have been attributed to the following:

- Reduced input from the catchment because of reduced surface run-off
- Increased uptake by algae and macrophytes
- Increased denitrification due to longer residence times of the water

On the other hand, if point-sources or polluted groundwater enter the system during dry conditions, concentrations of nutrients tend to **increase** due to reduced dilution capacity. Boulton et al. (2006) note that food webs in floodplains of larger desert rivers may be driven by production of zooplankton, which in turn is initiated by pulsed nutrients and carbon associated with inundation of the floodplain. The fate of nutrients in N-PRs can therefore be of major significance in determining the ecological functioning of the entire ecosystem.

Stratification in water bodies can lead to alterations in internal processing of nutrients. Mosley (2015) cites studies reporting **increases** in ammonia, organic nitrogen and total phosphorus in the hypolimnion (bottom water) of stratified systems exhibiting low DO concentrations, a consequence of desorption of nutrients from the sediments. Reduction in stream flow that allows sediments to settle may also change the nutrient status of the ecosystem and hence the composition of the biota, since nutrients can desorb into the water column (see below). Some papers, on the other hand, report the turn-over of nutrient-enriched hypolimnetic water and resuspension of sediments as lakes become shallow and are acted upon by wind, with subsequent enhancement of algal growth. Pools in non-perennial rivers probably behave in the same way.

With regard to nutrient cycling, very little information is available on the fate of nutrients entering N-PRs. Seaman et al. (2010) investigated particle-bound nutrients in pools in the Seekoei River, South Africa, and found no significant trends in concentrations, although some seasonal patterns were identified. Concentrations of dissolved organic nitrogen increased during the dry winter months, for instance, while dissolved organic phosphorus decreased, presumably due to deposition in the sediments. McLaughlin (2008) found in Sycamore Creek, Arizona, USA, that dissolved nutrients became evaporites when the river was dry, coating sandbars and sometimes being blown away by the wind, along with fine silt and organic particles.

Lillebø et al. (2007), studying nutrient dynamics in the seasonal Mediterranean Fuirosos stream in Spain, found that the hydrological sequence of drying followed by re-wetting was the most important factor controlling nutrient levels in the water. It was during the period of dis-connectivity, when the water column was shallow and temperatures high, that nutrient interactions took place at the sediment/water interface. Von Schiller et al. (2011) investigated nutrient availability in a non-perennial Mediterranean stream, finding that nitrate concentrations ranged from 0.01 to 2.0 mg N/L, peaking in winter (when the river would be flowing). On the other hand, ammonium and soluble reactive phosphorus exhibited concentrations ranging from 0.005 to 0.15 mg N/L and 0 to 0.02 mg P/L, respectively, peaking in summer (during no-flow) due to anoxic conditions. Dissolved organic nitrogen showed

no clear seasonal pattern. These same authors developed a conceptual model for the complex pattern of changes in dissolved nitrogen (N) and phosphorus (P) availability. During the drying part of the hydrological cycle, when the system consists of isolated pools, nutrient availability and speciation are dominated by processes occurring within the pools. Because of the chemically reducing conditions, ammonium concentrations are higher than the more oxidised forms of nitrogen and some nitrogen is lost due to denitrification. Organic forms of nitrogen are released due to leaching from organic material and tend to be higher in concentration than the inorganic forms. Phosphorus is released from sediments due to the anoxic conditions, but concentrations of inorganic phosphorus are higher than those of organic phosphorus. Overall, temporal changes in concentration are more pronounced for N than for P. The authors emphasise the high spatial variability in nutrient concentrations whilst the stream was in this hydrological phase.

The conceptual model of Von Schiller et al. (2011) indicates the complex, often competing, processes that are likely to control instream concentrations of the various species of N and P. The extent to which it can be extrapolated to other non-perennial rivers is unclear. It is likely that similar processes occur in other systems but that the relative magnitudes and rates may differ. For example, in the above model, groundwater sources of nutrients are reported only during the re-wetting phase, whereas in many other non-perennial systems, groundwater sustains pools during the dry period. The extent to which a river is forested (shaded), land use in the catchment, and pollutant loads could all be expected to affect nutrient dynamics in freshwater systems and serve to emphasise the site-specific nature of water quality in N-PRs.

Algae and blue-greens

General comments

Still-water conditions and high concentrations of nutrients promote the growth of algae, which occur in fresh waters in two forms: as microscopic members of the phytoplankton, and as long filaments. Elevated concentrations of nutrients in fresh waters frequently result in eutrophication, which manifests as excessive blooms of phytoplankton, causing the water to become green, and mats of filamentous algae that can cover the surface of a pool, affecting the light regime. Certain cyanobacterial forms are toxic.

Algae and blue-greens in N-PRs

Pools in N-PRs often support a large biomass of algae, particularly filamentous forms, which represent a food source and refuge for insects and fish (Rossouw et al., 2005), although they also obscure the surface and sometimes prevent free diffusion of oxygen into the water. In addition, moderate temperatures, low turbidity and, frequently, increased nutrient concentrations, all combine to promote the growth of algae in the water column (phytoplankton) and on the bottom (benthic algae) (Nouri et al., 2009). In some cases there may be shifts from relatively benign algae to toxic cyanobacterial species during drought (Mosley, 2015). An extreme example of this phenomenon was a bloom of the toxic blue-green *Anabaena circinalis*, which extended down almost a thousand kilometres of the Barwon-Darling River in Australia. According to Donnelly et al. (1997) the bloom was the result

of low river flow during hot conditions, accompanied by stratification, clarification due to settling and flocculation of sediments, and phosphorus release from anoxic sulphate-reducing sediments. Several features of harmful algal species (e.g. *Anabaena* spp.), including tolerance of high temperatures, salt tolerance and the ability to fix nitrogen, make them particularly adapted to non-flowing freshwater bodies during drought (Paerl and Paul, 2012).

Dissolved oxygen (DO)

General comments

DO concentrations vary diurnally due to the competing processes of respiration and photosynthesis, as well as due to the effect of temperature on gas solubility. As a result of the uptake of oxygen by living organisms during respiration, DO levels in natural water bodies are at a minimum at dawn. During daylight, plants photosynthesize, releasing oxygen and resulting in maximal concentrations at mid-afternoon in eutrophic systems (Dallas et al., 1998). Sudden decreases in the DO content of a water body ('anoxic events') may have a severe impact on living organisms, especially animals, leading, for example, to fish-kills (e.g. La and Cooke, 2011). As noted previously, the solubility of oxygen in water is inversely proportional to salinity and temperature (Malan and Day, 2002).

Other factors having a pronounced effect on DO levels include the presence of excess oxidizable organic matter. The source may be anthropogenic (e.g. a carbon-rich effluent such as sewage) or natural (e.g. decaying organic matter) and can lead to oxygen depletion due to blooms of heterotrophic, oxygen-consuming bacteria (Gromiec et al., 1983). Metals (e.g. iron and manganese), as well as sulphides, can appear in solution under conditions of oxygen depletion (in addition to nutrients as noted previously), a situation that may be exacerbated at high temperatures (Ure and Davidson, 1995). According to Mitsch and Gosselink (2007), transformations of N, S, Fe, Mn, C and P occur in anoxic wetland sediments as a result of chemically reducing conditions and some of those transformations, for example that of S to H₂S, can result in the formation of toxins. Other important transformations include the processes of denitrification and methanogenesis, which release N₂ and methane to the atmosphere. Many of these processes are catalysed by anaerobic microbes.

Dissolved oxygen in N-PRs

The extent to which redox-driven chemical transformations will occur in a N-PR is likely to depend on the length of time pools have been hydrologically isolated (Mosley, 2015). The diffusion rate of oxygen between air and water has a marked effect on the oxygen content of water and this in turn is strongly influenced by turbulence, which is a function of current speed and stream bed characteristics (Campbell, 1982), and wind. Slowly flowing or stagnant waters, on the other hand, are prone to anoxia. Dissolved oxygen is usually not limiting to aquatic organisms in flowing systems, but concentrations tend to drop as flow ceases. As noted previously, the decomposition of organic matter (e.g. leaves, organic pollutants) also contributes to low DO levels (Chapman and Kramer, 1991). In the case of pools formed during the drying-out period of N-PRs, the extent to which pools might become anoxic is likely to depend on site-specific features such as wind, water depth, the extent of water movement (e.g. from springs, or in the hyporheos),

the length of time that water persists in the pool, the amount of carbonaceous material in the sediments, and the type and biomass of living organisms present – and thus the rate of respiration versus photosynthesis. If algae and macrophytes are abundant, diurnal DO (and pH) may increase during the day (Williams, 2006), whereas in pools with little primary production, DO may steadily decrease due to respiration and decreased solubility.

Turbidity and suspensoids

General comments

Changes in the quantity of suspended particles can affect the concentrations of inorganic ions, including nutrients, dissolved in the water column (Sweeting, 1994; Dallas et al., 1998). Phosphate, for example, binds to sediment particles, sometimes resulting in deficiencies of this nutrient in rivers carrying high sediment loads. This ameliorating effect may be reversed if sediments are disturbed or other system variables, such as pH, change. High levels of suspended sediments can also decrease water temperature due to reflection of heat by the suspended particles, resulting in reduced heat absorption by the water molecules (Kirk, 1985). A more obvious negative correlation exists between suspended sediments in the water column and the extent to which light penetrates into a water body. Reduced light penetration limits the potential for primary production by phytoplankton and rooted macrophytes because the depth of the photic zone (the zone in which enough light is available for photosynthesis) is limited (e.g. Grobler et al., 1987), while visually-searching predators may be unable to locate prey.

Turbidity and suspensoids in N-PRs

In rivers, suspended sediment loads normally decrease with decreasing flow (Fig. 2; and Hrdinka et al., 2012) due to settling out of sediments from the water column and decreased supply of sediments from the catchment, resulting in clear water in pools. In both Australia (e.g. Bunn et al., 2006) and South Africa, however, many perennial as well as N-PRs are permanently turbid. This turbidity is the result not of suspended organic particles such as algae or their remains, but of very finely-divided clay particles that are so small (<0.45 µm or so) and highly charged that they remain permanently suspended in the water column (Kirk, 1985). Very high levels of turbidity significantly limit light penetration and thus photosynthesis in pools in many N-PRs (Bunn et al., 2006). Alexandrov et al. (2003) examined the relationship between suspended sediment concentrations and discharge in the Eshtemoa (a dryland ephemeral channel) in the northern Negev, Israel. They showed that only half of the observed variation in the concentration of suspended sediments could be linked to variations in discharge, implying that variations in sediment supply, and changes in the importance of source areas, are equally significant in this semi-arid catchment (Alexandrov et al., 2003).

pH and alkalinity

General comments

pH may vary naturally from time to time. Streams in the south-western Cape, for example, tend to be more acidic during winter (Britton et al., 1993) than summer because of release of acidic

humic substances from the soil and groundwater during winter rains (Raubenheimer and Day, 1991). In productive ecosystems, on the other hand, as a result of removal of CO₂ during the process of photosynthesis, pH may reach values of 10 or more during the day, falling to 8 or less at night (Sweeting, 1994).

The pH of water can have a marked effect on the species (chemical form) of metal ions and nutrients, and hence on bioavailability (Filella et al., 1995). Decreasing pH frequently also leads to increased levels of metal cations in the water column as a result of desorption from the surfaces of suspended particles. The metals most profoundly affected by pH shifts in the 4–7 range include Al, Cu, Hg, Pb and Fe (Sweeting, 1994). Alkaline conditions, on the other hand, can result in enhanced levels of un-ionised ammonia (NH₃), which is toxic to living organisms. Casey and Farr (1982) also noted that the equilibrium between sediment-bound and free phosphate ions was sensitive to pH: as pH increased, phosphate was displaced from the sediments into the water column.

pH in N-PRs

Values for pH seem to exhibit great variability in response to drying (Fig. 3). Mosley (2015) summarizes the conflicting literature for this topic and cites several studies where it would appear that site-specific factors (e.g. sulphide oxidation in one Finnish catchment, and decreased dilution of bicarbonate-dominated groundwater in another), are the causes of a decrease or, in the second case, an increase, in pH. Von Schiller et al. (2011) found that in a largely un-impacted, forested Mediterranean stream, pH decreased as flow decreased in summer. Boulton et al. (2000) report that in some Australian N-PRs, pH can fall as low as 4.5 due to the leaching of organic acids from leaf litter as pools start to dry-up. Diurnal changes in pH due to photosynthesis have been reported in systems in which algae and macrophytes are abundant (Williams, 2006).

Organic carbon

General comments

Dissolved organic matter (DOM) and particulate organic matter (POM) are derived largely from decaying plant and animal matter and occur naturally in streams, rivers and wetlands. These substances are critical components of the food-chain, forming the energy base for detritivores and decomposer bacteria and fungi.

Organic carbon in N-PRs

As flow decreases and disconnected pools appear, large amounts of detritus may accumulate in pools and in the dry channel (Larned et al., 2010). With further drying of the river bed, there is a shift to processes characteristic of the dry phase. The breakdown of organic matter slows down and changes from degradation arising from leaching, and the action of aquatic microbes and invertebrates, to processes controlled by photodegradation, physical abrasion and decomposition mediated by terrestrial microbes and invertebrates (Datry et al., 2014a). Once the river flows again, much of the DOM (Guarch-Ribot and Butturini, 2016) and POM are flushed downstream. In rivers, the dynamics of accumulation and decomposition of organic matter are affected by drying and hydrological contraction to form pools. Dahm et al. (2003) reported decreases in the levels of DOM due to hydrological

disconnection in semi-arid intermittent rivers and ephemeral streams, although contrasting results have been reported by various authors. For example, studies on a semi-arid intermittent stream in Spain reported no change in DOM levels in surface waters during contraction (Ylla et al., 2010; Von Schiller et al., 2015). Von Schiller et al. (2015) hypothesise that differences between N-PRs are related to the balance between autotrophy and heterotrophy in each stream. Low river flow (especially during hydrological contraction) increases the accumulation and retention of both coarse (e.g. small woody debris and leaves) and fine POM (Dewson et al., 2007). Isolated pools accumulate detritus, resulting in extreme respiration, and limited re-aeration rates lead to hypoxia (Von Schiller et al., 2011) resulting in significant changes in the concentration and forms of DOM and nutrients. Von Schiller et al. (2011) also noted chemical changes in DOM, which tended to become less aromatic with time in isolated pools, suggesting an increase in the influence of both algal and microbial processes on DOM as the pools fragment. In addition, the concentration of low-molecular and potentially labile DOM may increase in older isolated pools. Further, these authors noted rapid changes in the composition and concentration of DOM just before the stream totally desiccated at the surface. They suggest that this may be due to rapid exudation of DOM from algae and microbes under stress, and ultimately from cell lysis (Von Schiller et al., 2015). Quantitatively, however, Casas-Ruiz et al. (2016) consider degradation and leaching of accumulated leaf litter and animal remains to be the major sources of DOM in rivers, stream and wetlands.

Toxic substances

General comments

Toxic substances form a very diverse group of chemicals, from metals to organics, including pesticides, and whole effluents. Furthermore, there is an ever-growing list of emerging contaminants derived from pharmaceuticals, toiletries and other products (Richardson and Ternes, 2011). Particulate matter acts as both a source of, and a sink for, numerous anthropogenic and natural contaminants. Many contaminating organic wastes entering aquatic systems ultimately accumulate in sediments, where they may negatively affect the benthic biota, as well as entering human and pelagic food chains. Organic enrichment is probably one of the commonest and best documented forms of pollution in rivers (Day and Dallas, 2011). It is usually due to a heterogeneous range of chemical compounds, from excreta in sewage to contaminants in waste streams from industrial processes. The major consequence is usually a decrease in ambient oxygen levels due to microbial uptake of oxygen during the process of decomposition. Other potential effects of organic enrichment include increased turbidity, suspended solids and nutrients. On a more positive note, both DOM and POM can be effective in binding metal ions and reducing their toxicity (ANZECC/ARMCANZ, 2000).

Toxic substances in N-PRs

The pollution profiles of N-PR basins are becoming similar to watersheds in perennial rivers, which are characterised by point sources and non-point sources. Many non-perennial streams are becoming, or have already become, channels for raw wastewater and treated effluents as sewage treatment procedures in many semi-arid regions move from septic tanks

and cesspools to major sewage systems, which affect both the morphology and the water of the stream (Hassan and Egozi, 2001; Tal et al., 2010; Köck-Schulmeyer et al., 2011). Evaporation, which is likely to concentrate toxicants, is an important driver in N-PRs, as are interactions such as changes in pH, or increased temperature, that govern the release and adsorption of toxicants from and to sediments.

Major manipulation of river-associated wetlands may have unexpectedly serious consequences. The San Joaquin River in California is a case in point. As a result of urban runoff, spillage and drainage of irrigation water, and reduction in virgin discharges, concentrations of some trace elements have increased dramatically. For instance, the concentrations of boron have increased from roughly 30 to > 2 900 µg L⁻¹ and of selenium from <1 to >6 µg L⁻¹, resulting in mass deformities and death of thousands of waterbirds and fish in the Kesterson wetlands along the river (Kingsford et al., 2006).

WATER QUALITY IN N-PRS DURING THE RE-WETTING PHASE

When flow resumes in N-PRs, depending on runoff intensity and the geomorphology of the catchment, the advancing wetted front can take the form of a flood bore (a wave), small rivulets, or expanding pools of upwelling groundwater (Jacobson et al., 2000; Jacobson and Jacobson, 2013; Datry et al., 2014a) and may be gradual or rapid (Datry et al., 2014b). Such events are unpredictable and can often, for example with flash floods occurring in deserts, be spectacular (Corti and Datry, 2012). Note that advancing and retreating fronts can occur also in the hyporheic zones of temporary rivers, as well as in river channels (Larned et al., 2010).

With the onset of rainfall and river flow, the hydrological link between the catchment and the river is renewed. Several studies (see Mosley, 2015) have found that during the re-wetting phase of the hydrological cycle nitrate is flushed from the catchment soils in the immediate post-drought period. This is likely to be due both to N having been concentrated in the catchment during the drought phase, and stimulation of mineralisation and nitrification when dried soils are rewetted. Leaching of organic matter and evaporite dissolution also cause spikes in solute concentrations in the waters of the newly flowing river (Jacobson et al., 2000; Hladysz et al., 2011; Corti and Datry, 2012). McLaughlin also found that rewetting following cessation of flow sometimes resulted in a release of significant quantities of organic nitrogen back into the river system. In contrast, Arce et al. (2015), found that in microcosm experiments, the process of rewetting sediments triggered denitrification.

In a similar way to rivers, the banks of reservoirs in semi-arid areas undergo desiccation during dry periods. Organic material on the banks becomes mineralized and after re-wetting N, P and C are dissolved into the water and released from the detritus (Gunkel et al., 2015). Thus, when flow resumes in the river, the concentrations of solutes and loads of suspended solids at the edge of the advancing front can be very high, and can exceed base flow concentrations by orders of magnitude. Organic matter, nutrients, and plant material (terrestrial plants, leaf-litter) that have accumulated on dry channel surfaces are transported downstream by advancing wetted fronts (Datry et al., 2014a). Non-perennial rivers are also likely to accumulate sediment (and in some situations associated contaminants) during periods of low-flow and lose them to flushing during periods of higher flow. In semi-arid

regions, suspended sediment loads are highest at the beginning of the wet season as flood waters scour the beds of particulate matter that has accumulated during the dry period. It has been estimated that re-suspended sediments may constitute more than 90% of the very high sediment loads in ephemeral dryland rivers in Israel (Powell et al., 1996). According to Datry et al. (2014a), advancing fronts in intermittent rivers can move many tons of organic matter to downstream receiving waters or to locations where flood attenuation results in deposition within the river channel. These deposits can be important carbon sources for heterotrophic organisms but they can also cause hypoxia (Rossouw et al., 2005) and fish and invertebrate kills (Hladysz et al., 2011). According to Datry et al. (2014a), the issue of large loads of organic matter released from intermittent rivers in pulses has not been fully studied. There may be consequences to the ecosystems of down-stream lakes, reservoirs and coastal environments in terms of eutrophication and oxygen depletion as well as greenhouse gas production and carbon sequestration.

During the re-wetting phase of a non-perennial stream, the dispersal of chemical constituents by the flowing water tends to homogenise nutrient concentrations, decreasing spatial heterogeneity (Chapman and Kramer, 1991; Von Schiller et al., 2011). The longer that flow persists in a N-PR, the more hydrological connectivity there will be between isolated pools, between groundwater and surface water, and between the river and the hyporheos and, as noted by Von Schiller et al. (2011), the greater the homogenisation of nutrients, but also of other WQ constituents. Malan and Day (2002) reviewed the literature concerning the effect of river discharge on instream WQ. They found that both catchment wash-off processes and instream processes influenced the resultant concentration of chemical constituents and values of physical variables. These findings were taken from the literature for perennial rivers; the longer water remains flowing within a non-perennial river, however, the more the trends described in Malan and Day (2002) for individual parameters will be relevant and can be used to predict WQ.

OVERARCHING PROCESSES DRIVING WATER QUALITY IN N-PRS

From his review of studies on the effect of drought on WQ in freshwater ecosystems, Mosley (2015) was able to identify three key factors or processes controlling the concentrations of chemical constituents and the magnitudes of physical variables. The first factor is – fairly obviously – mass balance. In a well-mixed water body at any point in time, the concentration of a chemical constituent is equal to the sum of all the masses from different sources (inflow, surface run-off, groundwater, point sources) divided by the total dilution volume. Processes such as biological or chemical transformation and sedimentation may also remove a constituent from the pool. Thus during drought, as volume decreases, the concentration will increase if the loading remains the same. The residence time of the remaining water also increases, causing sediments to settle out and supplies from the catchment to be cut off.

The second key factor identified by Mosley (2015) is temperature. During drought, ambient air temperatures commonly increase, accompanied by increases in water temperature. Consequently, another commonly observed effect of drought is an increased incidence and persistence of stratification and harmful algal blooms due to lower flows and increased thermal energy. Temperature increases and increased stratification result in lower dissolved oxygen concentrations.

The third key factor identified by Mosley (2015) is the increased importance of instream processes. During drought, when there is little or no flow in a river, not only is the residence time of the water in a given pool increased, but also that of chemical constituents, both dissolved and particulate; as a consequence, instream processes become more important than they are during the lotic state. For example, macrophytes and benthic algae become more prolific, possibly because of the longer residence time of nutrients and also because of the increased penetration of sunlight and decreased scouring effects. Other instream processes that Mosley (2015) identifies as important in controlling water chemistry are the movement of chemical constituents across the sediment/water interface, and sediment-associated processes such as denitrification.

In a key paper describing emerging concepts in temporary river ecology, Larned et al. (2010) consider that in temporary rivers, flow pulses trigger biogeochemical and physiological processes causing such systems to act as longitudinal, biogeochemical reactors. They describe how, because of the low moisture content in dry river beds, there are also low levels of microbial activity and consequently low rates of decomposition. With the onset of a flood event in a dry river bed, an advancing wetted front develops longitudinally in channels and laterally on floodplains. The advancing wetted front carries high loads of particulate organic matter (POM), sediments and organisms and also has high concentrations of dissolved chemical constituents. Through repeated cycles of flood pulses, material is moved downstream to a new reaction site (pool or sand bar) where it gradually dries out and water-dependent degradation processes cease.

MANAGEMENT OF WATER QUALITY IN NON-PERENNIAL RIVERS

Water quality and hydrology

In order to balance competing demands for water, many countries require water allocations (environmental flow requirements) to be specified for rivers, including N-PRs. Suitable allocations, with regard to both quality and quantity of water, have proved to be challenging (Stefanidis et al., 2016). Choy et al. (2002) recommend that, in setting water allocations for N-PRs, attention is paid to maintenance of natural periods of hydrological connectivity and of fragmentation, since periods of both are necessary for the maintenance of ecological integrity (Bishop-Taylor et al., 2017). In South Africa, the development of a method known as DRIFT-Arid is documented by Seaman et al. (2016a) and applied (Seaman et al. 2016b) to the Mokolo River in South Africa. (The DRIFT method for assessing water requirements for perennial rivers is documented in King and Brown, 2009).

Even for organisms adapted to living in unimpacted N-PRs, natural fluctuations in WQ variables such as water temperature, oxygen concentration and salinity are potentially lethal. When such rivers are additionally affected by disturbances such as water abstraction or the discharge of pollutants, the challenges to surviving in these systems are magnified. Deksissa et al. (2003) discuss management options for WQ improvement in the Crocodile River, Mpumalanga, South Africa. Whilst the river was probably perennial under natural conditions, it no longer is, so this study has some useful lessons for N-PRs in semi-arid areas that are subjected to pollution of different kinds and from different sources. These authors suggest that limiting water abstraction, setting and implementing effluent limits, and

augmenting flow during low-flow periods with releases from dams, will improve WQ (mainly by decreasing salinity and nutrient concentrations in this case) during periods of low flow. Kingsford et al. (2006) further discuss the effects of dams on desert rivers, including N-PRs.

Froebrich (2005) points out that it is difficult to predict WQ in temporary rivers because of the spatial heterogeneity in catchment processes, which leads to variability in contaminant build-up during the dry period, coupled with variability in the timing and extent of rainfall events. By implication, water chemistry needs to be assessed for each river and plans for managing it need to be based on the empirical data generated. Because of the overall lack of 'dilution capacity' in non-perennial systems an obvious way to protect them is to limit input of both point- and non-point sources of pollutants. An established method for reducing the input of sediments, nutrients and other contaminants into surface waters, especially N-PRs (Rosado et al. 2012) expanding during the wet autumn and winter period and contracting, often to a series of large pools, during the spring and summer. There can be considerable variation in annual rainfall patterns and current knowledge about the hydrology and dynamics of temporary streams is very limited. This means the confidence in the ability to predict the likely ecological consequences of climate change and increased water demand is low. To improve the existing level of understanding and predictive capability, a hydrological modelling study was carried out on the River Pardiela catchment in southern Portugal. The main objectives were to study long-term patterns of air temperature and precipitation, to apply the Soil and Water Assessment Tool (SWAT, is through the preservation of well-developed buffer strips of natural vegetation (e.g. Macfarlane et al., 2014; Lee and Fisher et al., 2016). It is during floods that large loads of sediment and contaminants enter the rivers (Froebrich, 2005) and it is also under these conditions that riparian vegetation is uprooted and flushed downstream. In very arid areas, though, there may well be no natural riparian vegetation, or the vegetation may consist of well-established trees within the water course, neither of which protects the river from lateral movement of sediments and other pollutants.

Choy et al. (2002) investigated the ecosystem integrity of four N-PRs located in the arid centre of Australia. Whilst the rivers at most of the study sites were found to be little affected by human activities, damage by livestock to the physically fragile banks was commonly noted. Similar impacts have been noted in the Doring River in the Western Cape (Paxton, 2017, pers. comm.). Rosado et al. (2012) consider exclusion of livestock from temporary Mediterranean streams as being a key factor in management, since uncontrolled roaming of livestock can damage stream beds, pollute the water, and cause health hazards to other users. Note that livestock play an important role in the economy and culture of subsistence farmers in Africa, so excluding them from riparian areas of N-PRs is extremely challenging, particularly where pools in river beds are the only sources of drinking water for stock animals. Note, too, that until recently, large herbivores, including megaherbivores such as elephants, were an integral part of the landscape even in arid areas in Africa; by implication, trampling must have been a greater factor in the natural perturbation of N-PRs than it is now.

Assessing environmental condition of N-PRs

Assessing the environmental condition ('ecological integrity') of any ecosystem requires information about some base state against

which the current condition of an ecosystem can be assessed. This base state, the 'reference condition', is usually considered to be the unimpacted pristine condition and is normally assessed separately for each type of aquatic ecosystem in each ecoregion. In such assessments it is necessary to distinguish between biotic responses to natural variability in hydrology and climate (i.e. drought, drying and flooding) and those brought about by anthropogenic impacts. Several authors (Arthington et al., 2014; Cid et al., 2016) have noted, however, that it is particularly difficult to assess the reference condition of N-PRs because of their dynamic nature and the sometimes decades-long long (but unmeasured) periodicity of drought and flood. Recognising this variability, Seaman et al. (2010) used the present state as the equivalent of the reference condition when developing a method for ascertaining the environmental water requirements of N-PRs. The predicted effects of climate change in the region – decreased rainfall, increased temperatures, greater variability and wider extremes (e.g. Schulze 2011) – will add to the challenge of assessing environmental condition in N-PRs.

Various approaches to assessing the ecosystem condition of N-PRs are reported in the literature. For non-perennial systems in the arid centre of Australia, Choy et al. (2002) used four aspects: the level of human influence; habitat condition; water chemistry; and aquatic invertebrates. In contrast, given challenges in the availability of water quality data, and the variability in WQ itself, Seaman et al. (2013) proposed the use of expert knowledge based on useful catchment information, including land use, potential pollution sources, soil types and geology. (Note that a similar approach has been used to make predictions of likely WQ in South African wetlands – Malan and Day, 2012.)

It seems that no bioassessment tools have successfully been developed for estimating the WQ of non-perennial rivers and wetlands. The well-known SASS (South African Scoring System) tool, based on invertebrate families and widely used in South Africa for assessing WQ and the ecosystem condition of perennial rivers, has proved not to be effective in wetlands (Bowd et al., 2006; Bird et al., 2013). Our hypothesis is that invertebrate assemblages will be no more effective in assessing the condition of N-PRs. As part of the MIRAGE project, however, Prat et al. (2014) have developed tools for assessing the ecosystem status of temporary Mediterranean rivers. According to Cid et al. (2016), though, these tools suffer the disadvantage of needing hydrological data, which is frequently lacking. In response these authors have developed a prototype tool (Bio-AS) for assessing temporary Mediterranean streams, using macro-invertebrates. The ratio EPT/OCH (i.e. the ratio between Ephemeroptera + Plecoptera + Trichoptera, and Odonata + Coleoptera + Hemiptera) has also been used to assess flow connectivity in Mediterranean streams (Bonada et al., 2006). It should be noted, though, that although these taxa can be useful in seasonal streams, rivers in very arid areas are unlikely to support more than one or two species, if any, of these taxa.

Water quality guidelines

In most countries, WQ criteria for aquatic ecosystems are designed for the protection of surface waters, particularly rivers. It is evident from the preceding discussions, though, that WQ in non-perennial rivers is naturally highly variable. Furthermore, the concentrations of some chemical constituents (e.g. salts, and therefore salinity) can rise markedly at the same time that others (e.g. DO) decrease (Mosley et al., 2012) and become limiting to the growth and even the survival of

some organisms. Von Schiller et al. (2011), for instance, note that because of the intrinsically high temporal and spatial variability in nutrient concentrations, reference conditions used to evaluate the status of permanent streams may not be appropriate for temporary streams. In fact, uncritical use in non-perennial systems of WQ criteria developed for perennial rivers might be harmful to some members of the biota (Rossouw et al., 2005). In particular, some of the natural values recorded for DO and salinity in N-PRs would, in perennial rivers, indicate poor WQ. In this regard, Feio et al. (2014) examined the data from a large number of national reference sites for perennial and non-perennial Mediterranean rivers from seven countries in an effort to characterize and define thresholds for the Least Disturbed Condition. Based on their results they were able to recommend a lower threshold value for DO for non-perennial than for perennial rivers.

Principles of water quality management in non-perennial rivers

It is possible to identify some general features of N-PRs:

- Hydrological regimes are seldom predictable with any certainty
- WQ varies naturally over time and space
- Groundwater often determines the WQ of surface water, especially in pools
- WQ in non-perennial rivers and pools may be affected by activities far upstream in the catchment

Furthermore, we still have no more than a sketchy understanding of the extent to which data on any one system can be applied to any other.

We therefore offer the following basic principles as guides to the management of these systems.

- Rivers need to be assessed on a case-by-case basis until such time as we can apply general principles to an understanding of WQ in N-PRs.
- Abstraction of both surface and groundwater, and storage of water in upstream dams, needs to be strictly limited and understanding of the groundwater regime is crucial if we are to avoid unsustainable 'mining' of the resource.
- Effluents need to be controlled and conservative effluent standards need to be set, sometimes on a case-by-case basis, for both ground and surface waters.
- Flows may need to be augmented at certain times of the year.
- Buffer zones need to be set, and where possible these should be designed to encourage the growth of natural vegetation.

It seems to us that the most useful step towards improving management of these systems would be the development of a much-simplified version of DRIFT-Arid for assessing water requirements for N-PRs, using additional test cases across the spectrum from episodic to semi-permanent systems and in different biomes. This should be linked to the development of a suitable monitoring programme for a number of N-PRs, particularly those for which water allocations have already been set. In seeking to deepen our knowledge of the ecological functioning of N-PRs it needs to be understood that because of the inherent variability of these systems, short-term investigations are of limited use and that study projects need to be long-term (10 years or more).

