In-situ tryptophan-like fluorescence: A real-time indicator of faecal contamination in drinking water supplies

J.P.R. Sorensen [a, *], D.J. Lapworth [a], B.P. Marchant [b], D.C.W. Nkhuwa [c], S. Pedley [d], M.E. Stuart [a], R.A. Bell [b], M. Chirwa [c], J. Kabika [c], M. Liemisa [e], M. Chibesa [e]

[a] British Geological Survey, Maclean Building, Wallingford, OX10 8BB, UK
[b] British Geological Survey, Nicker Hill, Keyworth, Nottingham, NG12 5GG, UK
[c] University of Zambia, Great East Road Campus, P.O. Box 32379, Lusaka, Zambia
[d] Robens Centre for Public and Environmental Health, University of Surrey, Guildford, GU2 7XH, UK
[e] Lukanga Water and Sewerage Company Limited, P.O. Box 81745, Kabwe, Zambia

ARTICLE INFO

ABSTRACT

Enteric pathogens are typically inferred from the presence of surrogate indicator organisms such as thermotolerant (faecal) coliforms (TTCs). The analysis of TTCs requires time-consuming incubation in suitable laboratories, which can limit sampling resolution, particularly during critical pollution events. Here, we demonstrate the use of in-situ fluorimeters targeting tryptophan-like compounds as a rapid, reagentless indicator of TTCs in groundwater-derived potable water supplies in Africa. A range of other common indicators of TTCs were also determined including nitrate, turbidity, and sanitary risk survey scores. Sampling was conducted during both the dry and wet seasons to investigate seasonality. Tryptophan-like fluorescence was the most effective predictor of both presence/absence and number of TTCs during both seasons. Seasonal changes in tryptophan-like fluorescence in deeper supplies suggest it is transported more efficiently through the aquifer than TTCs. Moreover, the perennial elevated concentrations in some wells suggest it is more resilient than TTCs in groundwater. Therefore tryptophan-like fluorescence could also be a better indicator of some smaller, more easily transported, and long-lived, pathogenic enteric viruses. These sensors have the potential to be included in real-time pollution alert systems for drinking water supplies throughout the world, as well as for mapping enteric pathogen risks in developing regions.

© 2015 Published by Elsevier Ltd.

1. Introduction

Many pathogens transmitted through drinking water are of faecal origin and these continue to pose a threat to human health globally (Ashbolt, 2004). In the USA, up to half of all groundwater supplies have had evidence of faecal contamination resulting in 750,000 to 5.9 million cases of waterborne illnesses per year (Macler and Merkle, 2000). However, these risks are most serious in the developing world where 99.8% of the annual 1.7 million deaths relating to unsafe water supply, sanitation, and hygiene occur (WHO, 2002).

The assessment of enteric pathogens in drinking water has traditionally been inferred using surrogate indicator organisms (Savichtcheva and Okabe, 2006). Currently, the WHO Guidelines for Drinking Water Quality, adopted as standards in many countries, use the indicator group thermotolerant (faecal) coliforms (TTC), or specifically *Escherichia coli*, as a measure of the safety of drinking water supplies. Analysis for these organisms requires well-trained operators working with sterile equipment and reagents in laboratory conditions, which are not always easily accessible. Furthermore, the procedure is time-consuming (>18 h), owing to the necessity for culturing, which can be critical during pollution events when timely intervention and consumer advice is essential.

Fluorescence spectrophotometry offers a potential alternative with multiple studies highlighting its use as a rapid reagentless wastewater indicator (Baker, 2001; Lapworth et al., 2008; Henderson et al., 2009). These observations are based on the positive relationship between tryptophan-like fluorescence (TLF) and labile organic carbon and microbial activity (Cammack et al., 2004; Hudson et al., 2007; Lapworth et al., 2007; Hudson et al., 2008).

Technological advances now mean that portable fluorimeters targeting the TLF peak are commercially available as possible in-situ or roaming field sensors. Cumberland et al. (2012) demonstrated their potential in controlled laboratory conditions: observing tentative positive correlations with a range of indicator organisms, including *E. coli*, in river and effluent samples. Their use in groundwater is likely to be particularly promising given the general low turbidity and reasonably stable temperature, which can both significantly impact detectable fluorescent signals (Baker, 2005; Downing et al., 2012; Khamis et al., 2015).