CONCLUSION

Processes taking place in N-PRs are poorly understood. Lack of knowledge, combined with the dynamic and sometimes

unpredictable nature of N-PRs, makes them challenging to manage. In consequence, intermittent rivers are particularly vulnerable in many parts of the world because of a lack of legislation, and therefore a lack of adequate management practices, protecting them and their waters (Datry et al., 2014a). Because of climate change, in southern Africa temperatures are expected to rise, rainfall to decrease, and inter-annual variability in rainfall to increase (Schulze, 2011), especially in those arid parts where temporary streams are located. Such conditions will invariably lead to an increased demand for water by humans, while the biotas of the rivers themselves will be subject to increased selective pressures; balancing conservation of N-PRs with human water demands will become more and more difficult over time (Rosado et al., 2012).

ACKNOWLEDGEMENTS

This paper is a contribution to a current project on N-PRs (https://www.uwc.ac.za/Faculties/NS/Water_Studies/Pages/non-perennial-rivers-group.aspx#Units) being carried out at the University of the Western Cape in South Africa and funded by the Shine Foundation. We thank three reviewers for their thoughtful and helpful comments.

REFERENCES

- ALEXANDROV Y, LARONNE JB and REID I (2003) Suspended sediment concentration and its variation with water discharge in a dryland ephemeral channel, northern Negev, Israel. *J. Arid Environ.* **53** (1) 73–84. <https://doi.org/10.1006/jare.2002.1020>
- ANZECC/ARMCANZ (2000) *Australian and New Zealand Guidelines for Fresh and Marine Water Quality* (NWQMS4, 2000). Australian and New Zealand Environmental and Conservation Council/Agriculture and Resource Management Council of Australia and New Zealand, Canberra, ACT. <https://doi.org/10.1111/anzs.2000.42.issue-3>
- ARCE MI, SÁNCHEZ-MONTOYA M, DEL M and GÓMEZ R (2015) Nitrogen processing following experimental sediment rewetting in isolated pools in an agricultural stream of a semiarid region. *Ecol. Eng.* **77** (3) 233–241. <https://doi.org/10.1016/j.ecoleng.2015.01.035>
- ARTHINGTON AH, BERNARDO JM and ILHÉU M (2014) Temporary rivers: linking ecophysiology, ecological quality and reconciliation ecology. *River Res. Applic.* **30** 1209–1215. <https://doi.org/10.1002/rra.2831>
- BAILEY PCE, BOON PI, BLINN DW and WILLIAMS WD (2006) Salinisation as an ecological perturbation to rivers, streams and wetlands of arid and semi-arid regions. Chapter 10. In: Kingsford R (ed.) *Ecology of Desert Rivers*. Cambridge University Press, Cambridge. 354 pp. ISBN: 0521818257 (hbk.), 9780521818254 (hbk.). 280–314.
- BERNARDO JM and ALVES MJ (1999) New perspectives for ecological flow determination in semi-arid regions: A preliminary approach. *Reg. Rivers: Res. Manage.* **15** 221–229. [https://doi.org/10.1002/\(sici\)1099-1646\(199901/06\)15:1/3<221::aid-rrr537>3.3.co;2-1](https://doi.org/10.1002/(sici)1099-1646(199901/06)15:1/3<221::aid-rrr537>3.3.co;2-1)
- BIRD MS, MLAMBO M and DAY JA (2013) Macroinvertebrates as unreliable indicators of human disturbance in temporary depression wetlands of the south-western Cape, South Africa. *Hydrobiologia* **720** 19–37. <https://doi.org/10.1007/s10750-013-1618-2>
- BISHOP-TAYLOR R, TULBURE MG and BROICH M (2017) Surface-water dynamics and land use influence landscape connectivity across a major dryland region. *Ecol. Appl.* **27**(4) 1124–1137. <https://doi.org/10.1002/eap.1507>
- BONADA N, RIERADEVALL M, PRAT N and RESH VH (2006) Benthic macroinvertebrate assemblages and macrohabitat connectivity in Mediterranean-climate streams of northern California. *J. N. Am. Benthol. Soc.* **25** (1) 32–43. [https://doi.org/10.1899/0887-3593\(2006\)25\[32:bmaamc\]2.0.co;2](https://doi.org/10.1899/0887-3593(2006)25[32:bmaamc]2.0.co;2)
- BOULTON AJ, SHELDON F and JENKINS KM (2006) Natural disturbance and aquatic invertebrates in desert rivers. Chapter 6 (pp.) In: Kingsford R (ed.) *Ecology of Desert Rivers*. Cambridge University Press, Cambridge. 354 pp. ISBN: 0521818257 (hbk.), 9780521818254 (hbk.). 133–153
- BOULTON AJ, SHELDON F, THOMS MC and STANLEY EH (2000) Problems and constraints in managing rivers with contrasting flow regimes. In: Boon PJ, Davies BR and Petts GE (eds) *Global Perspectives in River Conservation: Science, Policy and Practice*. John Wiley and Sons, Chichester. 415–430.
- BOUWMAN AF, BEUSEN AHW, LASSALETTA L, VAN APELDOORN DF, VAN GRINSVEN HJM, ZHANG J and VAN ITTERSUM MK (2017) Lessons from temporal and spatial patterns in global use of N and P fertilizer on cropland. *Sci. Rep.* **7** 40366. <https://doi.org/10.1038/srep40366>
- BOWD R, KOTZE DK, MORRIS CD AND QUINN NW (2006) Testing the applicability of the SASS5 scoring procedure for assessing wetland health: a case study in the KwaZulu-Natal midlands, South Africa. *Afr. J. Aquat. Sci.* **31** (2) 229–246. <https://doi.org/10.2989/16085910609503893>
- BRITTON DL, DAY JA and HENSHALL-HOWARD M-P (1993) Hydrochemical responses during storm events in a South African mountain catchment: the influence of antecedent conditions. *Hydrobiologia* **250** 143–157. <https://doi.org/10.1007/bf00008585>
- BROWN C, COLVIN C, HARTNADY C, HAY R, LE MAITRE D and RIEMAN K (2003) Ecological and environmental impacts of large-scale groundwater development in the Table Mountain Group. Discussion document. Water Research Commission, Pretoria, South Africa.
- BUNN SE, BALCOMBE SR, DAVIES PM, FELLOWS CS and MACKENZIE-SMITH FJ (2006) Aquatic productivity and food webs of desert river ecosystems. In: Kingsford R (ed.) *Ecology of Desert Rivers*. Cambridge University Press, Cambridge. 354 pp.
- CAMPBELL IC (1982) Biological water quality monitoring: An Australian viewpoint In: Hart BT (ed.) *Water Quality Management: Monitoring Programs and Diffuse Runoff*. Water Studies Centre, Chisholm Institute of Technology and Australian Society for Limnology (joint publishers), Melbourne.
- CASAS-RUIZ JP, TITTEL J, VON SCHILLER D, CATALÁN N, OBRADOR B, GÓMEZ-GENER L, ZWIRNMANN E, SABATER S and MARCÉ R (2016) Drought induced discontinuities in the source and degradation of dissolved organic matter in a Mediterranean river. *Biogeochemistry* **127** (1) 125–139. <https://doi.org/10.1007/s10533-015-0173-5>
- CASEY H and FARR IS (1982) The influence of within-stream disturbance on dissolved nutrient levels during spates. *Hydrobiologia* **92** 447–462. <https://doi.org/10.1007/bf02391959>
- CHAPMAN LJ and KRAMER DL (1991) Limnological observations of an intermittent tropical dry forest stream. *Hydrobiologia* **226** 153–166. <https://doi.org/10.1007/bf00006857>
- CHOY SC, THOMSON CB and MARSHALL JC (2002) Ecological condition of central Australian arid-zone rivers. *Water Sci. Technol.* **45** (11) 225–232. <https://doi.org/10.2166/wst.2002.0399>
- CID N, VERKAIK I, GARCIA-ROGER EM, RIERADEVALL M, BONADA N, SANCHEZ-MONTOYA MM, GOMEZ R, SUAREZ ML, VIDAL-ABARCA MR, DEMARTINI D and co-authors (2016) A biological tool to assess flow connectivity in reference temporary streams from the Mediterranean Basin. *Sci. Total Environ.* **540** 178–190. <https://doi.org/10.1016/j.scitotenv.2015.06.086>
- CORTI R and DATRY T (2012) Invertebrates and sestonic matter in an advancing wetted front travelling down a dry river bed (Albarine, France). *Freshwater Sci.* **31** 1187–1201. <https://doi.org/10.1899/12-017.1>
- DAHM CN, BAKER MA, MOORE DI and THIBAUT JR (2003) Coupled biogeochemical and hydrological responses of streams and rivers to drought. *Freshwater Biol.* **48** 1219–1223. <https://doi.org/10.1046/j.1365-2427.2003.01082.x>
- DALLAS H (2008) Water temperature and riverine ecosystems: an overview of knowledge and approaches for assessing biotic responses, with special reference to South Africa. *Water SA* **34** (3) 393–404.
- DALLAS HF and DAY JA (2004) The effect of water quality variables on aquatic ecosystems: a review. WRC Report No. TT 224/04. Water Research Commission, Pretoria.

- DALLAS HF, DAY JA, MUSIBONO DE and DAY EG (1998) Water quality for aquatic ecosystems: Tools for evaluating regional guidelines. WRC Report No. 626/1/98. Water Research Commission, Pretoria.
- DATRY T, LARNED ST and TOCKNER K (2014a). Intermittent rivers: A challenge for freshwater ecology. *BioScience* **64** (3) 229–235. <https://doi.org/10.1093/biosci/bit027>
- DATRY T, LARNED ST, FRITZ KM, BOGAN MT, WOOD PJ, MEYER EI and SANTOS AN (2014b) Broad-scale patterns of invertebrate richness and community composition in temporary rivers: effects of flow intermittence. *Ecography* **37**(1) 94–104. <http://dx.doi.org/10.1111/j.1600-0587.2013.00287.x>
- DATRY T, PELLA H, LEIGH C, BONADA N and HUGUENY B (2016) A landscape approach to advance intermittent river ecology. *Freshwater Biol.* **61** (8) 1200–1213. <https://doi.org/10.1111/fwb.12645>
- DAVIES B and DAY J (1998) *Vanishing Waters*. University of Cape Town Press, Cape Town, South Africa. 487 pp. <https://doi.org/10.15641/ghi.v2i1.728>
- DAY JA and DALLAS HF (2011) Understanding the basics of water quality. In: Grafton Q (ed.) *Water Resources Planning and Management*. Cambridge University Press, Cambridge. 777 pp. <https://doi.org/10.1017/cbo9780511974304.007>
- DEODHAR AS, ANSARI A, SHARMA S, JACOB N, KUMAR US and SINGH G (2014). *Isotope Techniques for Water Resources Management* Newsletter **337** 29–35. Bhabha Atomic Research Centre, Department of Atomic Energy, Government of India.
- DEKSISSA T, ASHTON PJ and VANROLLEGHEM PA (2003) Control options for river water quality improvement: A case study of TDS and inorganic nitrogen in the Crocodile River (South Africa). *Water SA* **29** (2) 209–218. <https://doi.org/10.4314/wsa.v29i2.4858>
- DEWSON ZS, JAMES ABW and DEATH RG (2007) Stream ecosystem functioning under reduced flow conditions. *Ecol. Appl.* **17** 1797–1808. <https://doi.org/10.1890/06-1901.1>
- DONNELLY TH, GRACE MR and HART BT (1997) Algal blooms in the Darling–Barwon River, Australia. *Water Air Soil Pollut.* **99** 487–496. <https://doi.org/10.1007/bf02406888>
- DWAF (Department of Water Affairs and Forestry, South Africa) (1996) South African Water Quality Guidelines. Volume 7: Aquatic ecosystems. Department of Water Affairs and Forestry, Pretoria.
- FAZI S, VÁZQUEZ E, CASAMAYOR EO, AMALFITANO S and BUTTURINI A (2013) Stream hydrological fragmentation drives bacterioplankton community composition. *PloS one* **8** (5) e64109. <https://doi.org/10.1371/journal.pone.0064109>
- FILELLA M, TOWN R and BUFFLE J (1995) Speciation in fresh waters. In: Ure AM and Davidson CM (eds.) *Chemical Speciation in the Environment*. Blackie Academic and Professional, New York.
- FEIO MJ, AGUIAR FC, ALMEIDA SFP, FERREIRA J, FERREIRA MT, ELIAS C, SERRA SRQ, BUFFAGNI A, CAMBRA J, CHAUVIN C and co-authors (2014) Least Disturbed Condition for European Mediterranean rivers. *Sci. Total Environ.* **476–477** 745–756. <https://doi.org/10.1016/j.scitotenv.2013.05.056>
- FROEBRICH J (2005) Water quality dynamics in temporary streams and implications for water management strategies in semi-arid regions. *Geophysical Research Abstracts (GRA)* **7** 01857. <http://www.cosis.net/abstracts/EGU05/01857/EGU05-J-01857.pdf>
- GROBLER DC, TOERIEN DF and ROSSOUW JN (1987) A review of sediment/water quality interaction with particular reference to the Vaal River system. *Water SA* **13** 15–22.
- GROMIEC MJ, LOUCKS DP and ORLOB GT (1983) Stream quality modelling. In: Orlob GT (ed.) *Mathematical Modelling of Water Quality: Streams, Lakes and Reservoirs*. John Wiley and Sons, New York.
- GUARCH-RIBOT A and BUTTURINI A (2016) Hydrological conditions regulate dissolved organic matter quality in an intermittent headwater stream. From drought to storm analysis. *Sci. Total Environ.* **571** 1358–1369. <https://doi.org/10.1016/j.scitotenv.2016.07.060>
- GUNKEL G, LIMA D, SELGE F, SOBRAL M and CALADO S (2015) Aquatic ecosystem services of reservoirs in semi-arid areas: sustainability and reservoir management. *WIT Trans. Ecol. Environ.* **197** 1743–3541. <https://doi.org/10.2495/rm150171>
- HASSAN MA and EGOZI R (2001) Impact of wastewater discharge on the channel morphology of ephemeral streams. *Earth Surf. Process. Landf.* **26** (12) 1285–1302. <https://doi.org/10.1002/esp.273>
- HAWKES HA (1979) Invertebrates as indicators of river water quality. In: James A and Evison L (eds.) *Biological Indicators of Water Quality*. John Wiley and Sons, Chichester.
- HLADY Z, WATKINS SC, WHITWORTH KL and BALDWIN DS (2011) Flows and hypoxic blackwater events in managed ephemeral river channels. *J. Hydrol.* **401** 117–125. <https://doi.org/10.1016/j.jhydrol.2011.02.014>
- HRDINKA T, NOVICKÝ O, HANSLÍK E and RIEDER M (2012) Possible impacts of floods and droughts on water quality. *J. Hydro-Environ. Res.* **6** (2) 145–150. <https://doi.org/10.1016/j.jher.2012.01.008>
- HUMAN LRD, MAGORO ML, DALU T, PERISSINOTTO R, WHITFIELD AK, ADAMS JB, DYSEL SHP and RISHWORTH GM (2018) Natural nutrient enrichment and algal responses in near pristine micro-estuaries and micro-outlets. *Sci. Total Environ.* **624** 945–954. <https://doi.org/10.1016/j.scitotenv.2017.12.184>
- JACOBSON PJ (1997) An ephemeral perspective of fluvial ecosystems: viewing ephemeral rivers in the context of current lotic ecology. PhD dissertation, Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Virginia.
- JACOBSON PJ and JACOBSON KM (2013) Hydrologic controls of physical and ecological processes in Namib Desert ephemeral rivers: Implications for conservation and management. *J. Arid Environ.* **93** 1–14. <https://doi.org/10.1016/j.jaridenv.2012.01.010>
- JACOBSON PJ, JACOBSON KM, ANGERMEIER PL and CHERRY DS (2000) Variation in material transport and water chemistry along a large ephemeral river in the Namib Desert. *Freshwater Biol.* **44** 481–491. <https://doi.org/10.1046/j.1365-2427.2000.00604.x>
- JAEGER KL, OLDEN JD and PELLAND NA (2014) Climate change poised to threaten hydrologic connectivity and endemic fishes in dryland streams. *PNAS* **111** (38) 13894–13899. <https://doi.org/10.1073/pnas.1320890111>
- KEDDY P (2010) *Wetland Ecology: Principles and Conservation*. (2nd edn.). Cambridge University Press, Cambridge. 495 pp.
- KIBRIA G (2016) World rivers in crisis: water quality and water dependent biodiversity are at risk- threats of pollution, climate change & dams development. URL: https://www.researchgate.net/publication/263852254_World_Rivers_in_Crisis_Water_Quality_and_Water_Dependent_Biodiversity_Are_at_Risk_-_Threats_of_Pollution_Climate_Change_and_Dams_Development. <https://doi.org/10.24926/2471190x.1546>
- KING J and BROWN C (2009) Integrated basin flow assessments: concepts and method development in Africa and South-east Asia. *Freshwater Biol.* **55** (1) 127–146. <https://doi.org/10.1111/j.1365-2427.2009.02316.x>
- KINGSFORD RT, LEMLEY AD and THOMPSON JR (2006) Impacts of dams, river management and diversion on desert rivers. Chapter 8 (pp203–248) in: Kingsford R (ed.) *Ecology of Desert Rivers*. Cambridge University Press, Cambridge. 354 pp.
- KIRK JTO (1985) Effects of suspensoids (turbidity) on penetration of solar radiation in aquatic ecosystems. *Hydrobiologia* **125** (1) 195–208. <https://doi.org/10.1007/bf00045935>
- KÖCK-SCHULMEYER M, GINEBREDA A, POSTIGO C, LÓPEZ-SERNA R, PÉREZ S, BRIX R, LLORCA M, DE ALDA ML, PETROVIĆ M, MUNNÉ A and co-authors (2011) Wastewater reuse in Mediterranean semi-arid areas: the impact of discharges of tertiary treated sewage on the load of polar micro-pollutants in the Lobregat river (NE Spain). *Chemosphere* **82** (5) 670–678. <https://doi.org/10.1016/j.chemosphere.2010.11.005>
- LA VT and COOKE SJ (2011) Advancing the science and practice of fish kill investigations. *Rev. Fish. Sci.* **19** (1) 21–33.
- LARNED ST, DATRY T, ARSCOTT DB and TOCKNER K (2010) Emerging concepts in temporary-river ecology. *Freshwater Biol.* **55** (4) 717–738. <https://doi.org/10.1111/j.1365-2427.2009.02322.x>
- LARSEN LG, HARVEY JW and MAGLIO MM (2014) Dynamic hyporheic exchange at intermediate timescales: testing the relative importance of evapotranspiration and flood pulses. *Water Resour. Res.* **50** 318–335. <https://doi.org/10.1002/2013wr014195>

- LEE J and FISHER C (2016) Arid green infrastructure for water control and conservation: state of the science and research needs for arid/semi-arid regions. Report prepared by The Scientific Consulting Group, Inc. for the Environmental Protection Agency, Las Vegas, NV 89193. <https://doi.org/10.13031/2013.13834>
- LELIAVSKY S (1955) Sediment suspension. In: *Introduction to Fluvial Hydraulics*. Constable and Co., New York. 181–191.
- LILLEBØ AI, MORAIS M, GUILHERME P, FONSECA R, SERAFIM A and NEVES R (2007) Nutrient dynamics in Mediterranean temporary streams: A case study in Pardiela catchment (Degebe River, Portugal). *Limnologica* 37 (4) 337–348. <https://doi.org/10.1016/j.limno.2007.05.002>
- MACFARLANE DM, BREDIN IP, ADAMS JB, ZUNGU MM, BATE GC and DICKENS CWS (2014) Preliminary guideline for the determination of buffer zones for rivers, wetlands and estuaries: Consolidated report. WRC Report No. TT 610/14. Water Research Commission, Pretoria, South Africa.
- MALAN HL and DAY JA (2002) Linking discharge, water quality and biotic response in rivers: a literature review. WRC Report No. 956/2/02. Water Research Commission, Pretoria.
- MALAN HL and DAY JA (2012) Water quality and wetlands: defining Ecological Categories and links with land-use. WRC Report No. 1921/1/12. Water Research Commission, Pretoria.
- MCLAUGHLIN C (2008) Evaporation as a nutrient retention mechanism at Sycamore Creek, Arizona. *Hydrobiologia* 603 241–252. <https://doi.org/10.1007/s10750-007-9275-y>
- MITSCH WJ and GOSSELINK JG (2007) *Wetlands* (4th edn.). John Wiley & Sons, Inc., Hoboken, New Jersey. 582 pp.
- MOSLEY LM (2015) Drought impacts on the water quality of freshwater systems: review and integration. *Earth Sci. Rev.* 140 203–214. <https://doi.org/10.1016/j.earscirev.2014.11.010>
- MOSLEY LM, ZAMMIT B, LEYDEN E, HENEKER TM, HIPSEY MR, SKINNER D and ALDRIDGE KT (2012) The impact of extreme low flows on the water quality of the Lower Murray River and Lakes (South Australia). *Water Resour. Manage.* 26(13) 3923–3946. <https://doi.org/10.1007/s11269-012-0113-2>
- NOURI J, MIRBAGHERI SA, FARROKHIAN F, JAAFARZADEH N and ALESHEIKH AA (2009) Water quality variability and eutrophic state in wet and dry years in wetlands of the semiarid and arid regions. *Environ. Earth Sci.* 59 (7) 1397–1407. <https://doi.org/10.1007/s12665-009-0126-1>
- PAERL HW and PAUL VJ (2012) Climate change: links to global expansion of harmful cyanobacteria. *Water Res.* 46 1349–1363. <https://doi.org/10.1016/j.watres.2011.08.002>
- PARSONS R and WENTZEL J (2007) Groundwater Resource Directed Measures Manual. WRC Report No. TT299/07. Water Research Commission, Pretoria.
- PAXTON B (2017) Personal communication, April 2017. B. Paxton, Freshwater Research Centre, Imhoff's Gift, Cape Town.
- POWELL DM, REID I, LARONNE JB and FROSTICK L (1996) Bed load as a component of sediment yield from a semi-arid watershed of the northern Negev. *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences* 236 389–398.
- PRAT N, GALLART F, VON SCHILLER D, POLESSELLO S, GARCÍA-ROGER EM, LATRON J, RIERADEVALL M, LLORENS P, BARBERÁ GG, BRITO D and co-authors (2014) The MIRAGE toolbox: An integrated assessment tool for temporary streams. *River Res. Appl.* 30 1318–1334. <https://doi.org/10.1002/rra.2757>
- RAU GC, HALLORAN LJS, CUTHBERT MO, ANDERSEN MS, ACWORTH R I and TELLAM JH (2017) Characterising the dynamics of surface water-groundwater interactions in intermittent and ephemeral streams using streambed thermal signatures. *Adv. Water Resour.* 107 354–369. <https://doi.org/10.1016/j.advwatres.2017.07.005>
- RAUBENHEIMER CM and DAY JA (1991) *In vitro* leaching of two sclerophyllous fynbos plants. *Hydrobiologia* 224 167–174. <https://doi.org/10.1007/bf00008466>
- RICHARDSON SD and TERNES TA (2011) Water analysis: Emerging contaminants and current issues. *Anal. Chem.* 83 4616–4648. <https://doi.org/10.1021/ac200915r>
- RIDGE T, PILLINGER P and WALTERS S (1995) Alleviating the problems of excessive algal growth. In: Harper DM and Ferguson AJD (eds.) *The Ecological Basis for River Management*. John Wiley & Sons, Chichester. 630 pp.
- ROSADO J, MORAIS M, SERAFIM A, PEDRO A, SILVA H, POTES M, BRITO D, SALGADO R, NEVES R, LILLEBØ A and co-authors (2012) Key factors in the management and conservation of temporary Mediterranean streams: A case study of the Pardiela River, Southern Portugal. In: Boon PJ and Raven PJ (eds) *River Conservation and Management*. John Wiley & Sons, Ltd, Hoboken, New Jersey. <https://doi.org/10.1002/9781119961819.ch22>
- ROSSOUW L, AVENANT MF, SEAMAN MT, KING JM, BARKER CH, DU PREEZ PJ, PELSER AJ, ROOS, JC, VAN STADEN JJ, VAN TONDER GJ and co-author (2005) Environmental water requirements in non-perennial systems. WRC Report No 1414/1/05. Water Research Commission, Pretoria.
- ROSSOUW L and VOS T (2008) Water quality component. In: Seaman et al. (2010) Developing a method for determining the environmental water requirements for non-perennial systems. WRC Report No. TT 459/10. Water Research Commission, Pretoria.
- SCHULZE RE (2011) Approaches towards practical adaptive management options for selected water-related sectors in South Africa in a context of climate change. *Water SA* 37 (5) 621–645. <https://doi.org/10.4314/wsa.v37i5.1>
- SEAMAN MT, AVENANT MF, WATSON M, KING JM, ARMOUR J, BARKER CH, DOLLAR E, DU PREEZ PJ, HUGHES D, ROSSOUW L and co-author (2010) Developing a method for determining the environmental water requirements for non-perennial systems. WRC Report No. TT 459/10. Water Research Commission, Pretoria.
- SEAMAN MT, WATSON M, AVENANT MF, JOUBERT AR, KING JM, BARKER CH, ESTERHUYSE S, GRAHAM D, KEMP ME, LE ROUX PA and co-authors (2013) Testing a methodology for environmental water requirements in non-perennial rivers. The Mokolo River case study. WRC Report No. TT 579/13. Water Research Commission, Pretoria.
- SEAMAN M, WATSON M, AVENANT M, KING J, JOUBERT A, BARKER C, ESTERHUYSE S, GRAHAM D, KEMP M, LE ROUX P and co-authors (2016a) DRIFT-ARID: A method for assessing environmental water requirements (EWRs) for non-perennial rivers. *Water SA* 42 (3) 356–367. <https://doi.org/10.4314/wsa.v42i3.01>
- SEAMAN M, WATSON M, AVENANT M, KING J, JOUBERT A, BARKER C, ESTERHUYSE S, GRAHAM D, KEMP M, LE ROUX P and co-authors (2016b) DRIFT-ARID: Application of a method for environmental water requirements (EWRs) in a non-perennial river (Mokolo River) in South Africa. *Water SA* 42 (3) 368–383. <http://dx.doi.org/10.4314/wsa.v42i3.02>
- STEFANIDIS K, PANAGOPOULOS Y, PSOMAS A and MIMIKOU M (2016) A methodological approach for evaluating ecological flows using ecological indicators: a case study in River Pinios, Greece. In: *1st GLOBAQUA International Conference Managing The Effects Of Multiple Stressors On Aquatic Ecosystems Under Water Scarcity*. 11–12 January 2016, Freising (Germany).
- SWEETING RA (1994) River pollution. In: Calow P and Petts GE (eds.) *The Rivers Handbook*, Vol. 2. Blackwell Scientific Publications, Oxford. 23–32.
- TAL A, AL KHATEEB N, NAGOUEK N, AKERMAN H, DIABAT M, NASSAR A, ANGEL R, AL SADAH MA., HERSHKOVITZ Y, GASITH A and co-author (2010) Chemical and biological monitoring in ephemeral and intermittent streams: a study of two transboundary Palestinian–Israeli watersheds. *Intl. J. River Basin Manage.* 8 (2) 185–205. <https://doi.org/10.1080/15715124.2010.491796>
- URE AM and DAVIDSON CM (1995) *Chemical Speciation in the Environment*. Chapman & Hall, Glasgow. 408 pp.
- UYS MC and O'KEEFE JH (1997) Simple words and fuzzy zones: Early directions for temporary river research in South Africa. *Environ. Manage.* 21 (4) 517–531. <https://doi.org/10.1007/s002679900047>
- VON SCHILLER D, ACUÑA V, GRAEBER D, MARTÍ E, RIBOT M, SABATER S, TIMONER X and TOCKNER K (2011) Contraction, fragmentation and expansion dynamics determine nutrient availability in a Mediterranean forest stream. *Aquat. Sci.* 73 (4) 485–497. <https://doi.org/10.1007/s00027-011-0195-6>

- VON SCHILLER D, GRAEBER D, RIBOT M, TIMONER X, ACUÑA V, MARTÍ E, SABATER S and TOCKNER K (2015) Hydrological transitions drive dissolved organic matter quantity and composition in a temporary Mediterranean stream. *Biogeochemistry* **123** (3) 429–446. <https://doi.org/10.1007/s10533-015-0077-4>
- WALKER KF (2006) The lower River Murray, Australia. Chapter 9. In: Kingsford R (ed.) *Ecology of Desert Rivers*. Cambridge University Press, Cambridge. 354 pp. 248–279.
- WATSON I and BURNET AD (1993) *Hydrology: An Environmental Approach*. Buchanan Books, Fort Lauderdale, USA. 702 pp.
- WILLIAMS DD (2006) *The Biology of Temporary Waters*. Oxford University Press, Oxford, New York. 337 pp.
- WOELFLE-ERSKINE C, LARSEN LG and CARLSON SM (2017) Abiotic habitat thresholds for salmonid over-summer survival in intermittent streams. *Ecosphere* **8** (2) 1–23. <https://doi.org/10.1002/ecs2.1645>
- WARD RC and ROBINSON M (2000) *Principles of Hydrology* (4th edn.). McGraw-Hill, New York. 450 pp.
- YLLA I, SANPERA-CALBET I, VÁZQUEZ E, ROMANÍ AM, MUÑOZ I, BUTTURINI A and SABATER S (2010) Organic matter availability during pre- and post-drought periods in a Mediterranean stream. *Hydrobiologia* **657** (1) 217–232. <https://doi.org/10.1007/s10750-010-0193-z>
-

Methods for the estimation of extreme rainfall events

KA Johnson^{1,2*} and JC Smithers^{1,3,4,5}

¹Centre for Water Resources Research, University of KwaZulu-Natal, Pietermaritzburg, South Africa

²Civil Engineering, School of Engineering, University of KwaZulu-Natal, Durban, South Africa

³Bioresources Engineering, School of Engineering, University of KwaZulu-Natal, Pietermaritzburg, South Africa

⁴JG Afrika, Hilton, Pietermaritzburg, South Africa

⁵National Centre for Engineering in Agriculture, University of Southern Queensland, Toowoomba, Australia

ABSTRACT

The estimation of design rainfalls is necessary to estimate the exceedance probabilities of extreme floods required to design hydraulic structures and to quantify the risk of failure of these structures. New approaches to estimating extreme rainfall events are being developed internationally. This paper reviews methods for estimating design rainfalls, particularly extreme events, in South Africa and internationally, and highlights the need to update methods used for estimating extreme rainfall events in South Africa as a platform for future research.

Keywords: design rainfall, extreme events, estimation methods, probable maximum precipitation, South Africa

INTRODUCTION

The estimation of design rainfalls and design floods is required for the design of hydraulic structures such as bridges, culverts, dam spillways and urban stormwater drainage systems. Although establishing reliable estimates of design floods is challenging, they are necessary for flood risk management and to quantify the risk of failure of hydraulic structures (Kjeldsen et al., 2014; Madsen et al., 2009).

Design rainfall estimates are required as input for many methods used for design flood estimation, and are used in numerous engineering, as well as environmental and ecological conservation, design decisions, resulting in millions of rands of construction in South Africa each year (Schulze, 1984). Realistic and reliable estimates of design rainfalls, and hence design floods, are important for the preservation of human life and property (Pegram and Parak, 2004) and are required for adequate assessment of risk and economic impacts of failure, as the costs associated with repairs can be significant (Green et al., 2015).

According to Nathan and Weinmann (2013), design rainfall estimates associated with return periods ranging from 50 years to 2 000 years are considered to be in the large to rare range. Rare design rainfalls are essential for the design of high-hazard hydraulic structures as they provide more representative design floods for return periods between 100 years and probable maximum events (Green et al., 2016a).

In South Africa, numerous regional- and national-scale studies have focused on design rainfall estimations for daily and sub-daily durations. The most recent South African method for design rainfall estimation for short and long durations was developed by Smithers and Schulze (2003), which provides estimates for return periods ranging from 2 to 200 years. Since this study was completed, over a decade of additional rainfall data are available to update design rainfall estimates. Furthermore, Cullis et al. (2007) recommend that higher return period floods be used for the design of dams.

The Probable Maximum Precipitation (PMP) is an extreme rainfall quantity that is commonly used to derive the Probable Maximum Flood (PMF), which is used in the planning, design and risk assessment of high-hazard hydraulic structures (Chavan and Srinivas, 2015; Wang, 1984). Currently the South African National Committee on Large Dams (SANCOLD) Guidelines for the estimation of the PMF is dependent on the PMP (Smithers et al., 2014).

Numerous methods have been developed globally to estimate the PMP. In South Africa the PMP is determined using HRU report 1/72 (HRU, 1972), which was based on approximately 30 years of rainfall data from the 1930s to 1960s. The PMP is presented as an envelope of maximum observed storm values for meteorologically homogenous regions (Cullis et al., 2007). There has been no update in the development of PMP estimation since HRU (1972) and these guidelines are still used in professional practice today. The use of outdated PMP estimates potentially affects the accuracy of PMF estimations, and consequently the design of high-hazard hydraulic structures.

The urgent need to update the data and methods used for design flood estimation in South Africa has been highlighted by Smithers and Schulze (2003), Cullis et al. (2007), Görgens et al. (2007), Smithers (2012) and Van Vuuren et al. (2013). The need to update extreme design rainfalls is particularly highlighted by Smithers et al. (2014).

The objective of this paper is to present an overview of the existing methods and current practices of design rainfall estimation in South Africa and globally, with a particular focus on rare events, and to identify and discuss factors needed to improve the estimation of extreme rainfall quantities in South Africa as a platform for future research.

EXTREME RAINFALL ESTIMATION

This section introduces the concept of design rainfall estimation and Probable Maximum Precipitation, and briefly discusses the approaches for estimating these types of extreme rainfall events.

*Corresponding author, email: Johnsonk1@ukzn.ac.za

Received 27 November 2018; accepted in revised form 25 June 2019

Design rainfall estimation

For the planning and design of hydraulic structures such as water and flood control structures, design flood estimates are needed (Haddad et al., 2011a). As such, flood frequency analysis is important due to its environmental and economic impacts (Pilgrim and Cordery, 1993). Design rainfall is an essential input for design flood estimations used to quantify the risk of failure of hydraulic structures (Madsen et al., 2009; Mamoon et al., 2014). Although there is no universal method for design flood estimation, many countries have derived their own standard techniques and guidelines, which typically include a probabilistic approach involving statistical analysis of observed flood data and mathematical modelling using rainfall-runoff techniques (Smithers and Schulze, 2003). Often there are inadequate quantities of observed streamflow data at a site, leading to the frequent use of rainfall-runoff event-based methods, which require a probabilistic estimation of rainfall for a critical duration. This estimation is known as design rainfall and comprises a depth of rainfall over a period of time, or intensity, associated with a given likelihood of exceedance or return period (Mamoon and Rahman, 2014). Design rainfall Depth-Duration-Frequency (DDF) relationships are primary inputs to rainfall-runoff models to estimate design floods and are widely used by engineers and hydrologists for water resources planning and design applications (Haddad et al., 2011b). As such, DDF relationships are a key concept in the design of hydraulic structures as they quantify risk and maximise design efficiencies by using recorded events to predict future exceedance possibilities (Smithers and Schulze, 2003).

Rainfall frequency analysis – at-site vs regional approaches

Rainfall frequency analysis can be based on Annual Maximum Series (AMS) or Partial Duration Series (PDS) (Madsen et al., 1997), using either an at-site analysis or a regional analysis (Stedinger et al., 1993). The AMS is the largest rainfall depth in each year for a given duration whilst the PDS takes into account all events above a specified threshold (Madsen et al., 1997). An at-site approach to rainfall frequency analysis uses rainfall records from a single site of interest (Hosking and Wallis, 1997). Long periods of records are needed for reliable frequency analyses. However, challenges arise as often many rainfall stations have insufficient rainfall records and a high degree of variability in rainfall characteristics, making the spatial interpolation of design rainfall characteristics difficult (Haddad et al., 2011a). Stedinger et al. (1993) stated that it is necessary to utilise data from similar neighbouring regions to account for insufficient at-site data for frequency analysis. This concept is referred to as a regional frequency analysis and uses data from numerous locations to estimate the frequency distribution of observed data at each location (Hosking and Wallis, 1997). In this way, time-limited sampling records are supplemented by using data from a large number of rainfall stations in a region (Haddad et al., 2011a). In regional frequency analysis it is assumed that the data from all gauged sites within a homogenous region can be combined to produce a single regional rainfall frequency curve. After appropriate rescaling, this curve is applicable anywhere in the region, to both gauged and ungauged sites. Key elements of regional frequency analysis include: identification of homogenous regions, determination of regional frequency curves, and a method for estimation of the at-site mean (scaling factor) at any location in the region (Parrett, 1997).

A regional index-flood type approach to frequency analysis based on L-moments was developed by Hosking and Wallis (1997) and is termed the Regional L-Moment Algorithm (RLMA). L-moments are statistical quantities that are derived from Probability Weighted Moments (PWMs), which were defined by Greenwood et al. (1979). They summarise the theoretical probability distributions and observed samples and can be used for parameter estimation, interval estimation and hypothesis testing. Vogel and Fennessy (1993) concluded that for the purpose of applications in hydrology, L-moments are always preferred over product moments.

Probable Maximum Precipitation

For the design of high-hazard structures, such as dams located upstream of populated areas, structures are commonly designed for the theoretically largest possible flood that could occur at a particular location in order to maximise safety and reliability (Shaw, 1994). One of the key concepts in the assessment of dam safety under extreme flood conditions for high hazard dams is the Probable Maximum Flood (PMF), which is the theoretical maximum flood that is expected to occur at a particular location due to the most severe combination of critical meteorological and hydrological conditions (Wang, 1984).

Estimation of the PMF may be undertaken by empirical, probabilistic or deterministic methods. In the case of deterministic methods, the estimation of an extreme design rainfall is used to derive the PMF. This extreme design rainfall, to which no return period can be attached, is known as the PMP (Chavan and Srinivas, 2015). PMP is defined by the World Meteorological Organization (WMO, 2009b p. xxvi) as 'the greatest depth of precipitation for a given duration meteorologically possible for a design watershed or a given storm area at a particular location at a particular time of year, with no allowance made for long-term climatic trends.' Prior to the 1950s the concept of a potential upper limit of precipitation was referred to as Maximum Possible Precipitation (MPP). The idea that maximum limits exist for all rainfall-producing elements and that these limits could be quantified by studies of natural processes was considered. However, as nature is not constrained to limits, this was determined to be impossible as the upper limit of rainfall could not be determined accurately (WMO, 2009b).

PMP estimation using WMO method

The Manual on *Estimation of Probable Maximum Precipitation* was first published in 1973 by the WMO. Since then two revisions have been published, the latest in 2009 (WMO, 2009b). The manual provides basic methods for estimating PMP and conditions under which these methods may be applied, allowing professionals to apply the methods, or combinations of methods, to their specific design conditions. There are two general approaches to PMP estimation: direct and indirect. The former is based on catchment area whilst the latter is based on storm area (WMO, 2009a; 2009b).

Direct approach

In the direct approach the PMP is estimated directly for a given duration and is based on the specific design catchment area for a particular project. The main steps involved in this approach are: (a) the determination of a storm model, (b)

maximisation, and (c) derivation of the PMP. The storm model is a typical or ideal storm reflecting the characteristics of extreme precipitation over the design catchment. This extreme precipitation is likely to pose a serious threat of flooding in the project area. Maximisation involves maximising the storm performance. The PMP over a design catchment is then derived from the maximisation of the storm model. Thereafter the PMF can be calculated. WMO (2009a; 2009b) give details of the following methods for the determination of the storm model:

- (a) Local storm maximisation method
- (b) Storm transposition method
- (c) Combination method
- (d) Inferential method

Indirect approach

The indirect approach is based on the storm area, i.e., the area surrounded by isohyets. This approach is centred on the estimation of a group of PMPs, each varying in duration and area within a wide region. Thereafter the group of PMPs are converted into the PMP for the design catchment. WMO (2009b) highlights two commonly used methods based on storm area, viz., generalised estimation and statistical estimation.

Generalised estimation method

This method, referred to as the 'generalised storm maximisation and transposition method' in WMO (1984), is one of the most commonly used and accepted methods for PMP estimation. The United States National Weather Service (NWS) first introduced the concepts of storm maximisation and transposition in the late 1930s. During the early 1960s the NWS introduced the 'generalised' storm maximisation and transposition approach and later in 1973 the WMO documented these concepts in detail (Görgens et al., 2007). Generalised methods take into account site-specific meteorological conditions and, therefore, provide more reliable estimates when compared to statistical methods. However, significant time is required to perform generalised methods (WMO, 2009b).

The generalised estimation method is used for estimating the PMP for large meteorologically homogeneous regions. It is a physical approach that requires site-specific meteorological and geographical data. This method is applicable to specific catchments as well as large regions consisting of numerous catchments with a range of areas. It aims to generalise the areal mean precipitation depth within an isohyet. The observed rainfall of a storm is grouped into orographic and non-orographic rainfall. In orographic regions the precipitation occurs from the passing of weather systems in combination with orographic effects. For non-orographic zones it is assumed that precipitation occurs solely due to the passing of weather systems. Precipitation resulting from passing weather systems is known as convergence rainfall. The generalised estimation method involves the generalisation of convergence rainfall and utilises the depth-area-duration (DAD) generalisation of storms. In addition, this method involves the generalisation of the spatial/temporal distributions of PMP (WMO, 2009a).

Long periods of long-term rainfall data are needed for this method. Although this may be a time-consuming and expensive process, this method can lead to high accuracy and easy application of PMP results. PMP estimations can be

determined for durations of 6–72 hours for catchments under 13 000 km² in orographic regions and up to 52 000 km² in non-orographic regions (WMO, 2009b). The procedure for the generalised estimation method for non-orographic regions is as follows:

- (i) Identification of a high-efficiency storm: Selection of a major storm, based on observed data, with the assumption that their precipitation efficiency was near or at maximum.
- (ii) Moisture maximisation: The moisture factors of high-efficiency storms are adjusted to their maximum. The maximum persisting 12-hour–100-kPa dew point for each storm is used as an index of the maximum available quantity of atmospheric water vapour. Next the corresponding precipitable water content to the maximum dew point for each storm is determined. NWS (1980) defines the precipitable water content as the depth of water vapour condensed into liquid in a column of air of unit cross section. The maximum persisting 12-hour dew point within 15 days of the seasonal maximum, or for the actual storm date for the study region and its surroundings, is determined. Thereafter, the corresponding precipitable water content for the 12-hour maximum persisting dew-point is determined. The observed storm rainfall quantities are then multiplied by the moisture maximisation ratio, W_m/W_s where W_m represents the maximum precipitable water content (mm) and W_s is the precipitable water content estimated for the storm at dew point (mm). In some cases wind maximisation may be required. This is achieved by determining the maximum 24-hour average wind speed for each storm, which is multiplied by the precipitable water content to establish the maximum moisture inflow index.
- (iii) Transposition: In large regions there may be areas that have not experienced extreme storm events or lack recorded data of extreme events observed in adjacent regions. Transposition involves translating observed storm characteristics from a location where extreme storm events have been observed to a location with inadequate record of such major storms. The boundaries of the area of transposability of each storm are defined and then each storm is transposed within these boundaries. The boundaries or limits of transposability are influenced by climate and geographic features of the region, the frequency of major storm events and the concentration of rainfall gauges within the region.
- (iv) Enveloping: The maximised and transposed rainfall data from the largest values from the data set are plotted and used to identify a smooth curve of the rainfall data. Depth-area envelopes and depth-duration envelopes are plotted separately, and then used for the construction of DAD curves. For depth-duration envelopes, the largest maximised and transposed rainfall values are plotted for each duration and a smooth curve envelopes these values. In the case of depth-area envelopes, the maximum adjusted rainfall values for various areas and a specific duration are plotted. The depth-duration and depth-area enveloping can be done in one operation. The values plotted on the combination graph are taken from the enveloping of the depth-area and depth-duration curves, resulting in one value for each duration and area depicted on the graph. The PMP is then determined by the application of DAD enveloping values for the design catchment. PMF can then be determined based on the assumption that the flood created by the PMP is the possible maximum flood for the design area.

Statistical estimation method

Most PMP estimation procedures are based on meteorological analyses. However, sometimes site-specific meteorological data are not available. Statistical methods are useful in such cases (Fattahi et al., 2011). The most widely used statistical estimation method developed by Hershfield (1961; 1965), has been adopted by the WMO (2009b) as one of the standard methods for PMP estimation. This method takes into account actual historical data at a particular location and expresses it in terms of statistical parameters (Koutsoyiannis, 1999). The statistical estimation method is an approximate method applicable to small catchments with a collecting area of less than 1 000 km². The maximum value K_m in an observed storm series is statistically represented by the largest storm defined by Eq. 1:

$$K_m = \frac{X_m - \bar{X}_{n-1}}{S_{n-1}} \quad (1)$$

where

X_m = maximum observed storm value

\bar{X}_{n-1} = mean value excluding the maximum value

S_{n-1} = standard deviation excluding the maximum value

Hershfield (1965) developed a graphical relationship between enveloping values and means of annual series K_m for different durations. This was based on more than 2 600 rainfall stations, of which approximately 90% were based in the USA. The envelope curve varies with duration as shown in Fig. 1. Using Fig. 1 to determine K_m , the mean \bar{X}_n and the standard deviation S_n are determined using rainfall data from a specific station in the design catchment. Thus Eq. 2 is used to determine the PMP.

$$PMP = \bar{X}_n + K_m S_n \quad (2)$$

Furthermore, the coefficient of variation of the precipitation series for n years of data at a particular design station is given by Eq. 3:

$$C_{vn} = \frac{S_n}{\bar{X}_n} \quad (3)$$

The largest storm, K_m , is a statistical representation of the maximum observed storms. This value is then transposed and corrected to the design station by computing the mean \bar{X}_n and the coefficient of variation C_{vn} from the storm series in the design catchment. The PMP at any design station can be calculated using Eq. 4:

$$PMP = (1 + K_m C_{vn}) \bar{X}_n \quad (4)$$

As the frequency factors used in Hershfield's method were generated from observations mostly based in the US, many countries have modified the K_m factors to suit local conditions. Estimation of PMP by statistical methods has been successfully applied and documented by many countries, including India (Rakhecha and Clark, 2000), Malaysia (Desa et al., 2001), the Czech Republic (Rezacova et al., 2005) and Spain (Casas et al., 2011).

SOUTH AFRICA

Current practices of design rainfall and PMP estimation in South Africa are reviewed in this section. Furthermore, the

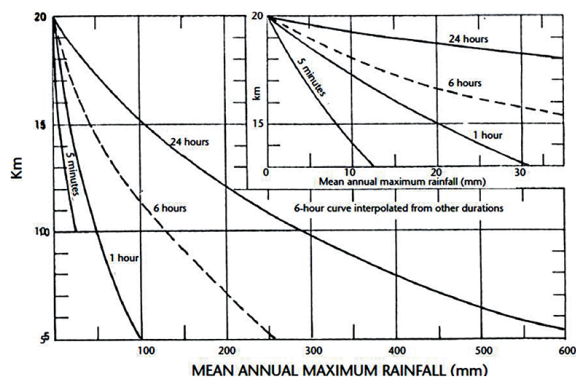


Figure 1. K_m as a function of rainfall duration and mean of annual series (WMO, 2009b)

need to update and modernise these methods and practices is highlighted.

Design rainfall estimation in South Africa

Numerous regional- and national-scale studies have focused on design rainfall estimation in South Africa, as cited by Smithers and Schulze (2003). Smithers and Schulze (2000a; 2000b; 2003) developed a regionalised index-storm approach based on L-moments for design rainfall estimation in South Africa. Smithers and Schulze (2000a) used digitised rainfall data from 172 recording rain gauges with at least 10 years of records for short duration (≤ 24 h) estimates. For longer durations (1–7 days), Smithers and Schulze (2000b) used daily rainfall data from 1 806 rainfall stations, all with at least 40 years of record. Prior to this, estimates for short-duration rainfall were last comprehensively produced in the 1980s and design rainfall depths for long durations were last estimated on a national scale by Adamson (1981).

Identification of homogenous regions involves the assignment of sites to regions whose frequency distributions are considered to be relatively homogenous after appropriate scaling. Grouped sites undergo standard multivariate statistical analysis based on the similarity of the vectors (Hosking and Wallis, 1997). Smithers and Schulze (2003) performed regionalisation of sites by cluster analysis, using appropriate site characteristics, viz., latitude, longitude, altitude, a monthly index of the concentration of precipitation (%), Mean Annual Precipitation (MAP) (mm), an index of rainfall seasonality and distance from sea (m). For short-duration rainfall, 15 relatively homogenous clusters were identified and for long-duration rainfall 78 clusters were identified. The index-storm value used was the mean of the AMS. It is necessary to estimate the mean of the AMS for a required duration at a particular location in order to estimate the design rainfall depths at an ungauged location. A cluster analysis of site characteristics was used to group the 78 long-duration clusters into 7 regions for estimating the mean of the 1-day AMS (Fig. 2). Multiple linear regression relationships with site characteristics such as MAP, latitude and altitude, allow for the estimation of the mean of the AMS at any location in the country. For each homogenous region the regression coefficients and variables differ (Smithers and Schulze, 2003).

Smithers (1996) and Smithers and Schulze (2000a; 2000b) determined the General Extreme Value (GEV) distribution to be the most appropriate distribution for extreme rainfall estimation in South Africa. Regional growth curves which

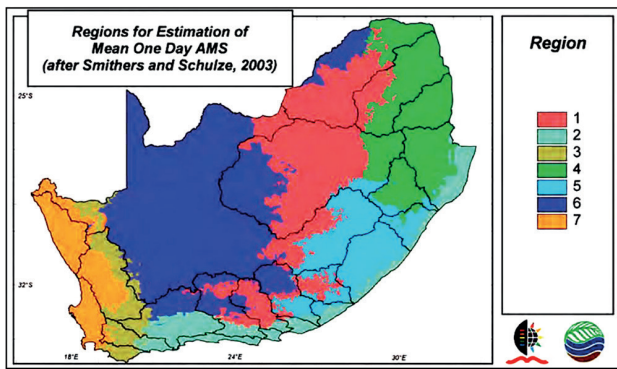


Figure 2. Relatively homogenous regions used for the estimation of the mean of the one day AMS for any location in South Africa (Knoesen et al., 2011)

relate design rainfall, scaled by the mean of the AMS, to duration were derived for each cluster for durations ranging from 15 minutes to 7 days. Inconsistencies in the growth curves for the 24-hour duration were evident due to the use of different databases for sub-daily and daily rainfall. To overcome these inconsistencies and to improve the reliability of estimates, scale invariance characteristics were investigated by Smithers and Schulze (2003) and it was concluded that the growth curves were scale invariant. The growth curve for the 1-day duration was established to be the most reliable and hence could be applied to durations ranging from 5 minutes to 7 days. The daily rainfall database is utilised to estimate the mean of the AMS for all durations (Smithers and Schulze, 2003). The application of the Regional L-Moment Algorithm together with a Scale Invariance approach is termed the RLMA&SI. Based on the RLMA&SI procedures, reliable and consistent design rainfall estimates for return periods ranging from 2 to 200 years and durations from 5 minutes to 7 days were produced for South Africa.

Need to update design rainfall values in South Africa

The Water Research Commission (WRC), together with SANCOLD, identified the urgent need to update the data and methods used for design flood estimation in South Africa. This consequently resulted in the initiation of a draft plan for the National Flood Studies Programme (NFSP) by Smithers et al. (2014), which highlights a wide range of issues for research. The Rainfall Analysis Working Group identified research needs in terms of design rainfall estimation, particularly extreme events. Since the study by Smithers and Schulze (2003), more than a decade of additional rainfall data are available. Van Vuuren et al. (2013) noted that extended record length may have significant impact on the estimation of design rainfall. As such, there is a need to update design rainfall estimates for all durations. Smithers and Schulze (2003) estimated design rainfall for return periods up to 200 years. Cullis et al. (2007) recommended that higher return period floods, i.e., $T > 200$ years, be used for the design of dams. Hence, there is a need for the estimation of design rainfall for $T > 200$ years in South Africa.

PMP estimation in South Africa

The PMP estimations in South Africa are currently determined using the only set of established guidelines reported by the Hydrological Research Unit (HRU) of the University

of Witwatersrand in 1972 (HRU, 1972; Van der Spuy and Rademeyer, 2014). The guidelines are based on approximately 30 years of rainfall data from 1932 to 1961. The guidelines were developed from a report published in HRU (1969), describing a detailed methodology for the determination of PMP in South Africa (Cullis et al., 2007). The estimations of PMP for large-area and small-area storms are determined separately. This is due to differences in internal mechanisms of the storms. In addition, there are differences in the spatial distribution of daily observed rain gauges required for large-area storms and the autographic rain gauges required for small-area storms. PMP estimation for large-area storms in South Africa is undertaken by the method of storm maximisation and transposition. For small-area storms, empirically derived curves generated from the highest recorded point precipitation, for a range of durations in various parts of the country, are used to estimate PMP (Görgens et al., 2007).

PMP estimation method for large-area storms

Large-area storms are linked to widespread rainfall over a long duration, and typically in catchment areas exceeding 5 000 km². For the analysis of large-area storms, South Africa was divided into 29 meteorologically similar sub-regions. The boundaries of the regions were defined according to orographic features and ranges of MAP. The PMP can be estimated for each region from maximised curves depicting PMP against area for various durations, starting from 1 day up to 6 days. HRU (1969) gives a detailed description of the steps involved in the derivation of the Depth-Area-Duration curves for PMP in South Africa:

- (a) Depth-Area-Duration analysis
 - (i) Storm selection: The Weather Bureau of South Africa provided rainfall records for gauging stations across the country, subdivided the country into regions and numbered the rain gauges according to these regions. One station from each region was selected based on rainfall record reliability and rainfall records of approximately 30 years, starting from 1932, were examined. A total of 170 storms were selected country-wide. The 12 highest rainfall events from each rainfall station were analysed to identify the dates and locations of the most severe storms.
 - (ii) Isohyetal patterns of storm rainfall: It was found to be impossible to draw isohyetal maps from individual storms with confidence due to the uneven distribution of rain gauge stations within the area of influence of a storm. Thus, an 'isopercental' procedure was implemented. This procedure is based on the concept that the areal distribution of rainfall depths during a storm is affected by topographic features of the study area, in the same manner as the MAP would be affected. Therefore, the MAP isohyetal pattern can be extended to indicate the spatial distribution of the storm. The determination of the storm isohyetal pattern by means of the general isopercental procedure involves the following steps:
 - Conversion of daily observed rainfall totals from each of the selected rain gauges into percentages of the average MAP from the 30 years of observed data
 - Plotting these percentages and determination of isopercental lines by means of interpolation
 - Determination of percentage values of rainfall for defined grid points and then conversion to equivalent percentage values using the estimates of MAP at the grid points
 - Plotting the observed storm rainfall and the calculated rainfall on a map
 - Drawing up isohyetal patterns

(iii) Conversion of isohyetal patterns to depth-area-duration graphs: A computer program was used to determine the depth-area relationship by means of numerical integration. The time distribution of the storm was determined by the derivation of the average mass curves at 1-day time intervals, and using the mass curves to proportion the depths of precipitation of the storm. This was done by first selecting the maximum 1-day precipitation, followed by the maximum 2-day precipitation, with this process being continued up to the total storm precipitation. In this way, the depth-area-duration curves were determined for each of the regions. For any given sub-region it is assumed that the DAD curve can be applied to any location within that particular region.

(b) Storm maximisation

The assumption that a high-efficiency precipitation would most likely have occurred during at least one of the analysed storms within each of the sub-regions is the basis for the approach adopted for PMP estimation in HRU (1969). As such, by maximising the prevailing moisture content for all the storms analysed, the PMP can be shown by an envelope of the maximised depth-area-curves. At the time HRU (1969) was written, the calculation of moisture content of the atmosphere was not possible due to limited availability of such measurements. Consequently, methods for computing the precipitable water content of a column of air from surface temperature and pressure readings were developed. Atmospheric moisture was computed for a range of surface temperatures and pressures and these variables were plotted on a graph. The storm-moisture-content and maximum-moisture-content ratio for each region for a time of the year which is relevant was adopted as a basis for maximisation. The maximum daily dew-point temperature was determined for multiple severe storms for each of the sub-regions and the computed atmospheric moisture graph was used to read the corresponding maximum storm moisture content. According to the storm-moisture-content and maximum-moisture-content ratio, the 1-day precipitation depths for each severe storm were maximised and the depth-area curves were adjusted upward. Furthermore, the 2-day to 6-day precipitation depths were maximised using Eq. 5:

$$X_i = PR_i \left(\frac{M_x}{M_c} \right) T_i \quad (5)$$

where

X_i = maximised precipitation depth

P = the 6-day precipitation depth

R_i = the ratio of the i -th precipitation depth to the 6-day precipitation depth

M_x = the maximum moisture content at the relevant time of year

M_c = the average 1-day maximum moisture content

T_i = the ratio of the i -th day to the 1-day moisture content

(c) PMP estimation

The maximised depth-area curves for each duration in each of the sub-regions were plotted on a single sheet and the upper envelope to all the depths was drawn. The envelope was assumed to represent the PMP for the entire sub-region. PMP-area-duration curves were generated for each of the sub-regions. HRU (1972) presents all the curves for each sub-region. An example of such a curve is shown in Fig. 3. It is estimated that these PMP-area curves may be subject to a 25% error. This is largely due to the inability to take into account possible

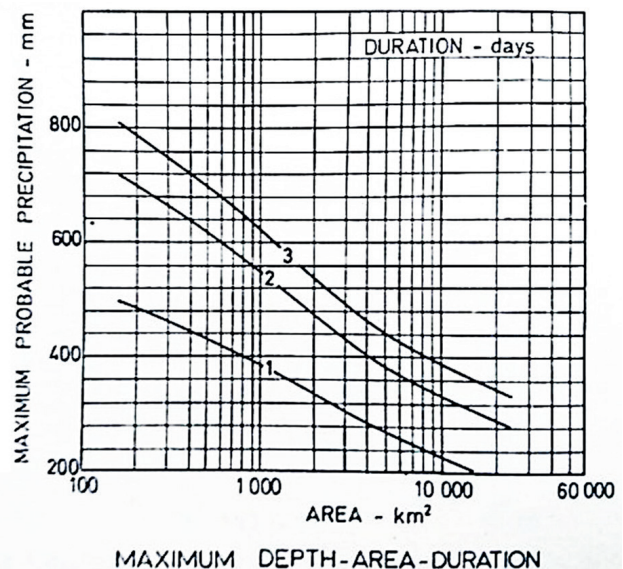


Figure 3. Example of PMP-area-duration curves (HRU, 1972)

atmospheric inversion, providing inaccurate estimation of moisture content (Görgens et al., 2007).

PMP estimation method for small-area storms

The estimation of PMP for small-area storms typically involves the analysis of point precipitation records and is associated with high-intensity, short duration, localised storms in areas of less than 15 km². Owing to the lack of sufficient meteorological data observed during short-duration storms, the short-duration point rainfall cannot be maximised in the same manner as for large-area storms. Instead, envelopes of the highest point rainfall of numerous durations observed throughout the country were used to develop an experience diagram. Once again, South Africa was divided into sub-regions. Each region has its own maximum rainfall envelope. This is shown in Fig. 4, together with envelopes of maximum rainfalls for the entire country and the world.

Performance of HRU PMP and recommendations

Since the development of the HRU PMPs, South Africa has experienced numerous severe flood events, and new and updated rainfall data are available (Lynch, 2004; Pegram et al., 2016; Van der Spuy and Rademeyer, 2014). Görgens et al. (2007) undertook to assess the applicability of the HRU PMP curves at that time. For the analysis of large-area storms, six severe storms were selected and a detailed site-specific analysis was completed for each storm. The analysis involved the generation of storm isohyets for critical durations. The average areal rainfall for specific storm durations was determined and compared to the applicable HRU PMP envelope curve of PMP versus area. Results of this analysis showed that the HRU PMP envelope curves were exceeded on numerous occasions, which suggests that these curves may be underestimating the maximum precipitation for large-area storms. The suggestion of underestimation is further supported by the fact that the storms utilised in the study were not maximised and transposed whilst the HRU PMP envelope curves were established based on storms which were both maximised and transposed. The maximum envelope curves for small-area

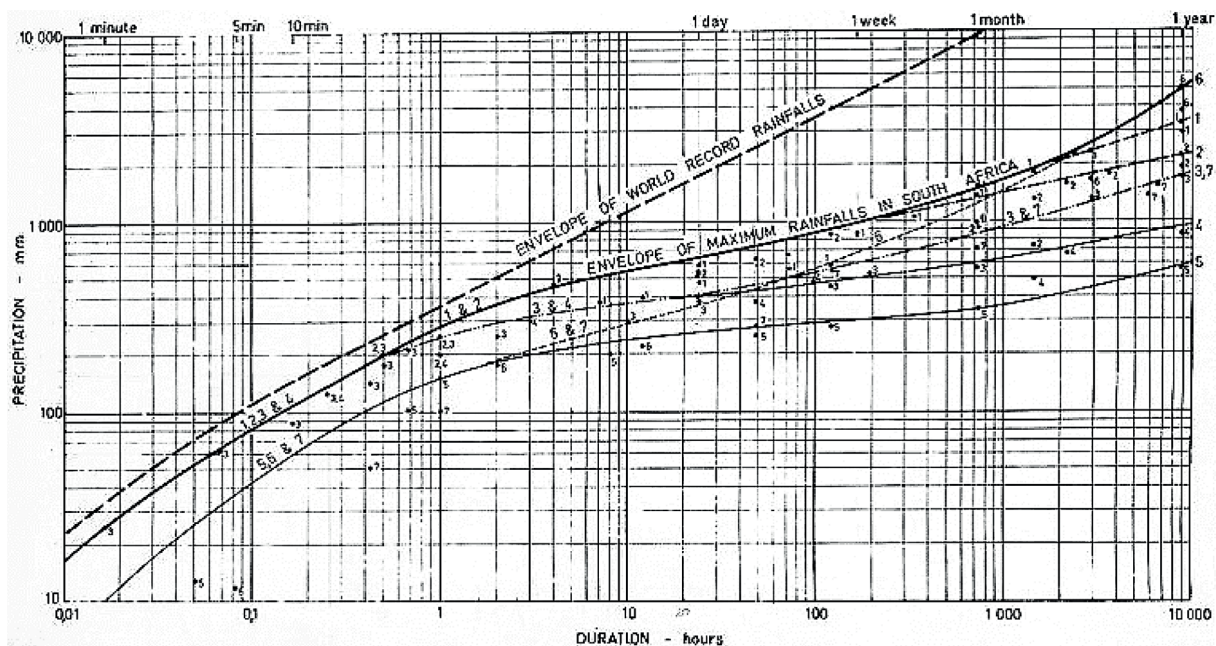


Figure 4. Maximum recorded point rainfalls in South Africa (HRU, 1972)

storms for the entire country were also exceeded on a number of occasions, once again highlighting that the HRU envelope curves may be underestimating extreme rainfall for short durations. Based on this study, Gørgens et al. (2007) suggested the HRU PMP curves may no longer represent the upper limit of design rainfall in many regions in South Africa, and that the PMF may be underestimated using these curves. Gørgens et al. (2007) recommend that the HRU PMP curves be modernised timeously to include longer and more current rainfall records, as well a more extensive rainfall gauge network. Additionally, research into addressing the inability to take into account possible atmospheric inversion, which affects the estimation of moisture content, was proposed.

AUSTRALIA

This section reviews Australia's methods for estimating design rainfalls for various probability ranges, as well as the various methods for estimating the PMP.

Design rainfalls in Australia

In Australia the guidelines for estimating design rainfalls are provided in Australian Rainfall and Runoff (ARR). The guidelines were first published in 1958 and updated thereafter in 1977 and 1978 (Pilgrim, 1984), and have been continually updated. Compared to the previous revision, the most recent

update makes use of nearly 30 years of additional daily rainfall data from 2 300 additional rainfall stations and was published in 2016 (Ball et al., 2016). Design rainfalls are categorised into five broad classes, based on the frequency of occurrence, as summarised in Table 1.

Different methods and data sets are used to estimate the various classes of design rainfalls. The traditional Intensity Frequency Duration (IDF) design rainfalls are the frequent and infrequent rainfalls in the probability range of 1EY to 1% AEP. The procedure for estimating these rainfalls involves assembling a quality-controlled rainfall database, selection and extraction of the extreme value series, frequency analysis using L-moments, regionalisation, and gridding (Green et al., 2016b).

For long-duration rainfalls, data from 8 074 daily-read rain gauges with at least 30 years of record were used. For short duration rainfalls 2 280 continuous rainfall gauges with more than 8 years of record were used. The AMS was selected to define extreme value series as large rainfalls are considered for estimating the IDF design rainfalls. For the at-site frequency analysis various distributions were trialled and it was determined that the GEV distribution was the best for the AMS fitted by L-moments (Green et al., 2016b)

For the regionalisation, Australia adopted the index-flood approach based on L-moments (Hosking and Wallis, 1997) in order to estimate the L-CV and L-skewness with greater confidence. The index value is the mean of the AMS (Green et al., 2016b). The homogenous regions can be defined using

Table 1. Classes of Design Rainfalls in Australia (Green et al., 2016b)

Design rainfall class	Frequency of occurrence	Probability range
Very Frequent Design Rainfalls	Very frequent	12EY* to 1 EY*
Intensity Frequency Duration (IFD)	Frequent	1 EY* to 10% AEP**
Intensity Frequency Duration (IFD)	Infrequent	10% to 1% AEP**
Rare Design Rainfalls	Rare	1 in 100 AEP** to 1 in 2000 AEP**
Probable Maximum Precipitation (PMP)	Extreme	< 1 in 2000 AEP**

*EY – exceedances per year

**AEP – annual exceedance probability

various methods. In areas where the station density is sparse a clustering or fixed region approach can be used, whereby stations are grouped into analysis areas with rigid boundaries based on spatial proximity. All stations in each analysis area were used to derive one regression equation that was then used for predictions within the region (Green et al., 2016b). Alternatively, when the station density is high a Region of Influence (ROI) approach (Burn, 1990) can be used, where an individual homogenous region is defined for each site. A ROI approach was used as the region sizes can be easily varied based on the station density and the available record length. A circular ROI and distances defined using latitude, longitude and elevation were used to define the membership of the ROI. The regionalisation was done using the 24-hour rainfall data and the same regions were used for long and short durations (Green et al., 2016b). In this way, the parameters of the GEV distribution (growth curve) were estimated at each location which was combined with the mean of the AMS (the index value) to estimate rainfall quantiles for any return period. In order to make the IDFs available for any point in the country, the ANUSPLIN software (Hutchinson, 2007) was used to grid the GEV parameters. Using this methodology, design rainfall estimates for return periods ranging from 1 to 100 years and durations ranging from 1 minute to 7 days are available for Australia (Green et al., 2016a).