This study is the first to pilot the use of in-situ TLF fluorimeters for the rapid assessment of the microbiological quality of drinking water supplies, with a focus on groundwater in the developing world. The aims of this work were to investigate the application of a portable fluorimeter to indicate i) whether a groundwater supply was faecally contaminated; ii) the extent of any contamination; and iii) compare TLF as an indicator of contamination against other common indicators such as NO_3, Cl, turbidity, and sanitary risk scores (SRS). Faecal contamination was determined through the presence of TTCs.

2. Materials and methods

2.1. Study site

Kabwe is located in Zambia's Central Province approximately 150 km north of the capital Lusaka. It has a population of over 200,000, with a high proportion residing within informal settlements on the outskirts of the city such as Makululu — regarded as one of the largest slums in southern Africa with an estimated 46,000 inhabitants (LgWSC, 2014). The city is predominantly underlain by several hundred metres of either the Lower Roan Group (quartzite, schist and pelite) or Upper Roan Group (dolomite). The bedrock is concealed beneath continuous saprolite and laterite superficial deposits that are typically 5–20 m thick (Houston, 1982). Groundwater is generally encountered 5–10 m below ground level, with the superficials typically in hydraulic connection with the deeper aquifer within the karstic bedrock. The local climate is sub-tropical with rainfall exhibiting strong seasonality: 95% falls between mid-November and mid-April (Nkhuwa et al., 2006). Natural surface waters are absent, as rainfall rapidly infiltrates into the subsurface.

Groundwater is the major source of drinking water supply for the city. The centralised supply system abstracts groundwater from deep boreholes within peri-urban wellfields, which is then treated and piped directly to properties, or dispensed via communal taps and water kiosks within the informal settlements. Households frequently also self-supply groundwater to some extent as the centralised supply can be unreliable and is charged on a per volume basis. Within informal settlements this is generally restricted to vulnerable shallow hand-dug wells and illegal connections to pre-treated water within the centralised supply network. In more affluent areas, self-supply includes tapping the bedrock or superficials through deeper boreholes or shallow wells, respectively, with limited use of piped supplies.

Low levels of sanitation coverage are a major cause for concern within newer parts of the Kabwe. This includes the burgeoning informal settlements where coverage is estimated at less than 11% of properties and restricted to pit latrines (LgWSC, 2014). In established parts of the city, the sewerage network is more extensive, but is ageing, in need of investment, and is therefore prone to leakage and overflow. Furthermore, waste collection is limited to the larger businesses in the town centre. Typically, household waste is buried within gardens, burned, or illegally dumped. It should be noted that informal settlements are

beginning to encroach into the wellfield areas, with concerns over the potential threats to the city's groundwater resources in the medium to long-term.

2.2. Groundwater sampling and analysis

A total of 117 groundwater samples were obtained from a mixture of supplies that were distributed across the city (Fig. 1). These were composed of 55 samples in the dry season (September 2013) and 62 in the subsequent wet season (January 2014), of which 45 were obtained from the same supplies. The dry season sampling included 25 boreholes and 30 shallow wells, whereas in the wet season 26 boreholes and 36 shallow wells were investigated. These supplies included the city wellfields (K1-12), a mixture of private supplies within both higher and lower cost residential areas, as well as those in the industrial zone (K26-28).

Groundwater samples were obtained once field measurements of pH, specific electrical conductance (SEC), Eh, dissolved oxygen (DO) and temperature had stabilised during pumping. Turbidity was also measured on an agitated pumped sample, as settling was rapid. In-situ TLF was undertaken by sampling 5 L of groundwater and immersing a portable UviLux Fluorimeter (λ_{ex} = 280 nm, λ_{em} = 360 nm) (Chelsea Technologies Group Ltd, UK) in the dark. The sensor utilises a UV light emitting diode (LED) light source and photomultiplier allowing a high level of sensitivity. The minimum detection limit for these sensors is 0.17–0.19 µg/L (Khamis et al., 2015) and repeatability is within ± 0.12–0.29 µg/L up to a concentration of 50 µg/L (Table S1).