For the very frequent design rainfalls, in the probability range of 1 to 12 exceedances per year, the overall approach was similar to that of IDFs, with the exception that the extreme value series selected was the PDS and the distribution chosen was the Generalised Pareto distribution (The et al., 2015).

Rare design rainfalls associated with return periods ranging from 1:100 to 1:2000 years are needed for the design of high-hazard hydraulic structures for which the impacts of potential failure could be significant. As these rainfalls are required for probabilities much rarer than the length of available records, they have to be extrapolated beyond recorded observed events. Daily rainfall stations with at least 60 years of record were used and the extreme value series selected was the AMS. The frequency analysis used the GEV distribution fitted to the AMS using LH-moments, developed by Wang (1997). LH-moments are a generalisation of L-moments which more accurately fit the upper tail of the distribution and increasingly focus on larger events in the data, depending on the degree of the shift, defined as η . When $\eta = 0$, LH-moments are simply L-moments. For Australia, LH-moments with a shift $\eta = 2$ were chosen. This is considered to be a good compromise between providing a better fit to the higher probabilities of the distribution and giving too much influence to the high outliers (Green et al., 2015). Many countries are exploring the use of LH-moments for estimating extreme rainfall and flood events (Ahmad et al., 2016; Deka et al., 2011; Hossein and Gheidari, 2013).

PMP estimation in Australia

The Australian Bureau of Meteorology has been continually developing procedures for estimating PMP since the first developments in the 1950s. Deterministic methods for PMP estimation have developed from an in-situ maximisation method, to a storm transposition method, and finally to the current generalised methods (Bureau of Meteorology, 1998; 2003).

The development of generalised methods began in the mid-1970s. Generalised methods utilise all the available data over a large area and include adjustments for moisture

availability and variable topographic effects on rainfall depth. Storm data undergo envelopment or smoothing over a range of durations and areas. The PMP estimates obtained via this method are generally higher than the other two methods. The generalised methods tend to set a more uniform standard. Currently, Australia has three generalised methods available: the Generalised Southeast Australian Method (GSAM) (Minty et al., 1996), the (Revised) Generalised Tropical Storm Method (GTSMR) (Walland et al., 2003) and the Generalised Short Duration Method (GSDM) procedures for small areas (Bureau of Meteorology, 2003).

UNITED KINGDOM

The methods used to estimate design rainfalls and the PMP in the United Kingdom are summarised in this section.

Design rainfalls in the United Kingdom

In 1975 the UK developed and published the Flood Studies Report (FSR) as the first guideline for flood estimation. For design rainfall estimation an index-storm approach was used with the index variable being the 5-year return period rainfall. The UK was split into two regions and regional growth curves were derived for 15 seconds to 25 days for each of the regions (NERC, 1975).

The FSR was replaced by the Flood Estimation Handbook (FEH) in 1999, which used longer rainfall records and contains updated methodology for rainfall and flood estimation in the UK (Stewart et al., 2014). This document was designed to estimate rainfalls up to return periods of 2 000 years and could be used to extrapolate rainfall estimates up to 10 000 years; however, this should be executed with caution. The Focused Rainfall Growth Extension (FORGEX) method is used to estimate rainfall frequency (Reed et al., 1999). It is an index-flood method where the index variable is the median annual maximum rainfall. The growth curves are derived using a complex empirical approach which combines a regional frequency analysis with an analysis of maximum points. The plotting positions of these points are shifted to accommodate spatial dependence of extremes.

PMP estimation in the UK

For PMP estimation using FSR methods, observed storms were examined and maximised across the UK for durations of 2 and 24 hours (Beran, 1987). Envelope growth factors for all durations were derived which allow for a quick estimation of PMP based on the 5-year return period for durations from 24 hours up to 25 days. Maximum rainfall depths for durations 2 and 24 hours can be determined by linear interpolation on a diagram of rainfall versus logarithm of duration. Rainfall depths for durations shorter than 2 hours and longer than 24 hours are estimated from factors related to average annual rainfall (Beran, 1987). The FEH does not make any revisions to the FSR PMP methodology and this remains the only available procedure for PMP in the UK (Babtie Group, 2000).

After brief use of the FEH design rainfalls in industry, concern was expressed by the dam profession in the UK regarding the results being obtained for high-return-period rainfalls. In a number of cases, the results of the 1 in 10 000 year rainfall depth exceeded the PMP assessed from the FSR (Babtie Group, 2000; MacDonald and Scott, 2001). Consequently, the UK Department for Environment, Food

and Rural Affairs (DEFRA) initiated an investigation into the irregularity and the appropriateness of the extrapolation methods in the FEH for the higher return period rainfall. The outcomes of the investigation led DEFRA to publish 'Revised Guidance to Panel Engineers' which states that the FEH should not be used to assess the 1 in 10 000 year return period rainfall; design rainfall values from FSR should be adopted whilst research continues. A new DDF model has been developed using a larger rainfall database. The FEH rainfall depths are in excess of the new DDF model, which is due to be released in the new FEH and made available through the web service from 2015 (Stewart et al., 2014).

USA

The USA's methods for estimating design rainfalls and the PMP are briefly discussed in this section.

Design rainfalls in USA

The National Oceanic and Atmospheric Administration's (NOAA) National Weather Service (NWS) provides the design rainfall estimates for the USA in the form of various atlases. Rainfall frequency estimations are based on the analysis of the AMS utilising an index-flood approach (Bonnin et al., 2011). Rainfall distributions at different sites are assumed to be the same with exception to a scaling factor, which is the index variable, the mean AMS rainfall. The method involves the estimation of the index for a particular site, and a standardised regional growth curve. The growth curve is the ratio of the quantiles to the index. Rainfall estimates are then obtainable by multiplying the site index value with the regional growth curve. The growth curve is estimated using L-moments following the procedure by Hosking and Wallis (1997). For shorter return periods an empirical formula is used to convert the AMS-based frequency estimates to Peak-Over-Threshold (POT) estimates.

Regionalisation was done by cluster analysis which takes into account topography, types of precipitation, and climatology in order to divide the area into fixed boundary regions. Various probability distributions were examined to determine the most suitable distribution for each region. Based on this approach, the rainfall frequency estimates have been published for durations ranging from 5 minutes to 60 days and for return periods of 1 to 1 000 years in NOAA Atlas 14 (Bonnin et al., 2011).

PMP in USA

The NWS also provides methods for determining the PMP for the USA in the form of various Hydrometeorological Reports (HMR). The first set of guidelines were developed in the late 1940s (Bonnin et al., 2011). The generalised method used involves the selection of maximum observed events, moisture maximisation, transposition, and envelopment. For the entire country over 500 storms have been analysed and the maximum observed aerial precipitation depths have been selected for various durations. The selected storm depths were then maximised based on the assumption that the extreme storms are sufficiently large that they have reached their maximum efficiency. The procedure is the same as that described in the WMO generalised method. Isohyetal patterns of storms are then relocated within meteorologically homogeneous regions. The isohyets were then smoothed to the data on several maps for durations of 6 to 72 hours

(NWS, 1980). Despite the fact that there have been many developments in PMP estimation since the development of the HMRS, there have been no updates to the methods used in the USA. Since 1999 the NWS no longer receives funding for research into PMP activities and as such developments into PMP estimation have ceased (Bonnin et al., 2011).

DEVELOPMENT OF OTHER APPROACHES FOR PMP ESTIMATION

Many other approaches to estimate PMP have been investigated. Douglas and Barros (2003) investigated an alternative approach for the estimation of PMP by means of multi-fractal analysis. The study involved the assessment of the value and utility of the application of multi-fractal analysis techniques to systematically compute estimates of extreme precipitation from observations in the eastern parts of USA. The multi-fractal approach provides a formal framework to estimate the extreme precipitation events empirically, referred to as Fractal Maximum Precipitation (FMP). Furthermore, this approach provides an objective estimate for the risk associated with the FMP. In the comparison of the multi-fractal estimates of the 1 in 1 000 000 year return period rainfall to the NWS PMP estimates, it was noted that the multi-fractal estimates were greater. Based on this it was concluded that the multi-fractal of extreme events should be taken as the upper bound of known risk to the standard NWS PMP (Douglas and Barros, 2003).

The storm model approach, developed by Collier and Hardaker (1996) in the UK, makes use of physical parameters, such as surface dew-point, height of storm cell, inflow and outflow, to represent the precipitation process. The maximised surface dew-point is used to determine the Maximum Precipitable Water content (MPW). The combination of the MPW with storm efficiency (E) produces the PMP for every duration t , as shown in Eq. 6. The storm efficiency is the ratio of the total rainfall at ground level to the total cloud condensed water:

$$PMP(t) = MPW \times E(t) \quad (6)$$

This method has been applied satisfactorily to produce PMP estimates in Barcelona, Spain, for durations ranging from 5 min to 30 hours. Similar results were obtained using the statistical approach (Casas et al., 2011).

UNCERTAINTY IN THE ESTIMATION OF EXTREME EVENTS

Montanari (2007) defines uncertainty in hydrology as a measure of the lack of accuracy regarding observed data and modelling outcomes. Data uncertainty results from measurement errors due to instrumental or human error. In cases where rainfall records are short, the use of a limited period of rainfall data introduces sample uncertainty. As such, estimates of higher order moments become more unstable due to the presence of outliers (Mamoon and Rahman, 2014). Tung and Wong (2014) explain that the sampling uncertainty is transferred to the DDF model, and eventually to the design rainfall estimation. Uncertainties related to regional frequency analysis include the degree of homogeneity assumed in a region, the record lengths of individual sites, the quantity of sites in a region, and the manner in which the data are pooled to compensate for the lack of data at other stations (Mamoon and Rahman, 2014). A changing climate has been noted to affect different aspects of the hydrological cycle, including rainfall and runoff. As there is a strong link between the

global climate system and the hydrological cycle, changes in any components of the climate system may result in potential changes to the magnitude and frequency of rainfall (Wang et al., 2013). Regional design rainfall estimates are based on recorded rainfall data. A non-stationary climate may modify regional rainfall characteristics, which may challenge the use of historical data for realistic long-term estimates (Mamoon and Rahman, 2014).

The representativeness of PMP estimations is dependent on the quantity and quality of data on extreme storm events and the depth of the analysis. Although methods for PMP estimation are designed to yield estimates to the nearest millimetre; this does not indicate the degree of accuracy. It must be noted that although storms and their associated floods have physical upper limits, due to limitations in data and hydrometeorological science and the physical complexity of the PMP, only approximations are available for these upper limits (WMO, 2009a). The lack of a standard approach for estimating the PMP as well as the perception that the PMP is an upper limit which cannot be exceeded has caused much criticism in industry as these values are used in the practice of designing and evaluating high-risk flood-related structures. The concept of a fixed upper limit with zero risk is unrealistic, as cases of recorded rainfalls exceeding the PMP have been documented globally (Salas et al., 2014). It is clearly impossible to produce values for PMP which are 100% representative. However, it is good practice to utilise a variety of methods simultaneously to estimate PMP and the results should be analysed, compared and synchronised from multiple perspectives to check consistency of the estimates and to select the best value (Koutsoyiannis, 1999).

DISCUSSION AND CONCLUSIONS

This review of literature indicates that it is clear that new and updated methods for estimating extreme design rainfall events are needed in South Africa. Design rainfall estimates are needed for the estimation of design floods which are not only used for the design of hydraulic structures, but to quantify the risk of failure of these structures as well. The potential loss of life as well as economic impacts associated with the failure of hydraulic structures can be significant. This highlights the importance of obtaining the best possible design rainfall and design flood estimates, which are dependent on the availability of reliable rainfall and flow records. At-site and regional design rainfall estimates are based on recorded rainfall data. However, a changing climate may modify local and regional rainfall characteristics, which may challenge the use of past data for realistic long-term estimates (Mamoon and Rahman, 2014). It is important that design rainfall estimates be updated periodically to include recent extreme events as they may have exceeded those previously recorded. Some countries are continually updating their design rainfalls to include updated rainfall databases with extended record lengths (e.g., Australia and UK). Typically a regional index-flood type approach to frequency analysis is used to estimate design rainfalls which involves the identification of homogenous regions, selection of a probability distribution function and determination of regional frequency curves, and a method for estimating a scaling factor.

The urgent need to update the data and methods used for design flood estimation in South Africa has been identified by SANCOLD and the WRC (Smithers et al., 2014). More than 10 years of additional rainfall data are available since design rainfalls were last estimated. Moreover, South Africa

has experienced changes in the occurrence of heavy rainfall events. This presents the opportunity to update design rainfall estimates. Unlike some countries (e.g. Australia, UK, and USA), South Africa does not provide design rainfall estimates for rare events, i.e., for return periods greater than 200 years. Australia has introduced new and different approaches for estimating rare design rainfalls. Cullis et al. (2007) recommend that higher return period floods be used for the design of dams, presenting a need to produce design rainfall estimates for return periods greater than 200 years. This review suggests that variations of L-moments which are better suited to account for larger events in the data be used to achieve this.

The PMP is an extreme rainfall quantity applied by hydrologists and engineers to determine the PMF, an extreme flood quantity used in the design of high-hazard hydraulic structures (Wang, 1984). WMO (2009a; 2009b) presents numerous methods which have been developed to estimate the PMP, which include deterministic (hydro-meteorological) and statistical methods. Although the concept of maximising and transposing extreme storms to estimate PMP is aligned with international practice, many countries have not revised their PMP estimates since their initial development, despite the introduction of the WMO guidelines. PMP estimates for the UK and USA were last estimated in the late-1900s. Consequently, these PMP estimates do not take into account recent extreme events.

In South Africa the only established guideline for PMP estimation is given in HRU (1972), in which envelope curves for regions which experience similar extreme rainfalls were developed. Although the HRU approach is practical and conservative, it is based on only 30 years of data from the 1930s to the 1960s. Over 5 decades of additional rainfall data are now available. Since these guidelines were first published, the country has experienced several significant rainfall events, some of which resulted in serious damage and loss of life (Cullis et al., 2007). Investigations by Gørgens et al. (2007) suggest the HRU PMP curves may no longer represent the upper limits of design rainfall estimations everywhere in South Africa and, consequently, the PMF may be underestimated in places. Underestimating PMP estimates and relying on outdated estimates can have severe effects on flood estimation and flood risk management. As such, the current South African PMP estimates are in need of urgent revision (Cullis et al., 2007). The latest modernised methods prescribed by the WMO which take into account factors previously not included in South Africa's PMP estimates, such as atmospheric inversion, should be used to develop updated PMP envelopes for the country.

REFERENCES

- ADAMSON PT (1981) Southern African storm rainfall. Technical Report TR 102. Department of Environmental Affairs, Pretoria.
- AHMAD I, ABBAS A, SAGHIR A and FAWAD M (2016) Finding probability distributions for annual daily maximum rainfall in Pakistan using linear moments and variants. *Pol. J. Environ. Stud.* 25 (3) 925–937. <https://doi.org/10.15244/pjoes/61715>.
- BABTIE GROUP (2000) Reservoir safety – Floods and reservoir safety: Clarification on the use of FEH and FSR design rainfalls. Final report to the Department of the Environment, Transport and the Regions (DETR), Glasgow, UK.
- BALL J, BABISTER M, NATHAN R, WEEKS W, WEINMANN E and TESTONI I (2016) *Australian Rainfall and Runoff: A Guide to Flood Estimation*. Commonwealth of Australia, Australia.
- BERAN M (1987) The UK Flood Studies Report: Continuing responsibilities and research needs. In: Singh VP (ed) *Flood Hydrology*. Springer, Dordrecht. 27–39. https://doi.org/10.1007/978-94-007-4000-0_2.

- [org/10.1007/978-94-009-3957-8_3](https://doi.org/10.1007/978-94-009-3957-8_3).
- BONNIN G, MARTIN D, LIN B, PARZYBOK T, YEKTA M and RILEY D (2011) NOAA Atlas 14, Precipitation-Frequency Atlas of the United States, Volume 1. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland, USA.
- BUREAU OF METEOROLOGY (1998) Generalised Probable Maximum Precipitation Estimates for the Katherine River Catchment to Katherine Town. GPMP/16. Bureau of Meteorology, Canberra.
- BUREAU OF METEOROLOGY (2003) The Estimation of Probable Maximum Precipitation in Australia: Generalised Short-Duration Method. Bureau of Meteorology, Canberra.
- BURN DH (1990) An appraisal of the "region of influence" approach to flood frequency analysis *Hydrol. Sci. J.* **35** (2) 149–165. <https://doi.org/10.1080/02626669009492415>.
- CASAS MC, RODRIGUEZ R, PROHOM M, GAZQUEZ A and REDANOC A (2011) Estimation of the probable maximum precipitation in Barcelona (Spain). *Int. J. Climatol.* **31** (9) 1322–1327. <https://doi.org/10.1002/joc.2149>.
- CHAVAN RS and SRINIVAS VV (2015) Probable maximum precipitation estimation for catchments in Mahanadi River Basin. *Aquat. Proced.* **4** 892–899. <https://doi.org/10.1016/j.aqpro.2015.02.112>.
- COLLIER CG and HARDAKER PJ (1996) Estimating probable maximum precipitation using storm model approach. *J. Hydrol.* **183** (1996) 277–306. [https://doi.org/10.1016/0022-1694\(95\)02953-2](https://doi.org/10.1016/0022-1694(95)02953-2).
- CULLIS J, GÖRGENS AHM and LYONS S (2007) Review of the selection of acceptable flood capacity for dams in South Africa in the context of dam safety. WRC Report No. 1420/1/07. Water Research Commission, Pretoria.
- DEKA S, BORAH M and KAKATY SC (2011) Statistical analysis of annual maximum rainfall in North-East India: an application of LH-moments. *Theor. Appl. Climatol.* **104** (2011) 111–122. <https://doi.org/10.1007/s00704-010-0330-7>.
- DESA M, NORIAH A and RAKHECHA P (2001) Probable maximum precipitation for 24 h duration over southeast Asian monsoon region—Selangor, Malaysia. *Atmos. Res.* **58** (1) 41–54. [https://doi.org/10.1016/S0169-8095\(01\)00070-9](https://doi.org/10.1016/S0169-8095(01)00070-9).
- DOUGLAS EM and BARROS AP (2003) Probable maximum precipitation estimation using multifractals: application in the eastern United States. *J. Hydrometeorol.* **4** (2003) 1012–1024. [https://doi.org/10.1175/1525-7541\(2003\)004<1012:PMPEUM>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1012:PMPEUM>2.0.CO;2).
- FATTAHI E, NOORIAN AM and NOOHI K (2011) Comparison of physical and statistical methods for estimating probable maximum precipitation in southwestern basins of Iran. *Desert* **15** (2010) 127–132.
- GÖRGENS A, LYONS S, HAYES L, MAKHABANE M and MALULEKE D (2007) Modernised South African design flood practice in the context of dam safety. WRC Report No. 1420/2/07. Water Research Commission, Pretoria.
- GREEN J, BEESLEY C, FROST A, PODGER S and THE C (2015) National estimates of rare design rainfall. (December 2015). Paper presented at: *Hydrology and Water Resources Symposium*, December 2015, Hobart, Tasmania, Australia.
- GREEN J, BEESLEY C, THE C, PODGER S and FROST A (2016a) Comparing CRCFORGE estimates and the new rare design rainfalls. Paper presented at: *ANCOLD Conference*, October 2016, Adelaide, Australia.
- GREEN J, JOHNSON F, BEESLEY C and THE C (2016b) Book 2 Rainfall Estimation: Chapter 3 – Design Rainfall. In: Ball J, M Babister, R Nathan, W Weeks, E Weinmann, M Retallick and I Testonis (eds) *Australian Rainfall and Runoff: A Guide to Flood Estimation*. Commonwealth of Australia, Australia.
- GREENWOOD JA, LANDWEHR JM, MATALAS NC and WALLIS JR (1979) Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form. *Water Resour. Res.* **15** (5) 1049–1054. <https://doi.org/10.1029/WR015i005p01049>.
- HADDAD K, RAHMAN A and GREEN J (2011a) Design rainfall estimation in Australia: a case study using L-moments and Generalized Least Squares Regression. *Stoch. Env. Res. Risk A.* **25** (6) 815–825. <https://doi.org/10.1007/s00477-010-0443-7>.
- HADDAD K, RAHMAN A, GREEN J and KUCZERA G (2011b) Design rainfall estimation for short storm durations using L-moments and generalised least squares regression-application to Australian data. *Int. J. Water Resour. Arid Environ.* **1** (3) 210–218.
- HERSHFIELD D (1965) Method for estimating probable maximum precipitation. *J. AWWA* **57** 965–972. <https://doi.org/10.1002/j.1551-8833.1965.tb01486.x>.
- HOSKING JRM and WALLIS JR (1997) *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press, Cambridge, UK.
- HOSSEIN M and GHEIDARI N (2013) Comparisons of the L- and LH-moments in the selection of the best distribution for regional flood frequency analysis in Lake Urmia Basin. *Civ. Eng. Environ. Syst.* **30** (1) 72–84. <https://doi.org/10.1080/10286608.2012.749870>.
- HRU (1969) Design storm determination in South Africa. 1/69. Hydrological Research Unit, University of the Witwatersrand, Johannesburg.
- HRU (1972) Design flood determination in South Africa. 1/72. Hydrological Research Unit, University of the Witwatersrand, Johannesburg.
- HUTCHINSON M (2007) ANUSPLIN version 4.37 User Guide. The Australian National University, Centre for Resources and Environmental Studies Canberra, Australia.
- KJELDSEN T, MACDONALD N, LANG M, MEDIERO L, ALBUQUERQUE T, BOGDANOWICZ E, BRÁZDIL R, CASTELLARIN A, DAVID V, FLEIG A and co-authors (2014) Documentary evidence of past floods in Europe and their utility in flood frequency estimation. *J. Hydrol.* **517** 963–973. <https://doi.org/10.1016/j.jhydrol.2014.06.038>.
- KNOESEN D, SCHULZE R and SMITHERS J (2011) Climate change and short duration day design rainfall: a 2011 perspective. Chapter 7.1. In: Schulze RS (ed) *A 2011 Perspective on Climate Change and the South African Water Sector*. WRC Report No. TT 518/12. Water Research Commission, Pretoria. 199–205.
- KOUTSOYIANNIS D (1999) A probabilistic view of Hershfield's method for estimating probable maximum precipitation. *Water Resour. Res.* **35** (4) 1313–1322. <https://doi.org/10.1029/1999WR900002>.
- LYNCH S (2004) Development of a raster database of annual, monthly and daily rainfall for southern Africa. WRC Report No. 1156/1/04. Water Research Commission, Pretoria.
- MACDONALD D and SCOTT C (2001) FEH vs FSR rainfall estimates: An explanation for the discrepancies identified for very rare events. *Dams Reservoirs* **11** 280–283.
- MADSEN H, ARNBJERG-NEILSEN K and MIKKELSEN PS (2009) Update of regional intensity-duration-frequency curves in Denmark: Tendency towards increased storm intensities. *Atmos. Res.* **92** 343–349. <https://doi.org/10.1016/j.atmosres.2009.01.013>.
- MADSEN H, RASMUSSEN PF and ROSBJERG D (1997) Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. *Water Resour. Res.* **33** (4) 747–757. <https://doi.org/10.1029/96WR03848>.
- MAMOON AA, JOERGENSEN NE, RAHMAN A and QASEM H (2014) Derivation of new design rainfall in Qatar using L-moment based index frequency approach. *Int. J. Sustainable Built Environ.* **3** 111–118. <https://doi.org/10.1016/j.ijbsbe.2014.07.001>.
- MAMOON AA and RAHMAN A (2014) Uncertainty in design rainfall estimation: A review. *J. Hydrol. Environ. Res.* **2** (1) 65–75.
- MINTY LJ, MEIGHEN J and KENNEDY MR (1996) Development of the Generalised Southeast Australia Method for estimating probable maximum precipitation. Hydrology Report Series, Report No. 4. Bureau of Meteorology, Melbourne, Australia.
- MONTANARI A (2007) What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. *Hydrol. Process.* **21** 841–845. <https://doi.org/10.1002/hyp.6623>.
- NATHAN R and WEINMANN E (2013) Australian rainfall and runoff. Discussion paper: Monte-Carlo simulation techniques. AR&R D2, Engineers Australia: Water Engineering, Barton, Australia.
- NERC (1975) *Flood Studies Report*. National Environment Research,

- London.
- NWS (National Weather Service) (1980) Probable maximum precipitation estimates – United States east of the 105th meridian. Hydrometeorological Report No. 51. National Weather Service, National Oceanic and Atmospheric Administration, Washington, DC, USA.
- PARRETT C (1997) Regional analysis of annual precipitation maxima in Montana. 97-4004. United States Geological Survey, Montana, USA.
- PEGRAM G and PARAK M (2004) A review of the regional maximum flood and rational formula using geomorphological information and observed floods. *Water SA* **30** (3) 377–392. <https://doi.org/10.4314/wsa.v30i3.5087>.
- PEGRAM G, SINCLAIR S and BÁRDOSSY A (2016) New methods of infilling southern African raingauge records enhanced by annual, monthly and daily precipitation estimates tagged with uncertainty. WRC Report No. 2241/1/15. Water Research Commission, Pretoria.
- PILGRIM D (1984) *Australian Rainfall and Runoff: A Guide to Flood Estimation*. Institution of Engineers Barton, ACT, Australia
- PILGRIM D and CORDERY I (1993) Chapter 9: Flood Runoff. In: Maidment DS (ed) *Handbook of Hydrology*. McGraw-Hill, New York.
- RAKHECHA P and CLARK C (2000) Point and areal PMP estimates for durations of two and three days in India. *Meteorol. Appl.* **7** (1) 19–26. <https://doi.org/10.1017/S1350482700001389>.
- REED D, FAULKNER D and STEWART E (1999) The FORGEX method of rainfall growth estimation – II: Description. *Hydrol. Earth Syst. Sci.* **3** 197–203. <https://doi.org/10.5194/hess-3-197-1999>.
- REZACOVA D, PESICE P and SOKOL Z (2005) An estimation of the probable maximum precipitation for river basins in the Czech Republic. *Atmos. Res.* **77** (1) 407–421. <https://doi.org/10.1016/j.atmosres.2004.10.011>.
- SALAS JD, GAVILAN G, SALAS FR, JULIEN PY and ABDULLAH J (2014) Uncertainty of the PMP and PMF. In In: Eslamian S (ed.) *Handbook of Engineering Hydrology*. Taylor & Francis Group, LLC, USA. 575–603.
- SHAW EM (1994) *Hydrology in Practice* (3rd edn.). Taylor & Francis e-Library, UK.
- SMITHERS JC (1996) Short-duration rainfall frequency model selection in Southern Africa. *Water SA* **22** (3) 211–217.
- SMITHERS JC (2012) Review of methods for design flood estimation in South Africa. *Water SA* **38** (4) 633–646. <https://doi.org/10.4314/wsa.v38i4.19>.
- SMITHERS JC, GÖRGENS A, GERICKE J, JONKER V and ROBERTS CPR (2014) The initiation of a national flood studies programme for South Africa. South African Committee on Large Dams (SANCOLD), Pretoria, South Africa.
- SMITHERS JC and SCHULZE RE (2000a) Development and evaluation of techniques for estimating short duration design rainfall in South Africa. WRC Report No. 681/1/00. Water Research Commission, Pretoria.
- SMITHERS JC and SCHULZE RE (2000b) Long duration design rainfall estimates for South Africa. WRC Report No. 811/1/00. Water Research Commission, Pretoria.
- SMITHERS JC and SCHULZE RE (2003) Design rainfall and flood estimation in South Africa. WRC Report No. 1060/1/03. Water Research Commission, Pretoria.
- STEDINGER J, VOGEL R and FOUFOULA-GEORGIOU E (1993) Chapter 18: Frequency analysis of extreme events. In: Maidment DS (ed) *Handbook of Hydrology*. McGraw-Hill, New York.
- STEWART L, VESUVIANO G, MORRIS D and PROSDOCIMI I (2014) The new FEH rainfall depth-duration-frequency model: results, comparisons and implications. Paper presented at: 12th British Hydrological Society National Symposium, 2–4 Sept 2014, Birmingham, UK.
- THE C, BEESLEY C, PODGER S, GREEN J, HUTCHINSON M and JOLLY C (2015) Very frequent design rainfalls – an enhancement to the new IFDs. Paper presented at: *Hydrology and Water Resources Symposium*, December 2015, Hobart, Tasmania, Australia.
- TUNG Y and WONG C (2014) Assessment of design rainfall uncertainty for hydrologic engineering applications in Hong Kong. *Stoch. Env. Res. Risk A.* **28** 583–592. <https://doi.org/10.1007/s00477-013-0774-2>.
- VAN DER SPUIY D and RADEMEYER P (2014) Flood frequency estimation methods as applied in the Department of Water Affairs, Pretoria, RSA. Department of Water Affairs, Pretoria.
- VAN VUUREN S, VAN DIJK M and COETZEE G (2013) Status review and requirements of overhauling flood determination methods in South Africa. WRC Report No. TT 563/13. Water Research Commission, Pretoria.
- VOGEL RM and FENNESSY NM (1993) L-Moments diagrams should replace product moment diagrams. *Water Resour. Res.* **29** (6) 1746–1752. <https://doi.org/10.1029/93WR00341>.
- WALLAND DJ, MEIGHEN J, XUEREB KC, BEESLEY CA and T HTM (2003) Revision of the Generalised Tropical Storm Method for Estimating Probable Maximum Precipitation. Hydrology Report Series, Report No. 8. Bureau of Meteorology, Melbourne, Australia.
- WANG B-H (1984) Estimation of probable maximum precipitation: case studies. *J. Hydraul. Eng.* **110** (10) 1457–1472. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1984\)110:10\(1457\)](https://doi.org/10.1061/(ASCE)0733-9429(1984)110:10(1457)).
- WANG D, HAGEN SC and ALIZAD K (2013) Climate change impact and uncertainty analysis of extreme rainfall events in the Apalachicola River basin, Florida. *J. Hydrol.* **480** 125–135. <https://doi.org/10.1016/j.jhydrol.2012.12.015>.
- WANG QJ (1997) LH moments for statistical analysis of extreme events. *Water Resour. Res.* **33** (12) 2841–2848. <https://doi.org/10.1029/97WR02134>.
- WMO (World Meteorological Organisation) (1984) Manual for Estimation of Probable Maximum Precipitation. Second edition. Operational Hydrology Report No.1. WMO Report No. 332. World Meteorological Organisation, Geneva.
- WMO (World Meteorological Organisation) (2009a) Guide to Hydrological Practices Volume II Management of Water Resources and Application of Hydrological Practices, Sixth edition. WMO Report No. 168. World Meteorological Organisation, Geneva.
- WMO (World Meteorological Organisation) (2009b) A manual on estimation of probable maximum precipitation (PMP). WMO Report No. 1045. World Meteorological Organisation, Geneva.

Climate change and the water footprint of wheat production in Zimbabwe

Simbarashe Govere^{1,*}, Justice Nyamangara¹ and Emerson Z Nyakatawa²

¹Department of Environmental Sciences & Technology, Chinhoyi University of Technology, P. Bag 7724, Chinhoyi, Zimbabwe

²Department of Crop Sciences & Technology, Chinhoyi University of Technology, P. Bag 7724, Chinhoyi, Zimbabwe

ABSTRACT

Reductions in the water footprint (WF) of crop production, that is, increasing crop water productivity (CWP), is touted as a universal panacea to meet future food demands in the context of global water scarcity. However, efforts to reduce the WF of crop production may be curtailed by the effects of climate change. This study reviewed the impacts of climate change on the WF of wheat production in Zimbabwe with the aim of identifying research gaps. Results of the review revealed limited local studies on the impacts of climate change on the WF of wheat production within Zimbabwe. Despite this, relevant global and regional studies suggest that climate change will likely result in a higher WF in Zimbabwe as well as at the global and regional level. These impacts will be due to reductions in wheat yields and increases in crop water requirements due to high temperatures, despite the CO₂ fertilization effect. The implications of a higher WF of wheat production under future climate change scenarios in Zimbabwe may not be sustainable given the semi-arid status of the country. The study reviewed crop-level climate change adaptation strategies that might be implemented to lower the WF of wheat production in Zimbabwe.

Keywords: water footprints, climate change, wheat, yield, crop water requirements

INTRODUCTION

Agro-ecosystems are the largest users of water accounting for 70% of global withdrawals and 90% of the global water consumption (Shiklomanov and Rodda, 2003; Haddeland et al., 2011).

Accelerated population growth, change in diets and demand for green fuels means that global freshwater demand in agro-systems will increase to cater for the rising need for food, fibre and biofuels (Falkenmark et al., 2008; Gleick, 2003). However, most countries of the world are located in already water-scarce basins with 2–3 billion people living in highly water-stressed areas (Oki and Kanae 2006; Kummu 2014). A possible solution to close the gap between agricultural water demand and availability might be to increase the crop water productivity (CWP), that is reduce the water footprint (WF) of crop production in agro-ecosystems (Hoestra and Mekonnen, 2012).

Efforts to decrease the WF in crop production of agro-ecosystems may be hampered by the possible implications of climate change and variability. In the last century global surface temperatures have risen by an average of 0.07°C whilst the global atmospheric CO₂ concentration is rising at an annual rate of 2 ppm (NASA, 2016). Since crop yields and evapotranspiration, and thus WF, are determined to a large extent by climatic conditions, future changes in climate are likely to affect the WF of crop production. The various non-linear ways in which climatic factors can affect crop production via geographic and crop-specific factors mean that the precise impact of climate change on the WF of crop production for many countries is not known. There is therefore a need for the assessment of the impacts of climate change on crop production WF at the national level (Sun et al., 2012).

Zimbabwe is already experiencing climate change and variability (GoZ, 2015). Climate change is anticipated to affect

the production of staple crops such as wheat (Chawarika, 2016; Manyeruke, 2013). An initial process in the assessment of climate change impacts is a comprehensive review of past studies to ascertain the current state of knowledge. To the researchers' knowledge, there is no comprehensive review of studies assessing impacts of climate change and variability on the WF of wheat production in Zimbabwe.

This paper provides a literature review on the potential of climate change and variability to impact the WF of wheat production in Zimbabwe with the broad aim of identifying research gaps and needs. It firstly provides a background to wheat production and climate change in Zimbabwe and then critically assesses the results, methods and models of local research on the effect of climate change and variability on crop water consumption of wheat production. Search results from local studies were then compared with regional and global results. In response to the results of the literature review, adaptation strategies to help combat the impact of climate change on the local WF of wheat production in Zimbabwe are suggested.

DATA AND METHODS

Theoretical framework

In agro-systems the classical method used to measure a crop's capacity to convert water into marketable yield is the crop water productivity (CWP). For cereals the CWP can be defined as the ratio of the harvested grain yield to the volume or depth of water applied in irrigation or lost in evapotranspiration (Morisson et al., 2008; Tambussi et al., 2007). It is measured empirically by the formula:

$$CWP = Y_a / Et_a \quad (\text{kg/m}^3) \quad (1)$$

where Y_a is the actual marketable crop yield (kg/ha) and Et_a is the actual seasonal crop water consumption by evapotranspiration (m³/ha).

*Corresponding author, email: simgovere2000@yahoo.com

Received 31 July 2018; accepted in revised form 13 June 2019

Another relatively new and innovative way of measuring a crop's capacity to convert water into marketable yield is the WF. The WF of crop production is defined as the freshwater volume (in cubic meters per ton of crop) required for crop growth and dilution of pollutants during the production process of crops (Hoekstra et al., 2009). Analogous to the ecological footprint and carbon footprint the WF is becoming increasingly popular in research and policy because of its universal application to non-agricultural sectors such as manufacturing as well as its critical linkages to human activities such as pollution and trade (Zhang et al., 2014). In some research circles the WF is also referred to as the virtual water content (VWC).

The WF has 3 components: green, blue and grey. The blue water footprint is the volume of freshwater that evaporated from blue water resources (surface water and ground water) to produce the crop. The green water footprint is the volume of water evaporated from green water resources (rainwater stored in the soil as soil moisture). The grey water footprint is the volume of polluted water that is associated with the production of the crop.

The blue WF can be calculated numerically as:

$$WF = Et_a / Y_a \quad (m^3/kg) \quad (2)$$

This can also be written as:

$$WF = 1/(Y_a/Et_a) = 1/(CWP) \quad (m^3/kg) \quad (3)$$

Thus the blue WF is the reciprocal of the WP (Amarasinghe and Smakhtin, 2014).

The two dependent components of the WF of crop production are crop yield and crop evapotranspiration, both of which are affected by the main climatic parameters which are temperature, rainfall and atmospheric CO₂ levels. Climate change and variability is thus a significant issue that can potentially affect WF in crop production. The overall effect of climate change on yields and evapotranspiration can either be positive or negative depending on geography, the particular crop and the degree of climate change. High temperatures in mid- and low-altitudes are anticipated to elevate crop evapotranspiration and reduce crop yields for C4 crops (e.g. maize). whilst in high latitudes they may increase C3 crop (e.g. wheat) yields resulting in lower WFs of production (Gornall et al., 2010; Adams et al., 1999). However, for C3 crops like wheat, increases in atmospheric CO₂ levels can increase crop yields, the so called CO₂ fertilization effect, whilst simultaneously reducing crop transpiration resulting in a net reduction in the WF of crop production (Cartwright, 2013; Degener, 2017).

Database search

The literature review was conducted by carrying out a literature search of peer-reviewed articles and grey literature published from 1985 to 20 February 2018. Many climate change studies on crops have focused on either crop yield or crop evapotranspiration, separately. For this reason, this study de-segregated its analysis of climate change impacts on WF by focusing separately on wheat yields, crop evapotranspiration and the actual WF. For peer-reviewed articles the databases EBSCO, PubMed, Web of Science, BioOne and Scopus were used. Grey literature searches were conducted using Eldis, Google Scholar and AGRIS (the Food and Agriculture Organization of the United Nations and the International Food Policy Research Institute) and UNESDOC (the UNESCO

database). Since most climate impact studies present crop water use as either crop evapotranspiration, crop water requirements, irrigation requirements or their variants, these terms were included in the literature review. The search terms used in the search were 'Zimbabwe', 'Africa' or 'global' and 'climate change' and 'wheat', 'temperate cereals' or 'cereals' followed by 'yields,' 'water footprints', 'virtual water contents', 'water productivity,' 'water use efficiency', 'irrigation requirements' or 'crop evapotranspiration'.

Only studies written in English were used due to limitations on resources. To be included, a study had to meet the following criteria:

- Any original peer-reviewed research, review paper or white/grey document that contained results on climate change impacts pertaining to Zimbabwe, Southern Africa or Africa
- The results of climate change impacts had to be specific to wheat, temperate cereals (barley and rye) or cereals in general
- Assess the effects of increasing CO₂ and temperature on the water footprints, virtual water contents, water productivity, water use efficiency, irrigation requirements, crop evapotranspiration or water use of wheat, temperate cereals (barley and rye) or cereals in general
- Provided quantitative or qualitative data on changes in yields of wheat, temperate cereals or cereals due to the effects of climate change

A schematic representation of the screening process is given in Fig. 1. The initial search produced 840 papers which were then subjected to filtering. The first filtering was based on the source title; a second filter was then applied based on the source abstract. Full documents (peer review articles, industry reports) were only reviewed after satisfying all inclusion criteria. Ultimately 34 articles were selected and analysed to provide 50 'observations' on wheat yield, water use and water footprints in Zimbabwe and the region.

Study area

Zimbabwe is a semi-arid country located in Southern Africa between latitudes 15° and 23°S, and longitudes 25° and 34°E. Zimbabwe has a tropical climate which is moderated in many places by the effect of local topography. As a result, wheat, a temperate crop, in Zimbabwe is mainly grown in the midveld and highveld areas (altitude above 600 m) where the cool winter season (0–20 °C) has proved ideal for wheat production (Havazvidi, 2006).

Significant wheat production also occurs in the lowveld (altitude below 600 m) in areas under government parastatals such as the Chisumbanje and Save-Valley estates.

These combined areas coincide with agro-ecological regions IIA, IIB and III which account for 75–80% of the area planted to crops in Zimbabwe (FAO, 2000). Wheat is the second-most strategic food security crop in Zimbabwe after maize, accounting for more than 50% of daily caloric intake in the country (Kapuya, 2010).

Wheat is mainly used to make flour and bran. Flour is used to make bread, which has become a staple food in the country, while bran is processed into a stock feed (Mutambara et al., 2013). Since there is very little or no rainfall during the winter months (May to August), irrigation is required to achieve a good wheat crop. Wheat is therefore a fully-irrigated winter crop in Zimbabwe. The irrigation requirements for wheat and its associated energy and water development costs mean that wheat production in Zimbabwe is mainly a commercial enterprise. The heavy reliance of wheat production

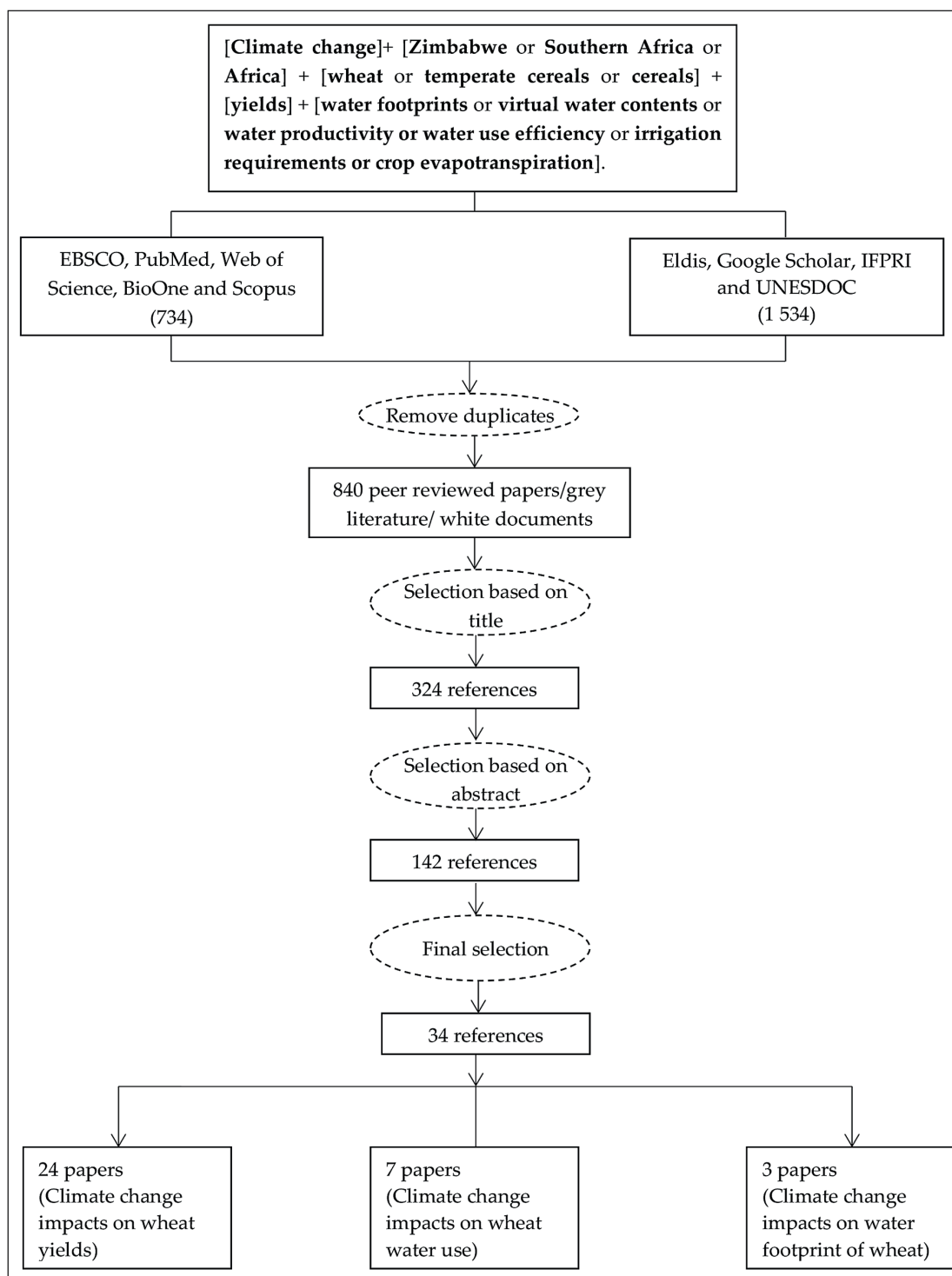


Figure 1. Schematic representation of the review process

on irrigation in the predominantly semi-arid climate of Zimbabwe, which is experiencing the effects of climate change, implies that the efficient utilization of water resources is a key issue. Zimbabwe adopted the concept of integrated water resources management (IWRM) at the 2nd World Water Forum, held in The Hague in March 2000 (Swatuk, 2005). One of the key Dublin policy principles that enshrine IWRM stipulates that water resources must be used in an efficient manner in all human endeavours.

RESULTS AND DISCUSSION

Climate change and variability in Zimbabwe

Compelling evidence exists which shows that climate change is already occurring in Zimbabwe. During the 20th century the annual mean temperatures over Zimbabwe have significantly increased. Unganai (1996) reports the mean centennial rise in temperature at 0.8°C. Rekecwicz (2005)

posits a conservative increase of 0.4°C but notes that there has been an increase in both the minimum and maximum temperatures over Zimbabwe represented by a decrease in the number of days with a minimum temperature of 12°C and a maximum of 30°C. Brown et al. (2012) estimate the overall rise in the daily minimum and maximum temperatures to be 2.6°C and 2°C, respectively.

Most researchers concur that the annual mean rainfall over Zimbabwe has declined (Makarau, 1995; Unganai, 1996; Rekacewicz, 2005; Chamaille-Jammes et al., 2007). Unganai (1996) used linear regression and noted a 10% decline in annual mean rainfall over the country in the past century. Makarau (1995) used quadratic and exponential analysis and noted a reduction of approximately 100 mm in the mean annual rainfall from 1901 to 1994. Rekacewicz (2005) concluded that the mean annual rainfall received during a rainy season has decreased by about 5% since 1900 and rainfall patterns have shifted; more rainfall is occurring at the beginning of the season, in October, and less rain is being received between January and March. Mazvimavi (2010), however, used the Mann-Kendal test and concluded that due to the high rainfall inter-annual variability over Zimbabwe it is meaningless at the moment to associate any rainfall change with global warming.

With regards to climate projections, the annual mean temperature over Zimbabwe is anticipated to increase by about 3°C for the 2050s, compared to the 1961–1990 normal, using various global circulation models (GCMs) (Hulme et al., 2001; Christensen et al., 2007). Similar results were obtained at a catchment level when Mujere and Mazvimavi (2012) projected a 3°C maximum temperature increase for Mazoe catchment for the 2050s. Seasons will likely change with hotter dry seasons and colder winters. Simulations using GCMs anticipate 5–18% less mean annual rainfall by the year 2080 compared to the 1961–1990 normal (Hulme et al., 2001; Christensen et al., 2007). The traditional onset and cessation of rainfall seasons will shift with fears of shorter and more erratic rainfall seasons. The reduction in precipitation means that in the long term yields from reservoirs will be reduced and there will be less water available for allocation across all sectors. Climate change is anticipated to cause a streamflow decline of up to 50% for the Gwayi, Odzi and Sebakwe catchments in Zimbabwe (Mazvimavi, 1998).

Growing research is showing that climate change at decadal timescales is closely linked with the increased frequency of the El Niño–Southern Oscillation (ENSO) phenomenon (Fedorov and Philander, 2000; Zhang et al., 2008; Cob et al., 2013; Cai, 2015; Wang et al., 2017). The frequency and strength of El Niño have been more variable during the 20th century than the preceding 7 000 years (Cob et al., 2013). In Zimbabwe 62% of El Niño occurrences are associated with below normal rains and droughts which result in reduced water availability in surface and groundwater sources (Gopo and Nangombe, 2016).

Zimbabwe is one of the many sub-Saharan African nations which are extremely vulnerable to climate change due to a combination of factors that include endemic poverty and constrained coping mechanisms (Chagutah, 2010; Madzwamuse, 2010). Agriculture is the mainstay of the economy, contributing an average of 16.76% (7.41–22.89%) of GDP and providing employment to 60–70% of the population (Malaba, 2013). Cross-national poverty profiles show that poverty is endemic in the country, with more than 70% of the population classified as poor and 84% of these living in

rural areas (Sakuhuni et al., 2012; Malaba, 2013; Manjengwa et al., 2012). Without adaptation strategies the impacts of climate change may be potentially severe for the country due to its heavy dependence on agriculture and lack of financial resources for mitigation and adaptation to climate change. Climate change adaptation strategies, especially in the agricultural sector, are therefore a principal development challenge in Zimbabwe.

Climate change impacts on wheat productivity

There is a general consensus among several studies, using different models and approaches, that projected climate change will negatively affect wheat yields for Africa and the world at large (Liu et al., 2008; Ringer et al., 2010; Nelson et al., 2009; Zhao et al., 2017; Challinor et al., 2014; Wheeler et al., 2012; Asseng et al., 2014; Matiu et al., 2017). This review found 24 regional and global studies that predicted declines in wheat yields due to climate change and variability (Table 1). Of the 24 studies reviewed no local studies were found suggesting that very little research has been carried out in Zimbabwe on climate impacts on wheat yield. Most local climate change impact studies have focused on maize due to its importance as the prime staple of the country (Makadho, 1996; Mano and Nhemachena, 2007; Muchena, 1994; Ciarns et al., 2016; Rurinda et al., 2015; Lebel et al., 2015; Makuvaro, 2014).

Nonetheless it was still possible to isolate climate change impacts on Zimbabwean wheat yields from global and regional studies.

Parry et al. (2004) estimated the future global wheat yields at a national level using yield transfer functions and concluded that wheat yields in Africa may decrease up to 30% in the 2080s. For Zimbabwe the study determined wheat reductions of from 10–30% for the 2080s under A1FI, A2, B1 and B2 scenarios, but using only one GCM, the HadCM3 model.

Konar et al. (2016) used climate outputs from 14 GCMs run under the high carbon emission SRES A2 scenario to force a process-based crop model, Global Hydrology Model (H08). Their global model predicted a decline in wheat yields in Zimbabwe of 26–63% for the 2030s. In another global study, Deryng et al. (2014) used the global crop yield model PEGASUS, driven by projected climate data from the MAGICC 6 GCM forced under all four RCPs emission scenarios. For Zimbabwe the study by Deryng et al. (2014) determined a 50% reduction in wheat yields.

With respect to climate variability studies, Zampieri et al. (2017), Iizumi and Ramankutty (2016) and Ray et al. (2015) concluded that year-to-year variability in climate has resulted in significant variations in wheat yields. At a global level as much as 75% of the wheat yield variability can be explained by climate variability whilst for Zimbabwe it was between 30 and 45%. Using statistical analysis the researchers correlate global historical wheat yields with historical climate data to detect patterns in yield changes. An advantage of statistical analysis is that it is based on observational historical data from individual farms or regions which implicitly takes into account farmer management behaviour as opposed to field experiments.

A statistical description of the results of the reviewed studies showed that there is wide variation in wheat yield reduction depending on spatial extent and time period (Fig. 2). However, there is general agreement that, regardless of spatial extent and time period, wheat yields will be negatively affected by climate change. The anticipated declines in wheat yields suggest that negative climate change impacts, like heat

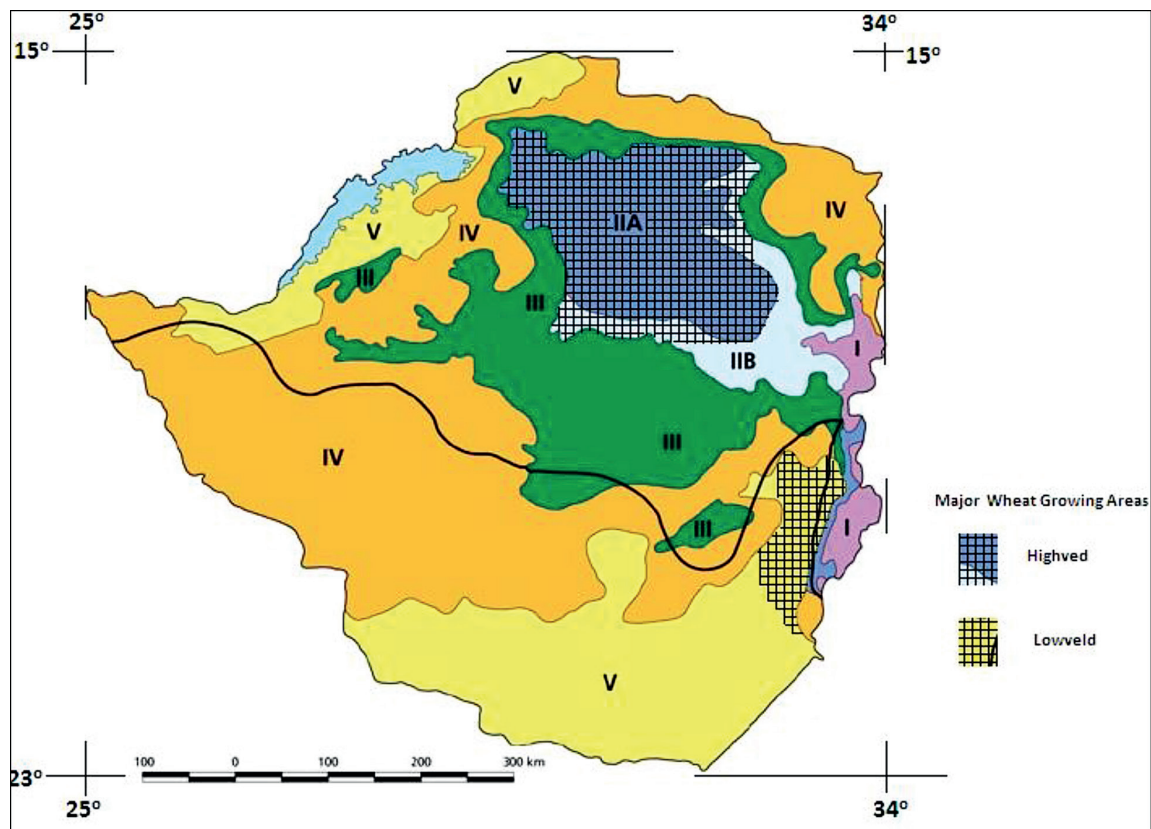


Figure 2. Agro-ecological zones and the major wheat-growing areas in Zimbabwe: After Morris (1988)

stress due to temperature rise, counter the beneficial effects of CO₂ fertilization (Siebert and Ewert, 2014; Matiu et al., 2017; Zampieri et al., 2017). For instance, in a meta-analysis Challinor et al. (2014) found that yield decline for a unit increase in temperature (°C) was 4.90% whilst yield increment for a unit increase in atmospheric CO₂ (ppm) was 0.06%, suggesting that at a global level heat stress overrides the CO₂ fertilization effect. The review shows that anticipated impacts of climate change might be higher at smaller spatial scales (national level) but lower at larger spatial scales (global and continental level). The average decline in wheat yields per time period in 46%, 22% and 8% for Zimbabwe, Africa and the world, respectively.

The implications of the decline in wheat yields are significant for Zimbabwe. Traditionally, Zimbabwe has had some of the highest national average wheat yields, of between 5–6 t/ha compared to the current global average of 2.5–3 t/ha (Bhasera and Soko, 2017). This has partly been due to the impact of wheat breeding research in Zimbabwe over the period 1969 to 1991 as well as introduction of agronomic practices which have resulted in a potential maximum wheat yield of 10 t/ha (Mashingwani and Mutisi, 1994). Wheat yield reductions of 46% would shift the national average wheat yields to 3.15–3.78 t/ha. Economic analysis using the current gazetted wheat prices for Zimbabwe in 2017 shows that the break-even yield that results in a gross profit of zero is 4 t/ha (Bhasera and Soko, 2017).

The literature review revealed that climate impact studies use a diverse variety of GCMs, emission scenarios and crop models. For this reason a time-series plot of wheat yield shocks produced no meaningful trend since diverse emission scenarios and GCMs can be used by various studies for the

same time-slice, resulting in varying results (Fig. 3). Earlier studies used a limited number of GCMs in their simulations (e.g. Parry et al., 2004 used only the HaDCM3 GCM); however, more recent studies are incorporating GCM ensembles in their analysis (e.g. Deryng et al., 2014).

The majority of the studies used process-based crop growth simulation models for assessing the impacts of climate change on wheat productivity. Common models used in the studies include the Decision Support System for Agro-technology Transfer (DSSAT), APSIM (Agricultural Production System Simulator), the GIS coupled Erosion Productivity Impact Calculator (GEPIC), the Agro-Ecological Zone (AEZ) model, WOFOST (World Food Studies) and the FAO's Cropwat model. Process-based crop models are based on crop physiological responses to environmental factors which is a key strength if external validation of the model to the environment is done (Roberts et al., 2017). A down-side of the models is that validation is based on experimental field plots but not on real farmer-managed fields where pest, weed and disease control strategies, fertilizer applications and other management practices significantly vary, depending on farmer behavior. Perhaps more importantly, process-based models were developed for finer spatial scales with homogeneous environmental conditions and their use over large spatial scales with multiple heterogeneities in environmental conditions might lead to errors (Abraha and Savage, 2006; Schulze and Walker, 2006).

In an effort to increase precision in modeling approaches a significant number of researchers are adopting a multimethod approach (Asseng et al., 2014; Zhao et al., 2017). A multimethod approach incorporates process-based crop models, statistical modeling as well as Free-Air

Table 1. Reviewed literature on climate change impacts on wheat yields

Sources	Methods	Region	Emission scenario	CO ₂ effect	Time period	Percentage per +1°C	Change per period
Asseng et al., 2014	Multimethod ensemble	Global	-	Included	2020s, 2050s	-6%	
Zhao et al., 2017	Multimethod ensemble	Global	RCP2.6, RCP4.5, RCP6, RCP8		1961–1990 2071–2100		-13.6%
IPCC, 2014	Meta-analysis	Global	-	-	1960–2013		-8%
Challinor et al., 2014	Meta-analysis	Global	-	-	-		
Moore et al., 2017	Meta-analysis	Global	-	Included	1995–2005	-4.90%	
Knox et al., 2012	Meta-analysis	Africa	-	-	-	-5%	
Lobell et al., 2008	Process-based	Southern Africa	A2, B1, A1b	Not included	2020–2040		-17%
Liu et al., 2013	Process-based	Africa	A1F1, A2, B1, B2	Included	2020–2040		-15%
Fischer et al., 2005	Process-based	Zimbabwe		Included	2020s, 2050s, 2080s		-17%
		Zimbabwe		Included			-50–75%
		Zimbabwe		Included			-25–50%
		Africa		Included			-12%
Wiebe et al., 2015	Process-based	Africa	RCP4.5, RCP6 RCP8	Included	2050		-10%
Muller et al., 2010	Process-based	Zimbabwe		Included	2046–2055		0–5%
Konar et al., 2016	Process-based	Global			2000–2030		-6.5%
Parry et al., 2004	Process-based	Zimbabwe					-26–63%
		Global	A1F1, A2, B1, B2	Included	2020s, 2050s, 2080s		-5%
		Zimbabwe		Included			-30–10%
Tatsumi et al., 2011	Process-based	Southern Africa	A1B	Included	2090		-36.83
		Global					+7%
Rosenzweig and Parry, 1994	Process-based	Zimbabwe		Included	2060		-10–30%
Nelson et al., 2014	Process-based	Global	550 ppm CO ₂	Included	2050s		-2.3–25%
Ringler et al., 2010	Process-based	Africa		Included	2000–2050		-22%
Nelson et al., 2009	Process-based	Africa		Included	2050		-28%
Deryng et al., 2014	Process-based	Global	RCP 2.6/8.5	Included	2050–2080		+sp
		Zimbabwe					-50%
lizumi and Ramankutty, 2016	Statistical	Global	-	-	1981–2010		±19%–33%
Lobell et al., 2011	Statistical	Global	-	Indirectly included	1980–2008		-5.5%
Ray et al., 2015	Statistical	Global	-		1961–2000		±75
		Zimbabwe					±30–45%
Matiu et al., 2017	Statistical	Global	-	-	1961–2014		-9.2%
Zampieri et al., 2017	Statistical	Global			1980–2010		±40%

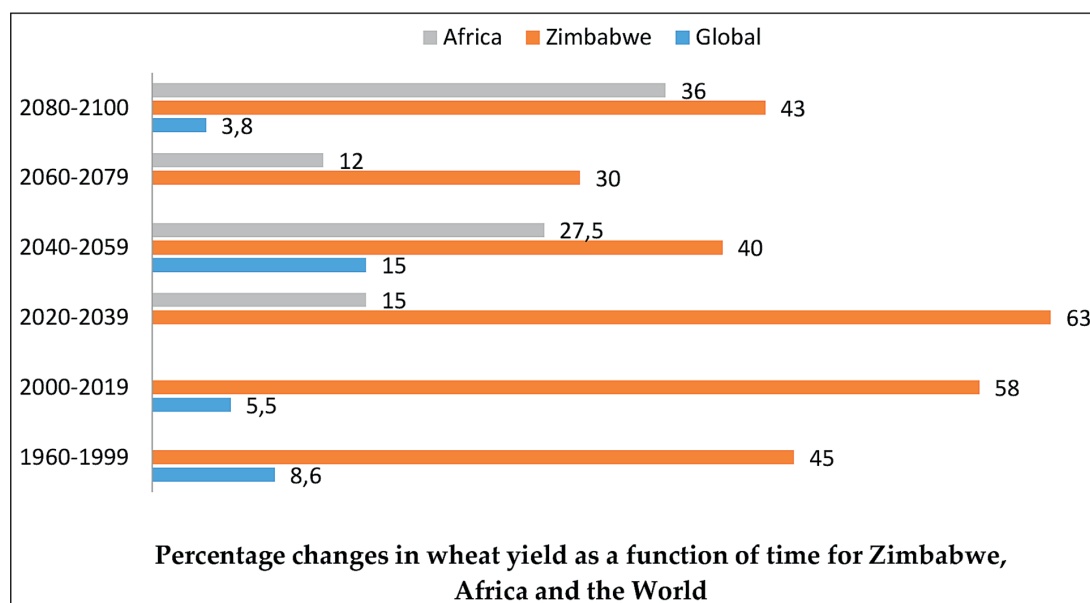


Figure 3. Wheat yield reductions per period from reviewed studies

Carbon-dioxide Enrichment (FACE) experiments where crops are grown in CO₂-rich environments to mimic the effect of climate change. However multimethod approaches may simply result in great precision in climate studies but not accuracy if the delimitations for each separate method are not resolved. For instance, Jones et al. (2014) have shown that though FACE experiments provide precise information about crop physiological and phenological responses to enriched CO₂, most of the experiments are spatially biased since they have been carried out in the temperate regions of Europe and America. Extrapolation of experimental results from these regions to the tropical Asian and African regions might result in inaccuracy since tropical areas have different biomes and environmental conditions.

Climate change impact on wheat water use

A literature search on the impact of climate change on water requirements of wheat in Zimbabwe yielded little or no local results. The search did, however, result in 7 global and regional studies that have attempted to quantify climate impacts on wheat water use. Of the 7 studies, 2 studies had results for Zimbabwe, and 2 of the 7 studies specifically focused on wheat whilst the remaining 5 collectively focused on cereals, temperate cereals or all crops (wheat included). Compared to the 24 papers reviewed under wheat productivity the relatively small number of papers on wheat water use may signal that food security issues have a higher research priority than water security issues. However, water security and food issues are intricately connected (Brazilian et al., 2011).

Global studies suggest that high temperatures lead to an increased irrigation water demand by increasing the overall crop transpiration rate. Using two GCMs (Hadley and CSIRO) under A2r scenarios Fischer et al. (2007) projected a 20% increase in net irrigation requirements for the world by 2080. They noted that about 65% of the global net irrigation requirement increases would emanate from higher crop water demands under the changed climate, and the remaining 35% from extended crop calendars. For Africa net irrigation

requirements for crops were expected to increase by 14%. The significantly lowered CO₂ concentrations may contribute to lower crop water demand.

Doll and Siebert (2002) used the GIM (Global Irrigation Model) to determine how irrigation requirements might change under the climatic conditions of the 2020s and the 2070s using two climate models, ECHAM4 and HADCM3. Their simulation gave contradictory results for southern Africa based on the climate models. The ECHAM4 model showed a consistent decline in irrigation requirements across the region for the 2020s and the 2070s whilst HADCM3 showed increments for both scenarios. The study shows that climate change impact predictions vary significantly, depending on the crop models used, climate change scenario and the number and types of global circulation model (GCM) used.