The factory calibration was used which showed a good agreement with a range of ten synthetic standards and a bench Varian™ Cary Eclipse fluorescence spectrophotometer for the same wavelength pair at 20 °C (Fig. 2). Standards were produced by dissolving laboratory grade L-tryptophan (Acros Organics, USA) in ultrapure water. An excellent linear relationship between the portable (in µg/L) and bench (in Raman units) fluorimeters (R^2 0.9952) can be used to convert all concentrations on the portable device to Raman units by multiplication by 0.0024 at 20 °C.

Microbiological samples were collected in sterile 60 ml brown-glass bottles. Brief interviews with the well owners, or other informed persons, confirmed that the supplies sampled had not been recently chlorinated, or were chlorinated at a point further down the distribution system. Samples were stored in a cool box and transported back to the laboratory for analysis. All samples were processed within seven hours of collection.

Thermotolerant coliforms (TTCs) were isolated and enumerated using the membrane filtration method and Membrane Lauryl Sulphate Broth (MLSB. Oxoid Ltd) as the selective medium. Typically 50 mL (giving a limit of detection of 2 c.f.u/100 mL), or an appropriate dilution of the sample, were filtered through a 0.45 µm nitrocellulose membrane (PALL Gelman). The membrane was transferred on to an adsorbent pad saturated with MLSB, pre-incubated for approximately one hour at ambient temperature (around 25 °C), and then incubated at 44 °C for a total incubation time of between 18 and 24 h. Plates were examined within 15 min of being removed from the incubator. All cream to yellow colonies with a diameter greater than 1 mm were considered to be TTCs and were counted. No further tests were done to confirm the isolates. The results from 26 duplicate measurements were used to calculate the average Relative Standard Deviation of Reproducibility (RSD_{RC}) and the Uncertainty of Measurement of the microbiological analysis.

Hydrochemical samples were collected for Cl, NO_3, NH_4, SO_4, soluble reactive P (SRP), and total dissolved P (TDP) in 60 mL HDPE bottles after passing through a 0.45 µm nitrocellulose filter, with the bottle for SRP and TDP pre-treated with 0.45 g of potassium
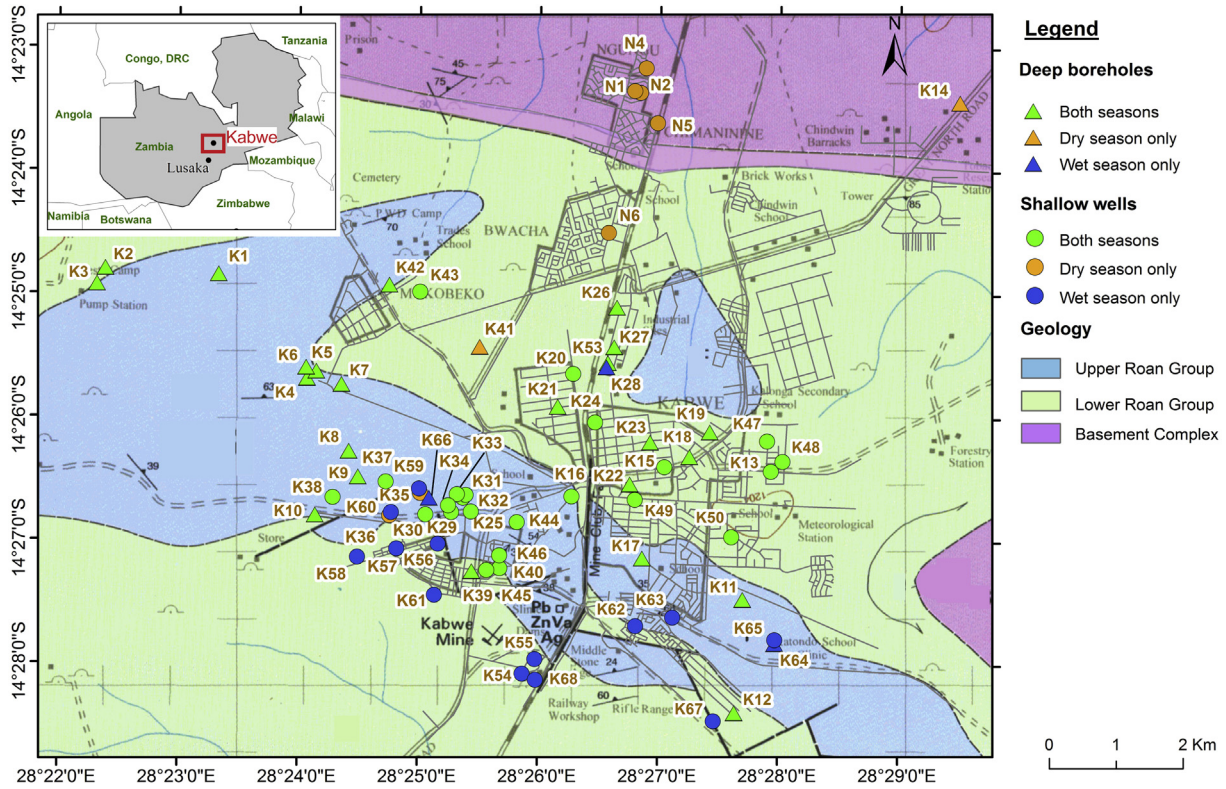
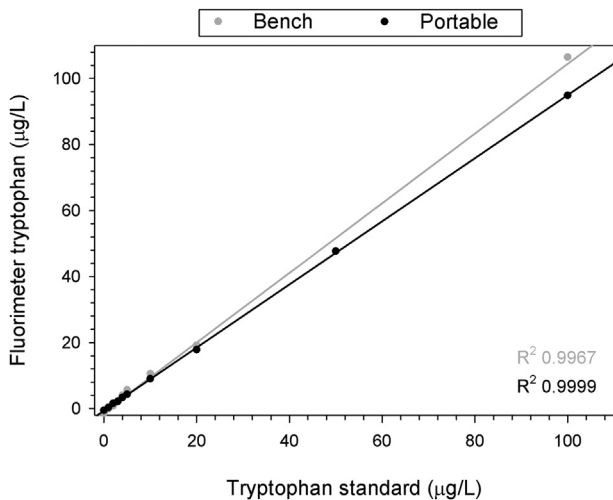**Fig. 1.** Bedrock geology and sampled groundwater supplies in Kabwe.