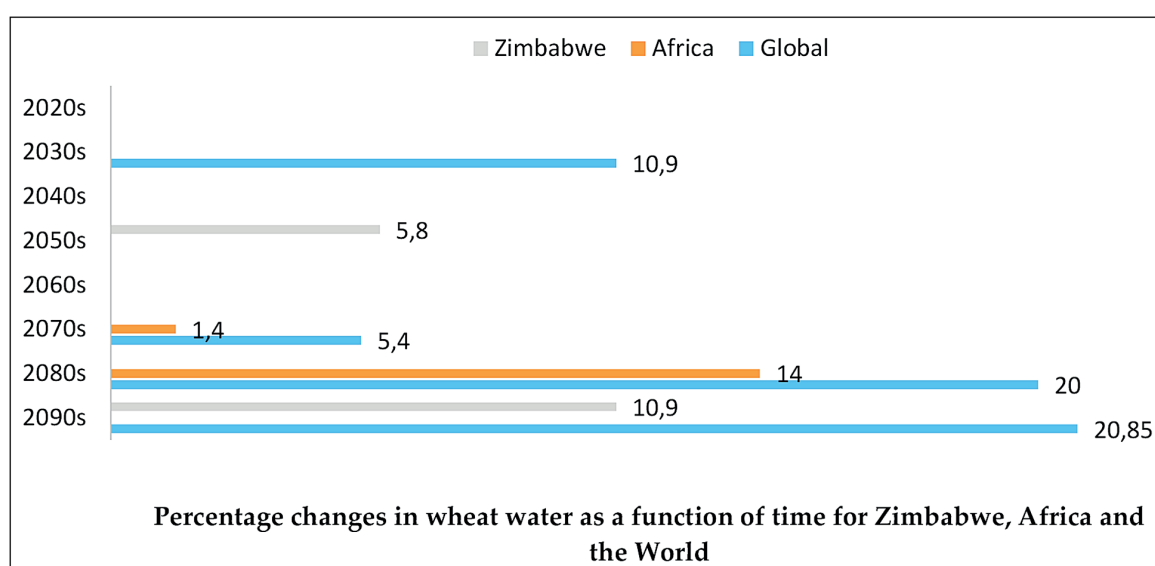
Pan et al. (2015) determined that global crop lands would experience an increase of about 38.9% and 14.5%, respectively, under both the A2 and B1 scenarios between the 2090s and the 2000s. Their analysis suggested that climate variability accounted for 91.3% of the inter-annual variation in evapotranspiration. Whilst the study pointed to increased evapotranspiration rates, it also showed that strength of the CO₂ fertilization effect would determine the magnitude of global terrestrial evapotranspiration during the 21st century. The CO₂ fertilization can result in reduced evapotranspiration through reduction in stomata conductance in plants.

Two global studies noted a decline in global irrigation requirements. Liu et al. (2013) noted that globally the net irrigation requirements for cereal crops (wheat, maize and rice) would decrease in the 2030s and 2090s. However net irrigation requirements would increase in southern Africa. For Zimbabwe, the study determined that crop water use would both increase (12.5% to 11.4%) and decrease (-45.7%–25.8%) for the 2030s and 2090s, respectively.

Zhang and Cai (2013) used 5 GCMs and noted that global irrigation requirements for major crops might decline slightly despite the anticipated rise in temperature. In their analysis wheat irrigation requirements for Zimbabwe and Africa as a whole decreased. This counter-intuitive effect noted by Zhang

Table 2. Percentage change in crop water use from reviewed papers and publications

Sources	Region	Crop	Methods	Emission scenario	CO ₂ effect	Time period	Water use percentage	Changes depth
Fischer et al., 2007	Global Africa	All crops	Process-based	o, A2r	Included	1990–2080	20% 14%	
Doll and Siebert, 2002	Global Africa	All excluding rice	Process-based	IS92a	Included	2020s, 2070s	5.4% –1.4%	
Pan et al., 2013	Global	General	Process-based	A2 and B1	Included	2090s	38.9–14.5 %	
Liu et al., 2013	Zimbabwe	Wheat, rice, maize	Process-based	A1FI and B2	Included	2030s, 2090s	–9.95–11.95%	
Zhang and Cai, 2013	Zimbabwe Global	Wheat	Process-based	A1B-SAM	Included	2099		0–99mm
Elliott et al., 2014	Global	All	Process-based	-	Included	2090s	–15%	
Fant et al., 2013	Zimbabwe	Wheat	Process-based	-	Included	2050	5.8%	

**Figure 4.** Irrigation demand changes per region from reviewed studies

and Cai (2013) and Liu et al. (2013) of reduction in irrigation requirements despite increments in temperature can be explained by the diurnal temperature range (DTR; difference between daily maximum temperature and daily minimum temperature). Zhang and Cai (2013) note that increments in temperature may not cause higher evapotranspiration in cases where there is a decline in the DRT.

Fant et al. (2013) explored the impacts of climate change on irrigation requirements in the Zambezi basin which overlaps with the south-eastern part of Zimbabwe. Their robust analysis used a large pool of climate projection (6,800) based on the full set of the CMIP-3 GCMs for the 2050s time slice. Their analysis predicts a 6% increment in irrigation demand for cereals (excluding maize) for 4 countries in the basin, including Zimbabwe.

A statistical description of the results using a bar graph is shown in Fig. 3. It shows that across all time periods and emission scenarios crop water use at the global, regional and local scale will increase. There is, however a lot of uncertainty

in these estimations since most of the studies focused on a range of crops and not just wheat.

The implications of the increase in water requirements for wheat in Zimbabwe are significant. Recommended water requirement for wheat per season in the country ranges from 350 to 600 mm/ha depending on method of irrigation and geography (Bhasera et al., 2017). Rahman et al. (2015) determined that for the highveld areas of Harare and Domboshava the crop water requirements ranged from about 550 to 990 mm per season and varied significantly with irrigation methods used. In a survey of 41 commercial farms, Longmire et al. (1987) determined that the average total water actually applied in wheat farms was 570 mm (range 360 to 800 mm). These irrigation requirements in Zimbabwe where water availability has always been erratic and highly variable have made water resources the most limiting factor in the production of winter wheat. An increment in the irrigation requirements due to climate change and variability can compound the situation.

Climate change impacts on the water footprint of wheat production

The literature search on the impact of climate change on the actual WF of wheat in production in Zimbabwe yielded limited or no local studies. At a global or regional level only 3 papers by Fader et al., 2011, Deryng et al., 2016 and Konar et al., 2016 have explored the possible climate change impacts on crop WFs.

Fader et al. (2011) modelled global VWC under climate change scenarios. Values of VWC are equal and comparable to values for WF. The global study showed that by the 2070s, under A2 emission scenarios of future climate change and increasing atmospheric CO₂ concentrations, the WF for temperate cereals like wheat and barley will decrease globally. However, for many arid regions, such as Australia, South Africa, Argentina and the Mediterranean, the WF would increase. Although Zimbabwe is classified as semi-arid it was not included in the study. In the arid regions it is expected that the negative effects of climate change exceed the positive effects of CO₂ fertilization. An interesting analysis by Fader et al. (2010) showed that yields are the main driver of WF, rather than evapotranspiration; decreases in WF by more than 1 m³/kg were highly correlated to yield increases, and increases in WF by more than 1 m³/kg⁻¹ were highly correlated to yield decreases.

The critical role of yields in determining the WF was also noted by Konar et al. (2016) who analysed VWC under future climate change scenarios. Konar et al. (2016) based their assessment on projected yield shock scenarios (low, medium and high yield) predicted by Hertel et al. (2010). In their analysis global WF decreased under the medium- and high-yield scenarios for all crops but increased under low yield scenarios.

Perhaps the most intensive study on climate impacts on global WF of wheat production was carried out by Deryng et al. (2016). The researchers analysed climate change impacts on the CWP of wheat production at a global level. Since the CWP is the inverse of the WF their results are comparable to this study by taking the inverse of the CWP. Using a network of field experiments and an ensemble of process-based crop models and GCMs, Deryng et al. (2016) determined that CO₂ fertilization decreased global WF of cereals by 10–27% by the 2080s. In sharp contrast to the study by Fader et al. (2010), the study determined that the WF of crops grown in arid climates benefits the most from the effects of elevated CO₂ leading to additional significant reductions in consumptive crop water use by 2080. The contrast in results by Fader et al. (2010) and Deryng et al. (2016) highlights the fact that the impact of higher atmospheric CO₂ concentration is a major source of uncertainty in crop yield projections.

From the conflicting and few results of the three global studies on WF of wheat production there is a lot of uncertainty on how climate change will impact WF of wheat production for Zimbabwe. However, it is possible to extrapolate the results from the literature review on local yield and wheat water use. The literature review showed both a median decrease in wheat yields and increase in water consumption for Zimbabwe. Pooled together these results might indicate that climate change may result in increased WF for wheat production in the country. The possible increase in the local WF of wheat production may not be sustainable for the country which is classified as semi-arid with limited water resources.

The severity of the impact of climate change on the WF of wheat production depends on the actual increase in the WF and the current WF determined for the country at the present moment; a high WF would be further exacerbated by climate

change whilst a low WF can absorb and significantly neutralize climate change impacts. Since no local research has been conducted, the actual increase in WF cannot be ascertained without a significant amount of error. Furthermore, information on the current WF of wheat production in Zimbabwe is not known. Notwithstanding, Fader et al. (2011) reported that WF differs significantly among regions with highest values > 2 m³/kg common in large parts of Africa. There is thus a need for local research to determine the effect of climate change and variability on the WF of wheat production.

Possible climate change adaptation strategies

The foregoing discussion suggests that there may be a need for climate change adaptation strategies in Zimbabwe that decrease the WF of wheat production. Options for increasing wheat yields that have been documented in literature and may be categorized as crop-level adaptations (e.g. increasing yields and decreasing crop water use by breeding for higher yields, drought resistance or heat resistance (Deressa et al., 2009) and planned adaptations (e.g. expanding irrigation infrastructure or improvement in agricultural markets [Mendelsohn, 2001]). This study will primarily focus on crop-level adaptations partly because they fall under the study scope of the researchers and have been shown to be effective. Research shows that crop-level adaptations have the potential to boost yields by 7–15% more compared to similar scenarios that do not utilize adaptation (PCIC, 2014). A list of possible crop-level adaptations are listed in Table 3.

Crop breeding to develop new wheat varieties with higher yields, drought or heat resistance is touted as a possible solution to combat the negative effects of climate change on WF. High-yielding wheat varieties would result in more production with less or equal amounts of water applied compared to lower yielding varieties. This would reduce the WF of crop production.

There is, however, general disagreement over whether crop breeding for wheat would have any significant effect on yield. Compelling evidence suggests that in certain regions of the world wheat yields are plateauing and the rate of yield progress is falling to as low as 1.16% (Graybosch and Peterson, 2010; Cassman et al., 2010; Mackay et al., 2011). Hawkesford et al. (2013) notes that the main route in crop breeding for higher yielding cereals was increasing the harvest index (HI), which is currently at 60%, hinting that further increases above that value may not be physiologically possible. The highest wheat yield ever recorded is 16.791 t/ha (New Zealand) setting the current wheat yield barrier at 17 t/ha (Agrifac, 2017).

Zimbabwe has a lot of local high-yielding varieties that include SC Nduna (White) (11t/ha) and SC Sky (Red) (12 t/ha) (Seed Co, 2018). Considering that the highest national average wheat yields have been between 5 and 6 t/ha a significant yield gap still exists suggesting that higher yields can currently be attained by adopting good farm management practices rather than crop breeding. Some of these practices are highlighted in Table 3 and include good fertilizer management as well as timely sowing. The prospects of adapting to a higher WF of wheat production under climate change by decreasing crop water use can be done through a number of interventions. One possible strategy with a lot of potential is the adoption of irrigation scheduling techniques by local farmers. Historically studies have shown that there has been inefficient water use among wheat farmers in Zimbabwe with a tendency to over-apply irrigation water (Morris, 1988). Irrigation scheduling is the technique of determining the time, frequency and quantity

Table 3. Possible climate change adaptation strategies to decrease WF of wheat production in Zimbabwe

Options	Strategies	Description
Increasing wheat yields	Crop breeding	Breeding for drought resistance and heat tolerance to higher temperatures
	Fertilizer management	More soil testing, variable rate application better matching rates to crop demand to improve efficiency of fertilizer use
	Timely sowing:	Timely sowing so that flowering and grain-filling occurs after the period of heightened frost risk, but before the effects of late season water limitation and high temperature can reduce yield
Reducing crop evapotranspiration	Irrigation scheduling	Farmers training on methods of scientific irrigation scheduling; more research carried out to demonstrate the benefits of improved water management
	Deficit irrigation	Higher yields per unit of irrigation water applied
	Irrigation technology	Phasing out the use of sprinklers in preference to centre pivot systems with high application efficiency and flexibility
	Partial root-zone drying	Irrigating approximately half of the root system of a crop; more research to be carried out
	Shift planting dates	Matching farm activities with changes in temperature
Tillage and residue management:	Crop mulching and stubble retention	Use of crop residues or synthetic material to curb soil evaporation
	Zero tillage	Soil surface management to minimize soil evaporation and maximize infiltration

of irrigation water applications to reduce overall crop water use by reducing crop evapotranspiration. Irrigation scheduling is particularly suitable for Zimbabwe because inefficient water use has been traced to farmers' ignorance on how to implement scientific scheduling and inadequate research carried out to demonstrate the benefits of improved water management. Tambo and Senzanje (1988) noted that most wheat farmers rarely practice irrigation scheduling.

Deficit (or regulated deficit) irrigation is another potential strategy that can be used to adapt to climate change in Zimbabwe and reduce the WF for higher yields per unit of irrigation water applied. Deficit irrigation is defined as the application of water below the crop water requirements exposing the wheat crop to a particular level of water stress for a specified period (Feres, 2006). The assumption of deficit irrigation is that possible yield reductions due to water stress will be insignificant compared to the benefits derived by diverting water to other crops.

Research by Nyakatawa and Mugabe (1996) (see Table 4) showed that wheat yields of more than 4 t/ha can be expected when wheat is grown on deficit irrigation of at least 170 mm/season.

Good tillage and residue management on wheat farms can also help curb the effects of climate change on WF of wheat production by increasing yields and lowering crop evapotranspiration. Conventional tillage on wheat farms in Zimbabwe involves deep ploughing (ripping or chisel plough), followed by basal fertilizer application (option for liming), disking and then rolling (Basera and Soko, 2017). Conventional tillage practices have been linked to decreased water infiltration, increased soil evaporation and reductions in crop yield (Esser, 2017). Adopting conservation tillage practices together with the use of crop residues as mulching material

can reduce soil evaporation, increase infiltration, and boost yields, resulting in a lower WF. Gwenzi et al. (2008) in the Save Valley demonstrated that minimum tillage and no-tillage in a irrigated wheat-cotton rotation was more sustainable than conventional tillage; they improve soil structural stability and carbon sequestration. No studies on mulching on wheat farms have been conducted locally. However, the potential exists. In a meta-analysis Nyamangara et al. (2013) showed that conservation tillage and residue management may increase maize yields.

The adoption of efficient irrigation systems at the farm level can assist in decreasing the WF of wheat production under climate change. The majority of wheat farmers in Zimbabwe use overhead sprinkler systems (Tembo, 1988; Gambarara, 2016). Overhead sprinkler systems are composed of laterals that have many joints and have to be moved from one position to the other. They are thus associated with water losses and leakages which can increase the WF. Overhead sprinkler systems have an average application uniformity efficiency of 75%.

Centre pivot and drip irrigation systems have been replacing traditional flood irrigation and subsurface drip irrigation. Centre pivots and drip irrigation systems are highly efficient with up to 95% efficiency in terms of application uniformity. A study by Maisiri et al. (2005) in the semi-arid Insinza district of Matabeleland South province showed that drip irrigation systems on wheat farms use only 35% of the water used by the surface irrigation systems resulting in low WF of crop production.

CONCLUSIONS

This paper fills an important gap in climate impact studies by providing a review of the anticipated climate change impacts

Table 4. Experimental wheat yields in Zimbabwe under deficit irrigation

Irrigation level	Days to maturity		Number of ears/m ²		Grain yield	
Full	114.3	112.0	375	350	5 537	5 290
Three quarter	112.1	110.5	367	345	4 862	4 894
Half	109.7	108.5	358	318	4 431	4 223

on wheat yields, crop water and WF of wheat production in Zimbabwe. From the study, and in relation to the objectives set, it can be concluded that there is a dearth of information on the possible impact of climate change on the WF of wheat production in Zimbabwe. No local studies have been carried out in the country to determine climate change impacts on wheat yields, crop water use and the WF of wheat production.

Despite the scarcity of information global and regional studies hint towards an increase in the WF of wheat production in Zimbabwe under future climate change scenarios. The increase in the WF stems from a decrease in local wheat yields and increase in crop water use.

A lot of uncertainty exists at on the global and regional level on the precise effect of climate change on the WF of wheat production. The uncertainty is partly due to the predominant use of process-based models which might not be suitable for large spatial scale with high heterogeneity with respect to environmental conditions.

There is therefore need for more local studies to be carried out within Zimbabwe to ascertain accurate impacts of climate change on the WF of wheat production.

ACKNOWLEDGMENTS

The researchers would like to appreciate Chinhoyi University of Technology Staff Development Fund for funding this research.

REFERENCES

- ADAMS RM, MCCARL BA, SEGERSON K, ROSENZWEIG C, BRYANT KJ, DIXON BL, CONNER R, EVENSON RE and OJIMA D (1999) The economic effects of climate change on US agriculture. In: Mendelsohn R and Neumann JE (eds) *The Impact of Climate Change on the United States Economy*. Cambridge University Press, Cambridge.
- AGRIFAC (2017) 50% higher yield, 50% less chemicals. URL: <https://www.agrifac.com/50-percent/19603> (Accessed May 2018).
- AMARASINGHE U and SMAKHTIN V (2014) Water productivity and water footprint: misguided concepts or useful tools in water management and policy? *Water Int.* **39** (7) 1000–1017. <https://doi.org/10.1080/02508060.2015.986631>
- ASSENG S, EWERT F, MARTRE P, RÖTTER RP, LOBELL DB, CAMMARANO D, KIMBALL BA, OTTMAN MJ, WALL GW, WHITE JW and co-authors (2014) Rising temperatures reduce global wheat production. *Nat. Clim. Change* **5** 143–147. <http://doi.org/10.1038/NCLIMATE2470>
- BAZILIAN M, ROGNER H, HOWELLS M, HERMANN S, ARENT D, GIELEND T, STEDUTO P, MUELLERA A, KOMOR P, TOL RS, YUMKELLA KK and co-authors (2011) Considering the energy, water and food nexus: Towards an integrated modelling approach. *Energ. Polic.* **39** (12) 7896–7906. <https://doi.org/10.1016/j.enpol.2011.09.039>
- BHASERA J and SHOKO T (2017) Basic principles of wheat irrigation. Commercial Farmers Union of Zimbabwe. URL: <http://www.cfum.org/index.php/agriculture/8392-basic-principles-of-wheat-irrigation>. (Accessed May 2018).
- BROWN DW, CHANAKIRA RR, CHATIZA K, DHLIWAYO M, DODMAN D, MUGABE PH, ZVIGADZA S, MASIIWA M, MUCHADENYIKA D and RANCE R (2012) Climate change impacts, vulnerability and adaptation in Zimbabwe. Working Paper. URL: <https://pubs.iied.org/10034IIED/> (Accessed May 2018).
- CAI W, WANG G, SANTOSO A, MCPHADEN MJ, WU L, JIN FF, TIMMERMAN A, COLLINS M, VECCHI G, LENGAINNE M and co-authors (2015) Increased frequency of extreme La Niña events under greenhouse warming. *Nat. Clim. Change* **5** 132–137. <https://doi.org/10.1038/nclimate2492>
- CAIRNS JE, MAGOROKOSHO D, COLSEN C, SONDER M, STIRLING M, KINDIE TESFAYE C, MAINASSARA ZAMAN-ALLAH T and PRASANNA BM (2013) Adapting maize production to climate change in sub-Saharan Africa. *Food Security* **5** (3) 345–360. <https://doi.org/10.1007/s12571-013-0256-x>
- CARTWRIGHT A, BLIGNAUT J, DE WIT J, GOLDBERG M, MANDER K, O'DONOGHUE M and ROBERTS (2013) Economics of climate change adaptation at the local scale under conditions of uncertainty and resource constraints: the case of Durban, South Africa. *Environ. Urbanization* **25** (1) 139–156. <https://doi.org/10.1177/0956247813477814>
- CASSMAN KG, GRASSINI P and VAN WART J (2010) Crop yield potential, yield trends and global food security in a changing climate. In: Illel D and Rosenzweig C (eds) *Handbook of Climate Change and Agroecosystems: Impacts, Adaptation, and Mitigation*. World Scientific, London.
- CHAGUTAH T (2010) Climate Change Vulnerability and Preparedness in Southern Africa: Zimbabwe Country Report. URL: https://za.boell.org/sites/default/files/downloads/HBF_web_Zim_21_2.pdf. (Accessed May 2018).
- CHALLINOR A, WATSON J, LOBELL D, HOWDEN BM, SMITH DR and CHETRI N (2014) A meta-analysis of crop yield under climate change and adaptation. *Nat. Clim. Change* **4** (4) 287–291. <https://doi.org/10.1038/nclimate2153>
- CHAWARIKA A, MUTAMBARA J and CHAMBOKO T (2017) Competitiveness of wheat production in Zimbabwe. *Afr. J. Sci. Technol. Innovation Dev.* **9** (3) 263–268. <https://doi.org/10.1080/20421338.2017.1322349>
- CHRISTENSEN JH, HEWITSON B, BUSUIOC A, CHEN A, GAO X, HELD I, JONES R, KOLLI RK, KWON WT, LAPRISE R and co-authors (2007) Regional Climate Projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M and Miller HL (eds) *Climate Change 2007: The Physical Science Basis Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge UK.
- COBB KM, WESTPHAL N, SAYANI H, WATSON JT, LORENZO ED, CHENG H, EDWARDS RL and CHARLES CA (2013) Highly variable El Niño–Southern Oscillation throughout the Holocene. *Science* **339** (6115) 67–70. <https://doi.org/10.1126/science.1228246>
- DEGENER JF (2015) Atmospheric CO₂ fertilization effects on biomass yields of 10 crops in northern Germany Front. *Environ. Sci.* **3** (1) 48. <https://doi.org/10.3389/fenvs.2015.00048>
- DERESSAA A, HASSANB TT, RINGLER L, YESUF C and TAM Y (2009) Determinants of farmers' choice of adaptation methods to climate change in the Nile Basin of Ethiopia. *Glob. Environ. Change* **19** (2) 248–255. <https://doi.org/10.1016/j.gloenvcha.2009.01.002>
- DERYNG D, CONWAY D, RAMANKUTTY N, PRICE J and WARREN R (2014) Global crop yield response to extreme heat stress under multiple climate change futures. *Environ. Res. Lett.* **9** (3) 13. <http://dx.doi.org/10.1088/1748-9326/9/3/034011>
- DERYNG D, ELLIOTT J, FOLBERTH C, MÜLLER C, PUGH TAM, BOOTE KJ, CONWAY D, RUANE AC, GERTEN D, JONES JW, KHABAROV N and co-authors (2016) Regional disparities in the beneficial effects of rising CO₂ concentrations on crop water productivity. *Nat. Clim. Change* **6** (3) 786–790. <https://doi.org/10.1038/nclimate2995>
- DÖLL P and SIEBERT S (2002) Global modeling of irrigation water requirements. *Water Resour. Res.* **38** (4) 8810. <https://doi.org/10.1029/2001WR000355>
- ELLIOTT J, DERYNG D, MÜLLER C, FRIELER K, KONZMANN M, GERTEN D, GLOTTER M, FLÖRKE M, WADA Y, BEST N and co-authors (2014) Constraints and potentials of future irrigation water availability on agricultural production under climate change. *Proc. Natl. Acad. Sci.* **111** (9) 3239–3244. <https://doi.org/10.1073/pnas.1222474110>
- ESSER KB (2017) Water Infiltration and Moisture in Soils under Conservation and Conventional Agriculture in Agro-Ecological Zone IIa, Zambia. *Agronomy* **7** (2) 40. <https://doi.org/10.3390/agronomy7020040>
- FADER M, ROST S, MÜLLER C, BONDEAU A and GERTEN D (2011) Virtual water content of temperate cereals and maize: Present and potential future patterns. *J. Hydrol.* **384** (34) 218–231. <https://doi.org/10.1016/j.jhydrol.2011.05.011>

- [org/10.1016/j.jhydrol.2009.12.011](https://doi.org/10.1016/j.jhydrol.2009.12.011)
- FALKENMARK M, ROCKSTRÖM J and KARLBERG L (2008) Present and future water requirements for feeding humanity. *Food Secur.* **1** (1) 59–69. <https://doi.org/10.1007/s12571-008-0003-x>
- FANT C, GEBRETSADIK Y and STRZEPEK K (2013) Impact of Climate Change on Crops, irrigation and hydropower in the Zambezi River Basin 2013. WIDER Working Paper No 2013/03. URL: www.wider.unu.edu (Accessed May 2018).
- FEDOROV AV and PHILANDER SG (2000) Is El Nino changing. *Science* **288** (5473) 1997–2002. <https://doi.org/10.1126/science.288.5473.1997>
- FERERES E and SORIANOAM T (2007) Deficit irrigation for reducing agricultural water use. *J. Exp. Bot.* **58** (2) 147–159. <https://doi.org/10.1093/jxb/erl165>
- FISCHER G, TUBIELLO FN, VELTHUIZEN VH and WIBERG A (2007) Climate change impacts on irrigation water requirements: Effects of mitigation, 1990–2080. *Technol. Forecast. Soc. Change* **74** (7) 1083–1107. <https://doi.org/10.1016/j.techfore.2006.05.021>
- GLEICK PH (2003) Global freshwater resources: soft-path solutions for the 21st century. *Science* **302** (5650) 1524–1528. <https://doi.org/10.1126/science.1089967>
- GOPO ML and NANGOMBE S (2016) El Nino not climate change. *The Herald. 20 January, 2016* URL: <https://www.herald.co.zw/el-nino-not-climate-change/> (Accessed December 2017).
- GORNALL J, BETTS RA, BURKE EJ, CLARK R, CAMP J, WILLET KW and WILTSHIRE AJ (2010) Implications of climate change for agricultural productivity in the early twenty-first century. *Phil. Trans. R. Soc. Lond. Ser. B, Biol. Sci.* **365** (1554). <https://doi.org/10.1098/rstb.2010.0158>
- GoZ (2014) Zimbabwe National Climate Change Response Strategy. URL: www.aczw/downloads/draftstrategy.pdf (Accessed May 2018).
- GRAYBOSCH RA and PETERSON CJ (2010) Genetic improvement in winter wheat yields in the Great Plains of North America, 1959–2008. *Crop Sci.* **50** (5) 1882–1890. <https://doi.org/10.2135/cropsci2009.11.0685>
- GWENZI W, GOTOSA J, CHAKANETSAS S and MUTEMA J (2008) Effects of tillage systems on soil organic carbon dynamics, structural stability and crop yields in irrigated wheat (*Triticum aestivum* L.)–cotton (*Gossypium hirsutum* L) rotation in semi-arid Zimbabwe. *Nutr. Cycl. Agroecosyst.* **83** (8) 211. <https://doi.org/10.1007/s10705-008-9211-1>
- HADDELAND I, HEINKE J, BIEMANS HA, EISNER S, FLÖRKE M, HANASAKI N, KONZMANN M, LUDWIG, F, MASAKI Y, SCHEWE J and co-authors (2014) Global water resources affected by human interventions and climate change. *PNAS* **111** (9) 3251–3256. <https://doi.org/10.1073/pnas.1222475110>
- HAVAZVIDI EK (2012) Selection for high grain yield and determination of heritable variation in biomass yield, grain yield and harvest index in different plant height classes of breeding populations of spring wheat (*Triticum aestivum* L). PhD thesis, Department of Crop Science, University of Zimbabwe, Harare.
- HAWKESFORD MJ, ARAUS JL, PARK R, CALDERINI D, MIRALLES T, JIANPING S and PARRY M (2013) Prospects of doubling global wheat yields. *Food Energ. Secur.* **2** (1) 34–48. <https://doi.org/10.1002/fes3.15>
- HERTEL TW and LOBELL DB (2017) Agricultural adaptation to climate change in rich and poor countries: Current modelling practice and potential for empirical contributions. *Energ. Econ.* **46** (C) 562–575. <https://doi.org/10.1016/j.eneco.2014.04.014>
- HOEKSTRA AY, CHAPAGAIN A, MARTINEZ-ALDAYA M and MEKONNEN M (2009) *Water Footprint Manual: State of the Art 2009*. Water Footprint Network, Netherlands.
- HOEKSTRA AY and MEKONNEN MM (2012) The water footprint of humanity. *PNAS* **109** (9) 3232–3237. <https://doi.org/10.1073/pnas.1109936109>
- HULME M, DOHERTY RM, NGARA T, NEW MG and LISTER D (2001) African climate change: 1900–2100. *Clim. Res.* **17** (2) 145–168. <https://doi.org/10.3354/cr017145>
- JONES AG, SCULLION J, OSTLE J, LEVY N and GWYNN-JONES PN (2014) Completing the FACE of elevated CO₂ research. *Environ. Int.* **7** (16) 252–258. <https://doi.org/10.1016/j.envint.2014.07.021>
- KAPUYA T, SARUCHERA D, JONGWE A, MUCHERI T, MUJEYI K, LULAMA N and FERDINAND M (2010) The grain industry value chain in Zimbabwe. URL: http://www.fao.org/fileadmin/templates/est/AAACP/estafrica/UnvPretoria_GrainChainZimbabwe_2010_1.pdf. (Accessed May 2018).
- KNOX J, WHEELER JK, HESS T and DACCACHE A (2012) Climate change impacts on crop productivity in Africa and South Asia. *Environ. Res. Lett.* **7** (3) 34032. <https://doi.org/10.1088/1748-9326/7/3/034032>
- KONAR M, HANASAKI M, REIMER JJ, HUSSEIN Z and NAOTA H (2016) The water footprint of staple crop trade under climate and policy scenarios. *Environ. Res. Lett.* **11** (3) 35006. <https://doi.org/10.1088/1748-9326/11/3/035006>
- KUMMU M, GERTEN D, HEINKE J, KONZMAN M and VARIS O (2014) Climate-driven interannual variability of water scarcity in food production potential: a global analysis. *Hydrol. Earth Syst. Sci.* **18** (2) 447–461. <https://doi.org/10.5194/hess-18-447-2014>
- LEBEL S, FLESKENS L, FORSTER PM, JACKSON LS and LORENZ S (2015) Evaluation of in situ rainwater harvesting as an adaptation strategy to climate change for maize production in rainfed Africa. *Water Resour. Manage.* **29** (13) 4803–4816. <https://doi.org/10.1007/s11269-015-1091-y>
- LIU J, FOLBERTH C, YANG H, RÖCKSTRÖM J, ABBASPOUR K and ZEHNDER AJB (2013) A global and spatially explicit assessment of climate change impacts on crop production and consumptive water use. *PLoS ONE* **8** (2) e57750. <https://doi.org/10.1371/journal.pone.0057750>
- LOBELL DB, BURKE MB, TEBALDI C, MASTRANDREA MD, FALCON WP and NAYLOR RL (2008) Prioritizing climate change adaptation needs for food security in 2030. *Science* **319** (5863) 607–610. <https://doi.org/10.1126/science.1152339>
- LOBELL DB, SCHLENKER W and COSTA-ROBERTS J (2011) Climate trends and global crop production since 1980. *Science* **333** (6042) 616–620. <https://doi.org/10.1126/science.1204531>
- LONGMIRE J, NGOBESE P and TEMBO S (1987) Wheat policy options in Zimbabwe and SADCC countries: preliminary findings. In: *Proceedings of the Third Annual Conference on Food Security Research in Southern Africa*, 1–5 November 1987. University of Zimbabwe, Harare.
- MACKAY I, HORWELL A, GARNER J, WHITE J, MCKEE J and PHILPOTT H (2011) Reanalyses of the historic series of UK variety trials to quantify the contributions of genetic and environmental factors to trends and variability in yield over time. *Theor. Appl. Genet.* **122** (1) 225–238. <https://doi.org/10.1007/s00122-010-1438-y>
- MACROBERT JF and SAVAGE MJ (1998) The use of a crop simulation model for planning wheat irrigation in Zimbabwe. In: Tsuji GY, Hoogenboom G and Thornton PK (eds) *Understanding Options for Agricultural Production*. Springer, Dordrecht. 205–220.
- MADZWAMUSE M (2010) Climate governance in Africa: Adaptation strategies and institutions. A synthesis report. URL: <https://www.boell.de/en/ecology/africa-climate-governance-in-africa-adaptation-strategies-and-institutions-10914.html>. (Accessed May 2018).
- MAISIRIN, SENZANJE A, ROCKSTROM J and TWOMLOW SJ (2005) On farm evaluation of the effect of low cost drip irrigation on water and crop productivity compared to conventional surface irrigation system. *Phys. Chem. Earth* **30** (11–16) 783–791. <https://doi.org/10.1016/j.pce.2005.08.021>
- MAKADHO J (1996) Irrigation timeliness indicators and application in smallholder irrigation systems in Zimbabwe. *Irrig. Drain. Syst.* **10** (4) 367–376. <https://doi.org/10.1007/BF01104900>
- MAKARAU A (1999) Zimbabwe climate: Past Present and Future In: Manzungu E, Senzanje A and Van Der Zaag P (eds) *Water for Agriculture: Policy and Management Options for the Smallholder Sector*. University of Zimbabwe Publications, Harare.
- MAKUVARO V (2014) Impact of climate change on smallholder farming in Zimbabwe, using a modelling approach. Master's thesis, University of Zimbabwe.
- MALABA J (2013) *Poverty and Poverty Datum Line Analysis in Zimbabwe 2011/12*. Zimbabwe Statistical Agency, Harare.
- MANJENGWA J, KASIRYE I and MATEMA C (2012) Understanding Poverty in Zimbabwe: A Sample Survey in 16 Districts. In: *Proceedings of the Centre for the Study of African Economies Conference*, 2012, Oxford, United Kingdom.

- MANO R and NHEMACHENA C (2007) Assessment of the economic impacts of climate change on agriculture in Zimbabwe: a Ricardian approach. Policy research working paper. URL: <https://doi.org/10.1596/1813-9450-4292>. (Accessed May 2018).
- MANYERUKE C, HAMAUSWA S and MHANDARA L (2013) The effects of climate change and variability on food security in Zimbabwe: a socio-economic and political analysis. *Int. J. Human. Soc. Sci.* **3** (6) 270. http://www.ijhssnet.com/journals/Vol_3_No_6_Special_Issue_March_2013/26.pdf.
- MASHIRINGWANI NA and MTISI E (1994) Wheat production and research in Zimbabwe: Constraints and sustainability. URL: <http://agrisfao.org/agris-search/searchdo?recordID=QY9400010>.
- MATIU M, ANKERST DP and MENZEL A (2017) Interactions between temperature and drought in global and regional crop yield variability during 1961-2014. *PLoS ONE* **12** (5) e0178339. <https://doi.org/10.1371/journal.pone.0178339>
- MAZVIMAVI D (2010) Investigating changes over time of annual rainfall in Zimbabwe. *Hydrol. Earth Syst. Sci.* **14** (12) 2671-2679. <https://doi.org/10.5194/hess-14-2671-2010>
- MENDELSON R, DINAR A and SANGHI P (2001) The effect of development on the climate sensitivity of agriculture. *Environ. Dev. Econ.* **6** (1) 85-101. <https://doi.org/10.1017/s1355770x01000055>
- MOORE FC, BALDOS UL and CHERTEL T (2017) Economic impacts of climate change on agriculture: a comparison of process-based and statistical yield models. *Environ. Res. Lett.* **12** (6) 065008. <https://doi.org/10.1088/1748-9326/aa6eb2>
- MORISON JI, BAKER N, MULLINEAUX P and DAVIES W (2008) Improving water use in crop production. *Phil. Trans. R. Soc. B: Biol. Sci.* **363** (1491) 639-658. <https://doi.org/10.1098/rstb.2007.2175>
- MORRIS LM (1988) Comparative advantage and policy incentives for wheat production in Zimbabwe. CIMMYT Economics Working Paper 88/02. URL: <https://repository.cimmyt.org/xmlui/bitstream/handle/10883/1125/19910.pdf>. (Accessed May 2018).
- MUCHENA P and IGLESIAS A (1995) Vulnerability of maize yields to climate change in different farming sectors in Zimbabwe. In: Rosenzweig C (ed) *Climate Change and Agriculture: Analysis of Potential International Impacts*. ASA Spec Publ 59 ASA, Madison.
- MUGABE F and NYAKATAWA E (2000) Effect of deficit irrigation on wheat and opportunities of growing wheat on residual soil moisture in southeast Zimbabwe. *Agric. Water Manage.* **46** (2) 111-200. [https://doi.org/10.1016/S0378-3774\(00\)00084-6](https://doi.org/10.1016/S0378-3774(00)00084-6)
- MUJERE N and MAZVIMAVI D (2012) Impact of climate change on reservoir reliability. *Afr. Crop Sci. J.* **20** (2) 545-551. <http://doi.org/10.4314/acsj.v20i2>
- MÜLLER C (2011) Climate change risks for African agriculture. *PNAS* **108** (11) 4313-4315. <https://doi.org/10.1073/pnas.1015078108>
- MURDOCK TQ and ZWIERS FW (2015) Regional Climate Impacts - Research Plan: 2015-2019. URL: https://www.pacificclimate.org/sites/default/files/publications/Murdock.RCI_Plan_Apr2012_0.pdf.
- MUTAMBARA J, ZVINAASHE AP and MWAKIWA E (2013) A critical review of the wheat industry in Zimbabwe. *GJ BAHs* **2** (2) 23-33. <https://www.longdom.org/articles/a-critical-review-of-the-wheat-industry-in-zimbabwe.pdf>
- NASA (2016) NOAA data show 2016 warmest year on record globally. URL: <https://www.nasagov/press-release/nasa-noaa-data-show-2016-warmest-year-on-record-globally> (Accessed May 2018).
- NELSON F, ROSEGRANT GC, KOO M, J ROBERTSON J, SULSER RD, ZHU T, RINGLER T, MSANGI C, PALAZZO S, BATKA T and co-authors (2009) Climate change: Impact on agriculture and costs of adaptation. Food Policy Report International Food Policy Research Institute (IFPRI). URL: <http://ebraryifpriorg/cdm/ref/collection/p15738coll2/id/130648> (Accessed May 2018).
- NELSON GC, VALIN RD, SANDS P, HAVLÍK H, AHAMMAD D, DERYNG D, ELLIOTT J, FUJIMORI S, HASEGAWA T, HEYHOE E and co-authors (2014) Climate change effects on agriculture: Economic responses to biophysical shocks. *Proc. Natl. Acad. Sci.* **111** (9) 3274-3279. <https://doi.org/10.1073/pnas.1222465110>
- NYAMANGARA J, NYENGERAI K, MASVAYA EN, TIRIVAVI R, MASHINGAIDZE N, MUPANGWA W, DIMES J, HOVE L and TWOMLOW S (2013) Effect of conservation agriculture on maize yield in the semi-arid areas of Zimbabwe. *Exp. Agric.* **50** (2) 159-177. <https://doi.org/10.1017/s0014479713000562>
- OKI T and KANAE S (2006) Global hydrological cycles and world water resources. *Science* **313** (5790) 1068-1072. <https://doi.org/10.1126/science.1128845>
- PAN S, TIAN H, DANGAL SR, YANG Q, YANG J, LU C, TAO B, REN W and OUYANG Z (2015) Responses of global terrestrial evapotranspiration to climate change and increasing atmospheric CO₂ in the 21st century. *Earth's Future* **3** (1) 15-35. <https://doi.org/10.1371/journal.pone.0112810>
- PARRY ML, ROSENZWEIG C, IGLESIAS A, LIVERMORE A and FISCHER G (2004) Effects of climate change on global food production under SRES emissions and socio-economic scenarios. *Glob. Environ. Change* **14** (1) 53-67. <https://doi.org/10.1016/j.gloenvcha.2003.10.008>
- RAHMAN AG, TALAAT AM and ZAWA C (2016) Water requirements for main crops grown under three different agro ecological zones, Zimbabwe. *Middle East J. Agric. Res.* **5** (1) 14-28. <http://www.wcurresweb.com/mejar/mejar/2016/14-28.pdf>.
- RAMANKUTTY TI and LUZIMU T (2016) Changes in yield variability of major crops for 1981-2010 explained by climate change. *Environ. Res. Lett.* **11** (3) 34003. <http://dx.doi.org/10.1088/1748-9326/11/3/034003>
- RAY DK, GERBER JS, MACDONALD GK and WEST PC (2015) Climate variation explains a third of global crop yield variability. *Nat. Commun.* **6** (12) (5989). <https://doi.org/10.1038/ncomms6989>
- REKACEWICZ P (2005) Climate change in Zimbabwe: Trends in temperature and rainfall. URL: www.gridano.com/graphicslib/detail/climate-change-in-zimbabwe-trends-temperature-and-rainfall_85e5 (Accessed May 2018).
- RINGLER C, TINGJU Z, XIMING C, JAWOO K and DINGBAO W (2010) Climate change impacts on food security in Sub-Saharan Africa: Insights from comprehensive climate change scenarios IFPRI Discussion. URL: <http://ebraryifpriorg/cdm/ref/collection/p15738coll2/id/6983> (Accessed May 2018).
- ROBERTSON MJ, HE D, WANG E and WANG J (2017) Data requirement for effective calibration of process-based crop models. *Agric. For. Meteorol.* **234-235** 136-148. <https://doi.org/10.1016/j.agrformet.2016.12.015>
- ROSENZWEIG C and PARRY ML (1994) Potential impact of climate change on world food supply. *Nature* **367** (234) 133-138. <https://doi.org/10.1038/367133a0>
- RURINDA J, VAN WIJK MT, MAPFUMO P, DESCHEMAEKER K, SUPIT I and GILLER KE (2015) Climate change and maize yield in southern Africa: what can farm management do? *Glob. Change Biol.* **21** (12) 4588-4601. <https://doi.org/10.1111/gcb.13061>
- SAKUHUNI RC, CHIDOKO NL, DHORO A and GWAENDEPI C (2011) Economic Determinants of Poverty in Zimbabwe. *Int. J. Econ. Res.* **2** (6) 1-12.
- SHIKLOMANOV IA and RODDA JC (2003) *World Water Resources at the Beginning of the Twenty-First Century*. Cambridge University Press, Cambridge.
- SIEBERT S, WEBBER H, ZHAO G and EWERT F (2017) Heat stress is overestimated in climate impact studies for irrigated agriculture. *Environ. Res. Lett.* **12** 054023. <http://dx.doi.org/10.1088/1748-9326/aa702f>
- SUN SK, WU PT, WANG YB and ZHAO XN (2012) Impacts of climate change on water footprint of spring wheat production: The case of an irrigation district in China. *Span. J. Agric. Res.* **10** (4). <http://dx.doi.org/10.5424/sjar/2012104-3004>
- SWATUK LA (2005) Political challenges to implementing IWRM in Southern Africa. *Phys. Chem. Earth* **30** (34) 872-880. <https://doi.org/10.1016/j.pce.2005.08.033>
- TAMBUSSI EA, BORT J and ARAUS JL (2007) Water use efficiency in C3 cereals under Mediterranean conditions: a review of physiological aspects. *Ann. Appl. Biol.* **150** (3) 307-321. <https://doi.org/10.1111/j.1744-7348.2007.00143.x>
- TATSUMI K, YAMASHIKI Y, VALMIR DA SILVA R, TAKARA K, MATSUOKA Y, TAKAHASHI K, MARUYAMA K and KAWAHARA N (2011) Estimation of potential changes in cereals production under climate change. *Hydrol. Process.* **25** (17) 715-725. <https://doi.org/10.1002/hyp.8012>
- TEMBO S and SENZANJE A (1987) Water-use efficiency on

- commercial wheat farms in Zimbabwe. In: *Proceedings of the Third Annual Conference on Food Security Research in Southern Africa*, 1–5 November 1987, University of Zimbabwe, Harare.
- UNGANAI L (1996) Historic and future climatic change in Zimbabwe. *Clim. Change* **6** 137–145. <http://dx.doi.org/10.3354/cr006137>
- WALKER NJ and SCHULZE RE (2008) Climate change impacts on agro-ecosystem sustainability across three climate regions in the maize belt of South Africa. *Agric. Ecosyst. Environ.* **124** (1–2) 114–124. <https://doi.org/10.1016/j.agee.2007.09.001>
- WANG G and HENDON HH (2007) Sensitivity of Australian rainfall to inter–El Niño variations. *J Clim.* **20** (5) 4211–4226. <https://doi.org/10.1175/JCLI4228.1>
- WIEBER KW, LOTZE-CAMPEN H, SANDS R, TABEAU AA and MEIJL JCM (2015) Climate change impacts on agriculture in 2050 under a range of plausible socioeconomic and emissions scenarios. *Environ. Res. Lett.* **10** (8) 34–38. <https://doi.org/10.1088/1748-9326/10/8/085010>
- ZAMPIERI M (2018) Understanding and reproducing regional diversity of climate impacts on wheat yields: current approaches, challenges and data driven limitations *Environ. Res. Lett.* **13** (2) 112–119. <https://doi.org/10.1088/1748-9326/aaa00d>
- ZHANG C and ANADON LD (2014) A multi-regional input–output analysis of domestic virtual water trade and provincial water footprint in China. *Ecol. Econ.* **100** (2) 159–172. <https://doi.org/10.1016/j.ecolecon.2014.02.006>
- ZHANG X and CAI X (2013) Climate change impacts on global agricultural water deficit. *Geophys. Res. Lett.* **40** (4) 1111–1117. <https://doi.org/10.1002/grl.50279>
- ZHAO C, ZHAO C, LIU C, PIAO S, WANG X, LOBELL DB, HUANG D, HUANG M, YAO Y, BASSU S and co-authors (2017) Temperature increase reduces global yields of major crops in four independent estimates. *PNAS* **114** (35) 9326–9331. <https://doi.org/10.1073/pnas.1701762114>
- ZWART SJ and BASTIAANSEN WGM (2004) Review of Measured Crop Water Productivity Values for Irrigated Wheat, Rice, Cotton and Maize. *Agric. Water Manage.* **69** (2) 115–133. <https://doi.org/10.1016/j.agwat.2004.04.007>
-

Sensors for the improvement of irrigation efficiency in nurseries

Nkosinathi D Kaptein^{1,2*}, Marnie E Light^{1,3} and Michael J Savage²

¹*Institute for Commercial Forestry Research, Pietermaritzburg, South Africa*

²*Soil-Plant-Atmosphere Continuum Research Unit, School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa*

³*Research Centre for Plant Growth and Development, School of Life Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa*

ABSTRACT

Traditional timer-based systems for irrigation management, which are more commonly used in commercial nurseries in South Africa, are not ideal as they may not irrigate seedlings efficiently. A sensor-based irrigation system is presented as an alternative, as this can provide several benefits to nurseries and nursery-grown seedlings. Small-sized soil water sensors that could fit in small-volume nursery containers (25 to 100 mL), and could be integrated into an automated irrigation system, are reviewed. Several experiments have been conducted internationally to measure soil water status of small-volume containers in soilless substrates, and a large body of knowledge is now available. In this review, we describe the principles of several currently commercially available sensors that can be adapted to this purpose, giving advantages and disadvantages of each type. We conclude that a sensor-based irrigation system has great potential to address the challenges associated with irrigation scheduling, while improving water usage in most nurseries.

Keywords: automation, porous substrates, seedlings, small-volume container, timer system

INTRODUCTION

Irrigation water is becoming increasingly scarce in South Africa, with the agricultural industry using approximately 62% of available fresh water (Pimentel et al., 2004; Fanadzo and Ncube, 2017). Ever-increasing demand for this limited resource for household and industrial use, as well as recent droughts, has resulted in an urgent need for water conservation (Gleeson et al., 2012). Thus, in commercial nurseries which support the agricultural and forestry sectors, it is becoming increasingly important to develop efficient irrigation systems. Improved irrigation practices by nurseries can save water and reduce the costs of irrigation, and have the potential to improve seedling quality, reduce leaching of nutrients, and lower the incidence of pathogen infestation (Belayneh et al., 2013; Lichtenberg et al., 2013; Bayer et al., 2015; Saavoss et al., 2016; Lea-Cox et al., 2017; Wheeler et al., 2017). Many biological and biophysical processes, such as seed germination, seedling nutrition and growth, as well as transpiration and evaporative cooling, are dependent on sufficient moisture via efficient irrigation (Bittelli, 2010; 2011). In most commercial nurseries, seedlings are grown in containers with small-volume cavities ranging from 25 to 100 mL (examples of some forestry seedling containers are shown in Fig. 1; Durner, 2013). This optimises the utilisation of nursery and transportation space while ensuring the best performance when planted in the field. In such containers, plant roots have a limited substrate volume to explore for water. Although space saving, one of the difficulties with these containers is that the substrate water status can range from near-saturation immediately after irrigation, to near-dryness after several hours without irrigation (Van Iersel et al., 2011; Montesano et al., 2016; Lea-Cox et al., 2017). Irrigation scheduling for these containers should, therefore,

aim at maintaining the media water content at levels that minimise seedling water stress and maximise irrigation efficiency during periods that require stress-free growth (Van Iersel et al., 2011).

Commercial nurseries generally irrigate their seedlings based on the visual appearance of the substrate (wet or dry), their intuition and experience (growers may pick up the container to feel its relative 'weight'), and often rely on a fixed timer-based system (Jones, 2004; 2008; Lea-Cox et al., 2011; 2017). Although the settings of a timer-based system can be adjusted, it is set to irrigate according to a predetermined schedule to apply water at a particular time of day and for a particular duration (Nemali and Van Iersel, 2006; Montesano et al., 2016). Such timer-based systems have been widely adopted due to their ease of use, relatively low cost, ease of programming and success in irrigating correctly when managed closely during periods of peak water demand (Nemali et al., 2007; Lea-Cox, 2012). However, they can be ineffective and inefficient since water requirements may vary by species, season, microclimate and changes in root density or leaf area as the plant grows (Lea-Cox, 2012). For example, such systems can waste water on cooler and cloudy days due to fixed irrigation schedules which operate irrespective of weather changes (Van Iersel et al., 2011). In addition, most nursery managers are risk averse, preferring to apply excess water to ensure against system failure or heterogeneity of application (Jones, 2004; 2008). However, with increasing water costs and restrictions, it is becoming more necessary to limit excessive water application. Therefore, scheduling irrigation under fixed timer-based systems may be inaccurate and costly and, considering these issues, there is a need for a more efficient method.

The use of soil water sensors to control automated irrigation systems is proposed as a useful alternative. However, sensor-based irrigation systems are not widely used in South African nurseries. This may be due to perceived (i) high start-up costs, too expensive to adopt on a large

*Corresponding author, email: nkosinathi.kaptein@icfr.ukzn.ac.za
Received 20 November 2017; accepted in revised form 26 June 2019



Figure 1. Examples of different sizes of small containers used in many commercial forestry nurseries showing their associated volumes

scale; (ii) difficulty in automation; and (iii) low reliability and maintenance issues (Annandale et al., 2011; Lea-Cox, 2012; Belayneh et al., 2013). For most commercial nurseries, the initial cost and the ease of use are the most crucial considerations for adopting a sensor-based irrigation system. Regardless of these challenges, numerous sensors with the potential to be adapted for use in nursery substrates are well documented in the literature (Burnett and Van Iersel, 2008; Chappell et al., 2013; Bayer et al., 2015; Montesano et al., 2016; Saasvoss et al., 2016; Wheeler et al., 2017). Increasing freshwater scarcity, and pressure from governments, society, and concerned bodies to conserve water, gives motive to investigate sensor-based irrigation systems. Thus, the aim of

this review was to evaluate commercially available soil water sensors that can be connected to an automated irrigation system to measure and control water status in small-volume nursery containers to improve nursery water management.

SCHEDULING IRRIGATION IN SMALL-VOLUME CONTAINERS

Soilless substrates, as used in many commercial nurseries, are generally porous and have relatively large particle sizes compared to mineral soils. They tend to release more water at very low matric potential (-1 to -40 kPa), which is 10 to 100 times lower than the matric potential in mineral soils (Lea-Cox et al., 2011). Also, the rooting system of a nursery plant is confined to the volume of the container. For these reasons, maintaining optimal water status for soilless substrates is critical for continued plant growth. A sensor-based system for scheduling irrigation is a logical choice, and consists of sensors linked to an automated irrigation system that periodically measure the substrate water status at a specified interval (Nemali and Van Iersel, 2006; Van Iersel et al., 2013; Lea-Cox et al., 2017; Kaptein et al., 2019). Since water use by the seedling, along with drainage and evaporation, causes a decrease in substrate water level over time, the sensors detect these changes in the substrate volumetric water content (VWC, $\text{m}^3 \cdot \text{m}^{-3}$) or soil water potential (Ψ , kPa), and relay that information to an irrigation controller (Fig. 2). The irrigation controller is programmed to control (open) selected irrigation valves for irrigation when the water level decreases below a set point (Belayneh et al., 2013; Van Iersel et al., 2013; Bayer et al., 2015; Wheeler et al., 2017). The irrigation is then allowed to continue until the upper set point of the programme is reached, at which time the irrigation valves are closed (Fig. 2).

Irrigation in small containers may be scheduled using two properties, VWC or soil water potential, as described in detail by Jones (2004; 2008). The VWC indicates the quantity of water available per unit volume of medium, which is easy to measure and control as long as the lower limit (LL) and drained upper limit (DUL) thresholds are defined (Smith

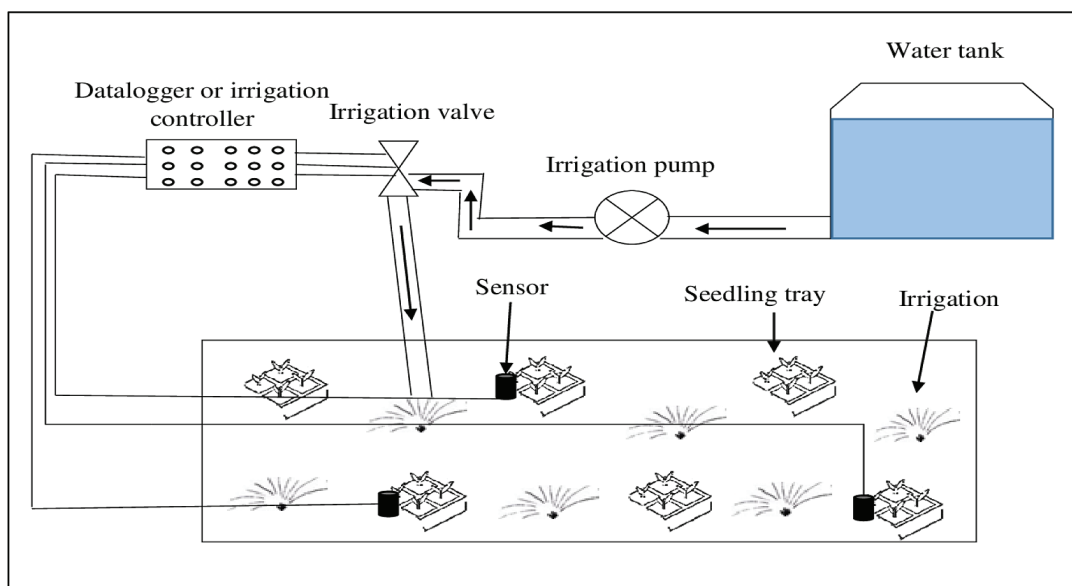


Figure 2. An automated irrigation system capable of measuring the water content of the substrate in the small-volume nursery containers, and controlling the opening and closing of irrigation valves according to lower and upper set points

and Mullins, 2000; Jones, 2004; Lal and Shuckla, 2004; Gebregiorgis and Savage, 2006). In contrast, the soil water potential is the energy status of water per unit volume of medium, which directly determines whether the substrate water is available to plants (Scanlon et al., 2002; Jones, 2004; Bittelli, 2010; Lea-Cox et al., 2011). Both are related to each other, but this relationship is different for different substrates and depends on pore size distribution (Gebregiorgis and Savage, 2006; Van Iersel et al., 2013). For example, substrates with larger pores hold less water compared to substrates with small pores. This relationship can be determined using a hydraulic conductivity relationship (Van Iersel et al., 2013; Schindler et al., 2016). There is no consensus as to which property is better suited for scheduling irrigation between the two (Van Iersel et al., 2013), although VWC is relatively easy to measure (Jones, 2004). If VWC is used, matric potential should also be estimated using hydraulic conductivity to determine the plant available water. This is critical in highly porous substrates such as pine bark due to its lower buffering capacity. Various VWC and soil water potential sensors that may be used in small-volume containers, and easily connected to a programmable control unit, will be discussed.

VOLUMETRIC WATER CONTENT MEASUREMENT SENSORS

The main aim of measuring VWC is to monitor water status as it diminishes and recharges within the root zone after each irrigation event (Charlesworth, 2005). The commercially available soil water content sensors do not measure soil water content directly; instead they detect changes in other soil properties that can be related to the soil water content (Charlesworth, 2005; Bittelli, 2010; 2011). The most common soil properties that change with soil water content and are easy to measure include (i) dielectric permittivity (Van Iersel et al., 2013); and (ii) thermal conductivity (Song et al., 1998). However, for accurate measurement, sensors need to be calibrated against the gravimetric method, which is the only direct soil water content measure (Smith and Mullins, 2000; Charlesworth, 2005).

Dielectric sensors

The term 'dielectric' refers to the ability of a substance to store charge from an electromagnetic field (Evelt and Parkin, 2005; Van Iersel et al., 2013). These sensors use an electromagnetic technique to determine the VWC, since water has a much higher dielectric permittivity relative to other constituents of the substrate (Smith and Mullins, 2000; Bogena et al., 2007; Bittelli, 2010). The main components of a substrate that affect the dielectric permittivity are air, solid matrix and water content. The dielectric permittivity of air and a solid matrix is 1 and 5, respectively, compared to that of water which is 80 at 20°C (Czarnomski et al., 2005; Bogena et al., 2007; Nemali et al., 2007). Hence, a small change in VWC can result in a significant change in the dielectric permittivity (Nemali et al., 2007; Van Iersel et al., 2013). Thus, the measurement of dielectric permittivity can be used to indirectly estimate the media VWC through predetermined calibration relationships using the equation of Topp et al. (1980). For this review, frequency domain reflectometry (FDR) and time domain reflectometry (TDR) sensors will be further discussed due to their suitability for use in small-volume nursery containers.

Frequency domain reflectometry

Most FDR sensors (also referred to as capacitance sensors) consist of two prongs (positive and negative electrodes) that produce an electromagnetic field (Bittelli, 2010) when placed parallel to each other (as shown in Fig. 3). The electromagnetic field is passed through dielectric material and then its ability to store charge is measured. The charge stored by the substrate and measured by a capacitor is directly related to the dielectric permittivity of a substrate (Bogena et al., 2007). The sensor circuitry then converts the capacitor charge to a voltage, so it can be measured using a control unit. Since water molecules store charge more than other particles in the substrate, this charge storing ability can be related to VWC through measurement of charge time t (s) using:

$$t = -RC \ln \left[\frac{(V - V_f + V_i)}{(V_i - V_f)} \right] \quad (1)$$

where R (Ω) is the series resistance, C (μF) the capacitance, V (V) the supply voltage, V_i the initial voltage and V_f the final voltage (Charlesworth, 2005).

When the VWC is high, the capacitor will charge slowly. This means that the capacitor of a sensor embedded in a wet substrate will reach a given voltage threshold later compared to a capacitor in a dry substrate. More specifications, advantages and disadvantages of FDR sensors are provided in Table 1.

The FDR sensors are laboratory-calibrated by the manufacturer in different soilless substrates to produce a generic equation. The factory generic calibration equation has an accuracy range of 0.03–0.05 $\text{m}^3\text{-m}^{-3}$ (Meter Group, 2018a, 2018b). Although a generic equation may be used with reasonable accuracy, a substrate-specific calibration may be conducted to estimate VWC to within 0.01 $\text{m}^3\text{-m}^{-3}$ of the actual VWC. For substrate-specific calibration, a single calibration equation could be used for a similar model FDR sensor regardless of the substrate type (organic or inorganic) (Nemali et al., 2007). Some of these sensors consist of additional sensors that measure soil temperature and electrical conductivity (EC) within the FDR sensor (Meter Group, 2018a, 2018b). These sensors have been used successfully to schedule irrigation in several studies with highly porous soilless substrates in nursery containers (Burnett and Van Iersel, 2008; Van Iersel et al., 2010; Belayneh et al., 2013; Bayer et al., 2015; Montesano et al.,

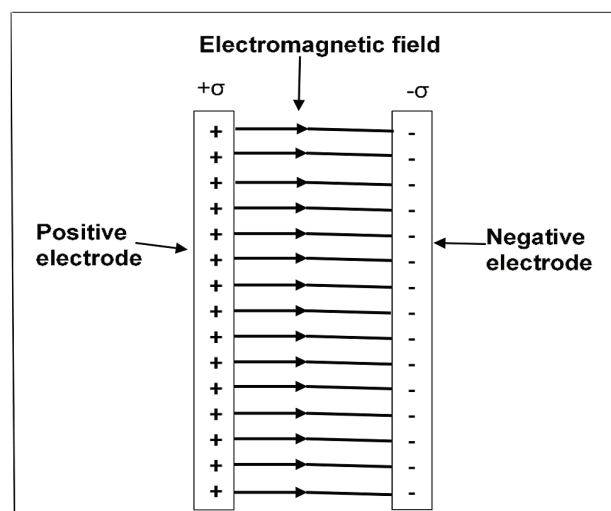


Figure 3. An electromagnetic field formed with a frequency domain reflectometry sensor (source: Campbell, 2012)

2016; Wheeler et al., 2017; Kaptein et al., 2019). Improvements in plant growth, reduction in water usage, and fewer incidences of pests and diseases were reported.

Time domain reflectometry

The TDR technology was initially used to find the break in power lines. Instead of manually inspecting a power line, a pulse is sent along the cable, and the time taken for the pulse to reflect indicates the precise point where the cable is broken. A TDR sensor works similarly by sending electromagnetic waves from the pulse generator of a cable tester to diffuse in the substrate where there is a parallel pair transmission line (Topp et al., 1984; Nobori, 1996). Electromagnetic waves are diffused through a coaxial cable to a probe inserted in a substrate. Some of these electromagnetic waves are reflected at the beginning of the probe due to impedance differences between the cable and the probe, whereas the rest of the waves diffuse through the probe until they reach the end of the probe where they are reflected (Noborio et al., 1996). Therefore, dielectric permittivity can be calculated considering that the transmission velocity is determined from the known length of the transmission line in the substrate using (Topp et al., 1984):

$$K_a = ct / l \quad (2)$$

where K_a is the soil dielectric constant, c the velocity of an electromagnetic signal in free space (i.e. speed of light, $3 \times 10^8 \text{ m}\cdot\text{s}^{-1}$), t the travel time of the voltage pulse and l the length of the soil transmission line (mm). Since substrate water content is the main factor that alters the dielectric permittivity (K_a), VWC can be calculated using (Topp et al., 1980):

$$\text{VWC} = -5.3 \times 10^{-2} + 2.92 \times 10^{-2} K_a - 5.5 \times 10^{-4} K_a^2 + 4.3 \times 10^{-6} K_a^3 \quad (3)$$

The specifications, advantages and disadvantages of TDR sensors are further outlined in Table 1.

Influence of substrate properties on dielectric permittivity

The dielectric permittivity may be influenced by factors other than the VWC (Smith and Mullins, 2000; Lukanu and Savage, 2006). Studies have indicated that dielectric sensors may be affected by the substrate temperature and EC (Scanlon et al., 2002; Bogena et al., 2007; Bittelli, 2010; 2011). However, the new generation of VWC sensors have a built-in temperature sensor that corrects for error in VWC estimation due to temperature changes, for example, a model 5TM or TEROS 10 (Meter Group, Inc., Pullman, WA, USA).

In a nursery, most substrates used as potting media are a unique blend of organic and inorganic materials, and seedlings are frequently irrigated with water-soluble fertilisers. This can result in an increased concentration of ions of fertiliser salts near the electromagnetic field of a dielectric probe which could weaken the electromagnetic energy and affect the measured VWC (Nemali et al., 2007; Rosenbaum et al., 2010). In a study by Kizito et al. (2008), a measurement frequency of 70 MHz mitigated the effects of EC and temperature on FDR sensors (model EC-5 and ECH₂O-TE, Meter Group). Latest FDR sensors measure dielectric permittivity using an oscillator operating at a frequency of 70 MHz making this sensor insensitive to EC below $10 \text{ dS}\cdot\text{m}^{-1}$ (Meter Group, 2018a, 2018b). To be able to measure in substrates with EC greater than $10 \text{ dS}\cdot\text{m}^{-1}$, substrate-specific calibration is a necessity, although such high EC is extremely uncommon in nursery substrates since it will most likely cause root death in most plant species. Increasing measurement frequency to higher frequencies (i.e. 150 MHz) decreases the sensor sensitivity to EC considerably; however, this increases sensor cost due to the increase in sensor electronics.

The effect of poor substrate-to-sensor contact caused by uneven packing of substrate in containers has been reported by Bogena et al. (2007), Nemali et al. (2007) and Van Iersel et al. (2010). This may be critical in some nurseries where the substrate is manually packed in small-volume containers

Table 1. Summary of techniques used to measure substrate water content (adapted from Jones, 2004)

	Frequency domain reflectometry (Capacitance)	Time domain reflectometry	Dual needle heat pulse
Cost (amounts as at Nov 2018: 1 USD = 14.30 ZAR) per sensor	2 200–3 550 ZAR	2 860 ZAR	3 490 ZAR
Accuracy ($\text{m}^3\cdot\text{m}^{-3}$)	0.01–0.03	0.02–0.05	0.05
Measurement range (%)	0–100	0–100	0–100
Measurement volume (mL)	240–715	100	50
Examples (model)	EC-5 ^a , 5TM ^a , GS1 ^a , TEROS 10 ^a	T-3 (T3R/F; mini-TDR) ^b	Specific heat capacity sensor ^b
Advantages	No calibration needed (3% VWC accuracy) Insensitive to salinity Simple readout device Inexpensive Low power usage Easy to install	No calibration needed Less sensitive to salinity $<3 \text{ dS}\cdot\text{m}^{-1}$	Compact size No calibration needed
Disadvantages	Sensitive to air gaps Need substrate-specific calibration for improved accuracy (1% VWC accuracy)	Expensive when pulse is included Their complexity requires expertise to set-up High power usage	Fragile Susceptible to substrate temperature gradients Needle deflection can impart high error (1 mm deflection = 6% error) Must be connected to good dataloggers

^a Meter Group, Inc., Pullman, WA, USA^b East 30 Sensors, Pullman, WA, USA

which may lead to a lack of solid contact between the sensor and the substrate causing inaccurate measurements of VWC. According to Lukanu and Savage (2006), differences in bulk density can have a small effect on FDR sensor measurements although VWC can be measured to within $0.02 \text{ m}^3 \cdot \text{m}^{-3}$. Van Iersel et al. (2013) reported that measurement errors due to air spaces can be improved by frequent irrigation schedules until substrate particles settle around the sensor.

Dual needle heat pulse sensors

The dual needle heat pulse (DNHP) method was first suggested by Campbell et al. (1991) and has since been used by several researchers (Tarara and Ham, 1997; Song et al., 1998; Ochsner et al., 2003). This technique measures changes in soil thermal properties (Song et al., 1998; Bittelli, 2010; 2011) caused by the variation in the VWC. Tarara and Ham (1997) describe the technique in detail, and Fig. 4 shows a probe with needle length of 30 mm. A heater and a temperature probe (usually a thermocouple) are used to determine the substrate volumetric heat capacity, which can be converted into VWC using:

$$VWC = [q / (\epsilon \pi r^2 \Delta T_m - (19.92x_m + 2.51))] / 4.18 \quad (4)$$

where q ($\text{J} \cdot \text{m}^{-1}$) is the heat applied per unit length of the line source, e is the natural logarithm base, r (m) is the distance between the heater probe and temperature probe, ΔT_m is the maximum temperature increase of the needle ($^{\circ}\text{C}$) and x_m is determined by dividing the substrate bulk density by particle size (Song et al., 1998).

The small compact size of the DNHP sensor enables measurements of small substrate volumes such as water content around a growing seed (Tarara and Ham, 1997). Bristow et al. (1993) compared the VWC as measured using DNHP sensors versus the VWC measured by gamma attenuation in repacked soil in the laboratory. The DNHP sensors estimated VWC to within $0.04 \text{ m}^3 \cdot \text{m}^{-3}$ of the gamma attenuation method at both the LL and DUL. However, large differences between VWC measured with DNHP sensors versus using the gravimetric method were observed by Tarara and Ham (1997) at high water content. At water contents less than $0.30 \text{ m}^3 \cdot \text{m}^{-3}$, the DNHP probe estimates were within $0.06 \text{ m}^3 \cdot \text{m}^{-3}$ of the gravimetric method. The probes of this sensor type are very fragile and special care needs to be taken so that the distance between the needle probes does not alter (Campbell et al., 1991; Bittelli, 2010). A needle deflection of 1 mm may cause a 6% error in VWC measurements (Song et al., 1998; Scanlon et al., 2002). The conversion of a temperature signal from analogue to digital may cause a measurement error in ΔT_m (Bristow et al., 1993;

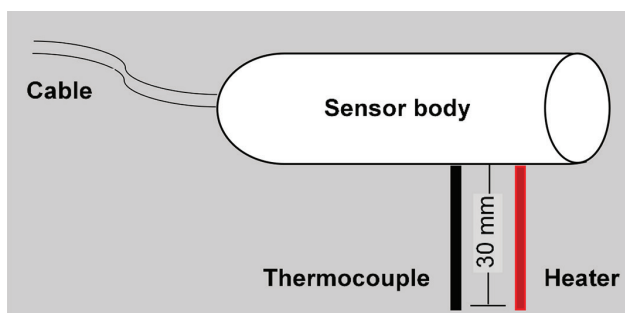


Figure 4. A dual needle heat pulse probe (adapted from: Tarara and Ham, 1997)

Song et al., 1998). This, however, can be minimised by increasing the datalogger sampling frequency by using a high resolution datalogger and applying sufficient power to the heater probe so that ΔT_m exceeds 0.5°C . This sensor is also susceptible to substrate temperature gradients (Scanlon et al., 2002), which may be significant in a nursery environment, and therefore precise substrate temperature measurements are needed. For these reasons, this type of sensor should be used with caution in scheduling nursery irrigation, and, in most cases, is not practical.

MEASURING THE SUBSTRATE WATER POTENTIAL

The concept of soil water potential is not new. As early as 1908, the first attempts to measure soil water potential were conducted using the Livingston disc (Crawford, 2015). A dry ceramic disc would be weighed, then placed in soil so that it equilibrated with the surroundings. After equilibration, it was removed, cleaned and weighed again. Soil water potential was then calculated using the water retention curve of the disc. Over 100 years later, there have been many advances in soil water potential sensor technology, but these sensors still rely on water potential equilibration. Most commercially available soil water potential sensors accurately measure within the 0 to -100 kPa range, which is ideal for most nursery crops since they generally exhibit signs of water stress at matric potential less than -100 kPa. Different techniques for measuring soil water potential in small-volume seedling containers are examined and summarised in Table 2.

Tensiometers

In 1960, a liquid calibration technique called tensiometry was discovered (Charlesworth, 2005). This method, along with soil psychrometry, was and still is the only direct method of measuring soil water potential. If correctly installed, a tensiometer is the most accurate soil water potential sensor in wet substrate (0 to -90 kPa) (Charlesworth, 2005; Bittelli, 2010). For an in-depth understanding of tensiometer measurement theory, see Smith and Mullins (2000) and Bittelli (2010). The advantages and disadvantages of tensiometers are outlined in Table 2. Small tensiometers fitted with transducers to measure and control the water potential in small-volume soilless substrates were reported by Bittelli (2010). A schematic diagram of such a tensiometer is shown in Fig. 5, with a diameter of 5 mm for an easy fit in a small-volume container. The transducer is used to convert the tensiometer tension to an electrical signal that may be sensed by a control unit.

When tensiometers are used, precautionary measures are necessary to ensure direct contact between the porous ceramic tensiometer tip and the substrate, and correct positioning of the sensor. If the substrate shrinks, or the tensiometer is disturbed, this contact may be interrupted after which air may enter and break the water column resulting in inaccurate measurements. In soilless substrates, fine roots might develop around the ceramic tip. This is common when the substrate dries out and the ceramic tip becomes an unreliable but assured water source (Meter Group, 2009). When the substrate becomes drier than -90 kPa, cavitation occurs, where liquid water pressure inside the tensiometer tube changes to water vapour pressure causing spontaneous evaporation and formation of air bubbles as described by Bittelli (2010). However, the new commercially available tensiometers can measure up to -500 kPa without cavitation, for example, a model T5x mini tensiometer (Meter Group). This is made

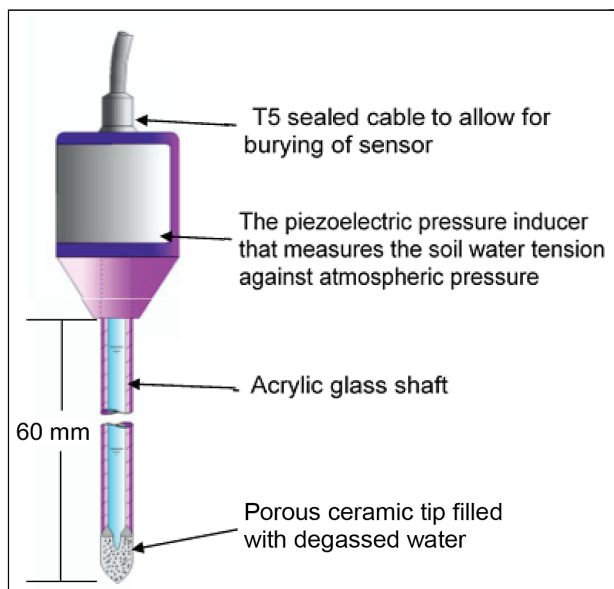


Figure 5. A transducer fitted mini tensiometer (Meter Group model T-5 laboratory tensiometer) (Source: Meter Group)

possible using a special ceramic tip with smaller pores and a gas-free filling. Van der Ploeg et al. (2010) also investigated polymer tensiometers that are not affected by cavitation and have an extended measurement range to $-2\,000$ kPa. These tensiometers consist of a swelling polymer that creates a positive pressure, and when the substrate is wet it absorbs water, creating a pressure offset. These tensiometers are, however, not yet well researched for use in nursery soilless substrates. Charlesworth (2005) reported on equitensiometers that use the principle of dielectric sensors to measure the water potential. These sensors do not require substrate-specific calibration since they measure the water potential of the ceramic material and not the surrounding substrate.

Heat dissipation sensors

As early as 1940, Shaw and Baver (1940) proved that soil water

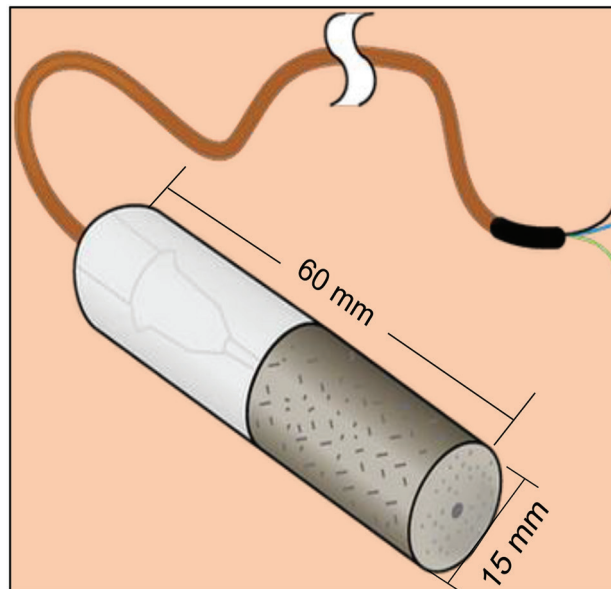


Figure 6. A heat dissipation sensor (Campbell Scientific, model 229-L) (Source: Campbell Scientific, 2009)

potential could be inferred from the rate of heat dissipated in the soil, and many sensors have been developed based on this relationship. Flint et al. (2002) described heat dissipation probe measurement theory in detail and a sensor is shown in Fig. 6. The sensor has a ceramic cup which equilibrates with the substrate. Following equilibration, a heating element is heated for a specific period, and a change in the substrate temperature is then measured. The temperature change depends on the thermal conductivity of the ceramic cup, which is affected by the VWC of the substrate (Flint et al., 2002; Scanlon et al., 2002; Bittelli, 2010). Thus, the VWC, which is related to the water potential of the ceramic cup, can be determined through a hydraulic conductivity relationship. The commonly used function is:

$$\Psi = \exp(\alpha \Delta T + \beta) \quad (5)$$

where Ψ is the soil water potential, \exp the exponential function, ΔT the increase in temperature for specified period of time, α the

Table 2. Summary of techniques used to measure substrate matric potential (adapted from Jones, 2004)

	Water potential	Tensiometers	Heat dissipation
Cost (amounts as at Nov 2018: 1 USD = 14.30 ZAR) per sensor	4 150 ZAR	11 390 ZAR	7 000 ZAR
Accuracy (kPa)	± 2 below -100	± 0.5	3–5
Measurement range (kPa)	$-100\,000$ to -9	-500 to 0	$-2\,500$ to -10
Equilibration time	10 min to 1 h	5 s	1 h
Measurement volume (mL)	1	> 100	1
Examples (model)	TEROS 21 ^a	T-5x ^a	229-L ^b
Advantages	Moderate to good accuracy Wide measurement range Insensitive to salinity below $10\text{ dS}\cdot\text{m}^{-1}$ Different sizes available	Direct measure Excellent accuracy between -90 to 0 kPa	No maintenance needed Insensitive to salinity Moderate to good accuracy
Disadvantages	Air entry in ceramic discs limit accuracy Sensitive to air gaps Low accuracy between -100 and 0 kPa	Need re-filling with degassed water after dry periods Air entry in ceramic tip limit upper range measurements High maintenance requirements in porous substrates	Need complex heating controller Slow reaction time High power usage High variability between sensors

^a Meter Group, Inc., Pullman, WA, USA ^b Campbell Scientific, Inc., Logan, Utah, USA

slope and β the intercept (Campbell Scientific, 2009). The relationship between the natural logarithm of soil water potential and temperature increase is linear (Campbell Scientific, 2009). The heat transfer properties between the heater and the ceramic cup of the sensor vary vastly between sensors, thus necessitating individual sensor calibration (Flint et al., 2002; Scanlon et al., 2002). These sensors can be used for scheduling irrigation in a laboratory and in greenhouses. However, due to their complexity, their use has been limited to research applications.

IDEAL SENSORS FOR USE WITH A NURSERY AUTOMATED IRRIGATION SYSTEM

The first step in selecting a sensor to use with an irrigation system is to identify the substrate property to measure, either VWC or water potential. There are a variety of commercially available sensors that may be used to measure either of these properties, as described previously. Soil water potential sensors should, however, be used with caution in small volume nursery containers. Several studies have indicated that these sensors are less accurate at low soil water potentials (< -30 kPa), which is when most seedlings show signs of water stress. Mini-tensiometers require high maintenance and are also less reliable in automated irrigation systems.

An ideal sensor should provide high accuracy and resolution, enable rapid and precise measurements, be easy to install, allow for custom calibration, have a long life expectancy, and be cost effective (Gebregiorgis and Savage, 2006; Jones, 2008). Cost (both purchase and maintenance), life expectancy, ease of use and media-specific calibration are key aspects to the adoption and use of sensor-based automated irrigation by nursery managers (Lea-Cox, 2012).

Cost-effectiveness

A typical seedling nursery consists of different seedling types at different ages. For example, a forestry nursery may have different seedling clones ranging from 1-day-old to 3-months-old. To control irrigation, multiple sensors may be required for different irrigation zones. In this case, many inexpensive sensors may be an option. In addition, multiple sensors may be needed to improve the quality of collected data. As an example, low-cost FDR sensors (model EC-5, Meter Group) have been successfully used in soilless substrates by Montesano et al. (2016), Saavoss et al. (2016), Wheeler et al. (2017) and Kaptein et al. (2019) due to their cost-effectiveness and ease of use. The challenge is selecting trade-offs between various sensor properties. For example, durable high-accuracy sensors with good precision and resolution may be more expensive than less durable lower-accuracy sensors. Sensor manufacturers may attempt to find the balance between these properties, but to select the best sensor for an application, it is important to ensure that minimum requirements for a particular purpose are met (Lea-Cox, 2012; Van Iersel et al., 2013). It may be necessary for a sensor used in a research study to have high accuracy, good precision and resolution, whereas a sensor used for scheduling irrigation may not necessarily need to have the same level of accuracy, as long as measurements are repeatable with good resolution and precision.

Specific media calibration

Most sensors are supplied with a generic calibration equation. However, not all soilless substrates used in nurseries have identical electrical properties due to variations in bulk

density and salinity. Thus, the generic calibration may result in an approximately 5% error in VWC estimation, whereas a substrate-specific calibration can reduce this error to 1%. Furthermore, different soilless substrates have different water-holding capabilities (Van Iersel et al., 2013), and it is important to understand the relationship between VWC and plant available water for different substrates. For example, a 20% VWC measurement in sand indicates a moist to saturated substrate and much of this water is available to seedlings. Irrigating this substrate further will simply result in excess water drainage. However, a value of 20% in coir indicates a much drier medium, meaning there is little or no plant available water in the substrate (Van Iersel et al., 2013).

Manufacturers commonly calibrate their sensors in the laboratory with minimal variation in environmental factors, in particular temperature. When the sensor is in a natural environment and exposed to more environmental variation, its accuracy may not meet the specified manufacturer specifications (Van Iersel et al., 2013). The effect of temperature is more pronounced in sensors such as DNHP and TDR. However, most of the new model commercially available sensors have a built-in temperature and EC sensor to correct for substrate temperature effects. Although media-specific calibration improves VWC measurement accuracy, generic calibration could be used for irrigation scheduling purposes with slightly less accuracy (Kaptein et al., 2019). Sensor-to-sensor variability has been shown to be relatively low within soilless substrates, even with influencing factors such as air gaps and bulk density (Belayneh et al., 2013; Kaptein et al., 2019). Also, it is beneficial to install several sensors and use their average output for scheduling irrigation decisions.

All sensors have measurement errors and there is no one sensor that is ideal. The best sensor for an application is the one that measures what the end-user needs. There are a variety of good commercially available sensors, and if the user understands what the sensor measures, then it becomes easier to compare and make the best choice for the application. Also, most sensors can be integrated into an existing nursery irrigation system, ensuring economic feasibility. For small-volume soil water measurements in nursery applications, sensor size and substrate-to-sensor contact are the most important considerations in this environment in which large temperature gradients may occur. This then limits sensor choice to a few capacitance sensors.

Potential economic impact

The use of a sensor-based automated irrigation system may have a positive impact within a nursery in different ways, such as improved water use efficiency, reduction in water and energy costs, labour savings, improved crop quality and shorter crop production cycles, as reported in detail by Belayneh et al. (2013). A reduction in irrigation water usage by 40–70% was reported by Bayer et al. (2013) and Chappell et al. (2013). This resulted in reduction in leaching of nutrients, thereby reducing the nursery fertiliser costs by 50%. Overall, in a study by Saavoss et al. (2016) a sensor-based irrigation system showed an increase in nursery annual profits by 156% compared to a standard timer-based system. This profit was mainly attributed to improved plant growth, improved crop quality and shorter production cycles.

CONCLUSIONS

In this review, different techniques for measuring VWC and water potential that are appropriate for use in small-volume

nursery containers were described. Many researchers have concluded that the sensor-based irrigation systems are more efficient and effective compared to timer-based systems. Also, the quality of seedlings produced using a sensor-based irrigation system is generally better than that from conventional timer systems. To select an optimal sensor for use in small-volume seedling containers, sensor cost-effectiveness and the ability to calibrate against the gravimetric method should be prioritised. Moreover, the sensor should meet the minimum requirements for an application. Utilising sensor-based irrigation systems may benefit the commercial nursery industry and reduce the high water consumption, thereby improving overall water management.

ACKNOWLEDGEMENTS

Funding from the Institute for Commercial Forestry Research (ICFR) is gratefully acknowledged. Many thanks also go to Professor Colin Dyer and Dr Andrew Morris of the ICFR for supporting this research. Dr Louis Titshall and Dr Steven Dovey are thanked for their constructive input.

REFERENCES

- ANNANDALE JG, STIRZAKER RJ, SINGELS A, VAN DER LAAN M and LAKER MC (2011) Irrigation scheduling research: South African experiences and future prospects. *Water SA* **37** 751–764. <https://doi.org/10.4314/wsa.v37i5.12>
- BAYER A, RUTER J and VAN IERSEL MW (2015) Automated irrigation control for improved growth and quality of *Gardenia jasminoides* ‘Radicans’ and ‘August Beauty’. *HortScience* **50** 78–84.
- BAYER A, WHITAKER K, CHAPPELL M, RUTER J and VAN IERSEL M (2013) Effect of irrigation duration and fertilizer rate on plant growth, substrate solution EC, and leaching volume. *Acta Hort.* **1034** 477–484. <https://doi.org/10.17660/ActaHortic.2014.1034.59>
- BELAYNEH BE, LEA-COX JD and LICHTENBERG E (2013) Costs and benefits of implementing sensor-controlled irrigation in a commercial pot-in-pot container nursery. *HortTechnology* **23** 760–769.
- BITTELLI M (2010) Measuring soil water potential for water management in agriculture: A review. *Sustainability* **2** 1226–1251. <https://doi.org/10.3390/su2051226>
- BITTELLI M (2011) Measuring soil water content: A review. *HortTechnology* **21** 293–300.
- BOGENA HR, HUISMAN JA, OBERDÖRSTER C and VEREECKEN H (2007) Evaluation of a low-cost soil water content sensor for wireless network applications. *J. Hydrol.* **344** 32–42. <https://doi.org/10.1016/j.jhydrol.2007.06.032>
- BRISTOW KL, CAMPBELL GS and CALISSENDORFF K (1993) Test of a heat-pulse probe for measuring changes in soil water content. *Soil Sci. Soc. Am. J.* **57** 930–934. <https://doi.org/10.2136/sssaj1993.03615995005700040008x>
- BURNETT SE and VAN IERSEL M (2008) Morphology and irrigation efficiency of *Gaura lindheimeri* grown with capacitance sensor-controlled irrigation. *HortScience* **43** 1555–1560.
- CAMPBELL GS, CALISSENDORFF C and WILLIAMS JH (1991) Probe for measuring soil specific heat using a heat-pulse method. *Soil Sci. Soc. Am. J.* **55** 291–293. <https://doi.org/10.2136/sssaj1991.03615995005500010052x>
- CAMPBELL SCIENTIFIC (2009) 229 Heat dissipation matrix water potential sensor manual. URL: <https://s.campbellsci.com/documents/br/manuals/229.pdf> (Accessed 23 October 2018).
- CAMPBELL C (2012) Soil moisture 201: Water content measurements, methods, and applications theory and application. URL: <https://www.metergroup.com/environment/events/soil-moisture-201-measurements-methods-and-applications/> (Accessed 1 December 2018).
- CHAPPELL M, DOVE SK, VAN IERSEL MW, THOMAS PA and RUTER J (2013) Implementation of wireless sensor networks for irrigation control in three container nurseries. *HortTechnology* **23** 747–753.
- CHARLESWORTH P (2005) Soil water monitoring. Irrigation Insights No. 1, Second Edition. URL: <http://www.insidecotton.com/xmlui/bitstream/handle/1/1726/pr000236.pdf> (Accessed 17 October 2018).
- CRAWFORD L (2015) The history and future of water potential. URL: <http://www.environmentalbiophysics.org/the-history-and-future-of-water-potential/> (Accessed 16 August 2014).
- CZARNOMSKI NM, MOORE GW, PYPKER TG, LICATA J and BOND BJ (2005) Precision and accuracy of three alternative instruments for measuring soil water content in two forest soils of the Pacific Northwest. *Can. J. For. Res.* **35** 1867–1876. <https://doi.org/10.1139/x05-121>
- DURNER EF (2013) *Principles of Horticultural Physiology*. CABI Publishers, London. 416 pp.
- EVETT SR and PARKIN GW (2005) Advances in soil water content sensing. *Vadose Zone J.* **4** 986–991. <https://doi.org/10.2136/vzj2005.0099>
- FANADZO M and NCUBE B (2017) Challenges and opportunities for revitalising smallholder irrigation schemes in South Africa. *Water SA* **44** 436–447. <https://doi.org/10.4314/wsa.v44i3.11>
- FLINT AL, CAMPBELL GS, ELLET KM and CALISSENDORFF C (2002) Calibration and temperature correction of heat dissipation matrix potential sensors. *Soil Sci. Soc. Am. J.* **66** 1439–1445. <https://doi.org/10.2136/sssaj2002.1439>
- GEBREGIORGIS MF and SAVAGE MJ (2006) Field, laboratory and estimated soil-water content limits. *Water SA* **32** 155–162. <https://doi.org/10.4314/wsa.v32i2.5256>
- GLEESON T, ALLEY WM, ALLEN DM, SOPHOCLEOUS MA, ZHOU Y, TANIGUCHI M and VANDERSTEEN J (2012) Towards sustainable groundwater use: Setting long-term goals, backcasting, and managing adaptively. *Groundwater* **50** 19–26. <https://doi.org/10.1111/j.1745-6584.2011.00825.x>
- JONES HG (2004) Irrigation scheduling: advantages and pitfalls of plant-based methods. *J. Exp. Bot.* **55** 2427–2436. <https://doi.org/10.1093/jxb/erh213>
- JONES HG (2008) Irrigation scheduling – comparison of soil, plant and atmosphere monitoring approaches. *Acta Hort.* **792** 391–403. <https://doi.org/10.17660/ActaHortic.2008.792.46>
- KAPTEIN ND, SAVAGE MJ and LIGHT ME (2019) An irrigation control system with a web-based interface for the management of *Eucalyptus* planting stock in a nursery. *South. For.* **81** 31–37. <https://doi.org/10.2989/20702620.2018.1479922>
- KIZITO F, CAMPBELL CS, CAMPBELL GS, COBOS DR, TEARE BL, CARTER B and HOPMANS JW (2008) Frequency, electrical conductivity and temperature analysis of a low-cost capacitance soil moisture sensor. *J. Hydrol.* **352** 367–378. <https://doi.org/10.1016/j.jhydrol.2008.01.021>
- LAL R and SHUCKLA MK (2004) *Principles of Soil Physics*. Taylor and Francis Library, New York. 736 pp.
- LEA-COX JD, ARGUEDAS-RODRIGUEZ FR, RISTVEY AG and ROSS DS (2011) Relating real-time substrate matrix potential measurements to plant water use for precision irrigation. *Acta Hort.* **891** 201–208. <https://doi.org/10.17660/ActaHortic.2011.891.23>
- LEA-COX JD (2012) Using wireless sensor networks for precision irrigation scheduling. In: Kumar M (ed.) *Problems, Perspectives and Challenges of Agricultural Water Management*. IntechOpen, Croatia. <https://doi.org/10.5772/31236>
- LEA-COX JD, BELAYNEH BE, MAJSZTRIK J, RISTVEY AG, LICHTENBERG E, VAN IERSEL MW, CHAPPELL M, BAUERLE WL, KANTOR G, KOHANBASH D and co-authors (2017) Demonstrated benefits of using sensor networks for automated irrigation control in nursery and greenhouse production systems. *Acta Hort.* **1150** 507–514. <https://doi.org/10.17660/ActaHortic.2017.1150.70>
- LICHTENBERG E, MAJSZTRIK J and SAAVOSS M (2013) Profitability of sensor-based irrigation in greenhouse and nursery crops. *HortTechnology* **23** 770–774.
- LUKANU G and SAVAGE MJ (2006) Calibration of a frequency-domain reflectometer for determining soil-water content in a clay loam soil. *Water SA* **32** 37–42.
- METER GROUP (2009) User Manual: T5/T5x Pressure Transducer

<https://doi.org/10.17159/wsa/2019.v45.i3.6750>

Available at <https://www.watersa.net>

ISSN 1816-7950 (Online) = Water SA Vol. 45 No. 3 July 2019

Published under a Creative Commons Attribution 4.0 International Licence (CC BY 4.0)

- Tensiometer. URL: http://library.metergroup.com/Manuals/UMS/T5_Manual.pdf (Accessed: 20 October 2018).
- METER GROUP (2018a) 5TM Manual. URL: http://publications.metergroup.com/Manuals/20424_5TM_Manual_Web.pdf (Accessed 22 October 2018).
- METER GROUP (2018b) EC5 Manual. URL: http://publications.metergroup.com/Manuals/20431_EC-5_Manual_Web.pdf (Accessed 22 October 2018).
- MONTESANO FF, VAN IERSEL MW and PARENTE A (2016) Timer versus moisture sensor-based irrigation control of soilless lettuce: Effects on yield, quality and water use efficiency. *Hortic. Sci. (Prague)* **43** 67–75. <https://doi.org/10.17221/312/2014-HORTSCI>
- NEMALI KS, MONTESANO F, DOVE SK and VAN IERSEL MW (2007) Calibration and performance of moisture sensors in soilless substrate: ECH₂O and Theta probes. *Sci Hortic.* **112** 227–234. <https://doi.org/10.1016/j.scienta.2006.12.013>
- NEMALI KS and VAN IERSEL MW (2006) An automated system for controlling drought stress and irrigation in potted plants. *Sci. Hortic.* **110** 292–297. <https://doi.org/10.1016/j.scienta.2006.07.009>
- NOBORIO K, MCINNIS KJ and HEILMAN JL (1996) Measurements of soil water content, heat capacity, and thermal conductivity with a single TDR probe. *Soil Sci.* **161** 22–28. <https://doi.org/10.1097/00010694-199601000-00004>
- OCHSNER TE, HORTON R and REN T (2003) Use of the dual-probe heat pulse technique to monitor soil water content in the vadose zone. *Vadose Zone J.* **2** 572–579. <https://doi.org/10.2136/vzj2003.05720>
- PIMENTEL D, BERGER B, FILIBERTO D, NEWTON M, WOLFE B, KARABINAKIS E, CLARK S, POON E, ABBETT E and NANDAGOPAL S (2004) Water resources: Agricultural and environmental issues. *BioScience* **54** 909–918. [https://doi.org/10.1641/0006-3568\(2004\)054\[0909:WRAAEI\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2004)054[0909:WRAAEI]2.0.CO;2)
- ROSENBAUM U, HUISMAN JA, VRBA J, VERECKEN H and BOGENA HR (2010) Correction of temperature and electrical conductivity effects on dielectric permittivity measurements with ECH₂O sensors. *Vadose Zone J.* **10** 582–593. <https://doi.org/10.2136/vzj2010.0083>
- SAAVOSS M, MAJSZTRIK J, BELAYNEH B, LEA-COX J and LICHTENBERG E (2016) Yield, quality, and profitability of sensor-controlled irrigation: a case study of snapdragon (*Antirrhinum majus* L.) production. *Irrig. Sci.* **34** 409–420. <https://doi.org/10.1007/s00271-016-0511-y>
- SCANLON BR, ANDRASKI BJ and BILSKIE J (2002) Miscellaneous methods for measuring matric or water potential. In: Dane JH and Topp GC (eds.) *Methods of Soil Analysis, Part 4, Physical Methods*, SSSA Book Series No. 5. Soil Science Society of America, Madison. 643–670.
- SCHINDLER U, MÜLLER L and EULENSTEIN F (2016) Measurement and evaluation of the hydraulic properties of horticultural substrates. *Arch. Agron. Soil Sci.* **62** 806–818. <https://doi.org/10.1080/003650340.2015.1083982>
- SHAW B and BAVER LD (1940) An electrothermal method for following moisture changes of the soil *in situ*. *Soil Sci. Soc. Am. Proc.* **4** 78–83. <https://doi.org/10.2136/sssaj1940.036159950004000C0014x>
- SMITH KA and MULLINS CE (2000) *Soil and Environmental Analysis: Physical Methods* (2nd edn). Marcel Dekker, New York. 637 pp.
- SONG Y, HAM JM, KIRKHAM MB and KLUITENBERG GJ (1998) Measuring soil water content under turfgrass using the dual-probe heat-pulse technique. *J. Am. Soc. Hortic. Sci.* **123** 937–941.
- TARARA JM and HAM JM (1997) Measuring soil water content in the laboratory and field with dual-probe heat-capacity sensors. *Agron. J.* **89** 535–542. <https://doi.org/10.2134/agronj1997.0002196200890040001x>
- TOPP GC, DAVID JL and ANNAN AP (1980) Electromagnetic determination of soil water content: Measurement in coaxial transmission lines. *Water Resour. Res.* **16** 574–582. <https://doi.org/10.1029/WR016i003p00574>
- TOPP GC, ZEBCHUCK WD, DAVIS JL and BAILEY WG (1984) The measurement of soil water content using portable TDR hand probe. *Can. J. Soil Sci.* **64** 313–321. <https://doi.org/10.4141/cjss84-033>
- VAN DER PLOEG MJ, GOOREN HPA, BAKKER G, HOOGENDAM CW, HUISKES C, KOOPAL LK, KRUIDHOF H and ROOIJ GH (2010) Polymer tensiometers with ceramic cones: direct observations of matric pressures in drying soils. *Hydrol. Earth Syst. Sci.* **14** 1787–1799. <https://doi.org/10.5194/hess-14-1787-2010>
- VAN IERSEL MW, DOVE S, KANG J and BURNETT SE (2010) Growth and water use of petunia as affected by the substrate water content and daily light integral. *HortScience* **45** 277–282.
- VAN IERSEL MW, DOVE S and BURNETT SE (2011) The use of soil moisture probes for improved uniformity and irrigation control in greenhouses. *Acta Hortic.* **893** 1049–1056. <https://doi.org/10.17660/ActaHortic.2011.893.119>
- VAN IERSEL MW, CHAPPELL M and LEA-COX JD (2013) Sensors for improved efficiency of irrigation in greenhouse and nursery production. *HortTechnology* **23** 735–746.
- WHEELER WD, WILLIAMS-WOODWARD J, THOMAS PA, VAN IERSEL M and CHAPPELL MR (2017) Impact of substrate volumetric water on *Pythium aphanidermatum* infection in *Petunia xhybrida*: A case study on the use of automated irrigation in phytopathology studies. *Plant Health Prog.* **18** 120–125. <https://doi.org/10.1094/PHP-01-17-0006-RS>