Fig. 2. Comparison of bench and portable fluorimeters with ten synthetic tryptophan standards.

peroxodisulfate preservative. Samples for dissolved organic carbon (DOC) were passed through 0.45 μm silver filters into sterile acid washed glass vials. All samples were stored within cool boxes before transfer to a refrigerator at the end of each day. Cl, NO_3, SO_4 were analysed by ion chromatography. NH_4 was determined by automated calorimetry using the indophenol blue method. Total dissolved phosphate and soluble reactive phosphate were determined via the methods of Eisenreich et al. (1975) after treatment with sulphuric acid. DOC analysis was conducted using a Thermalox™ C analyser after acidification and sparging.

2.3. Sanitary risk score

A sanitary risk assessment was undertaken at every supply during the wet and dry season surveys. The assessment was carried out using the methods published by the WHO and the sanitary risk assessment forms appropriate to the supply type (WHO, 1997). The supplies and surrounding areas were assessed for the presence of defined hazards, such as latrines and waste dumps, and pathways for contaminants to reach the sample point, using pre-set questions. The sanitary risk score (SRS) was calculated from the number of positive responses.

2.4. Statistical analyses

Statistical models and tests were performed with MATLAB v14.1 to determine the extent to which TLF concentration and other predictors indicate the presence of TTC contamination and the accuracy with which they could be used to model TTC counts. The TTC count in water supply i was denoted y_i and the binary variable, y_i^I, indicated whether or not the TTC exceeded the detection limit of 2 counts, i.e.

$$y_i^I = \begin{cases} 1 & \text{if } y_i \geq 2 \\ 0 & \text{otherwise} \end{cases}.$$

2.4.1. Characteristics of TTC contaminated water supplies

A series of Mann–Whitney rank sum tests (Mann and Whitney, 1947) were used to assess whether or not the values of each of the predictors differed between the water supplies with TTC contamination ($y_i^I = 1$) and those which were uncontaminated ($y_i^I = 0$). The non-parametric Mann–Whitney rank sum test does not

require the predictors to conform to a particular statistical distribution. Therefore, it was preferred to the T-test which requires the assumption of a Gaussian distribution. In each test the null hypothesis was that the expected value of the predictor was the same for contaminated and uncontaminated supplies.

2.4.2. Predicting the presence/absence of TTCs

The probability that a water supply was contaminated with TTCs was represented by a logistic regression model (Dobson, 2001). If p_i is the probability that there is contamination at supply i, then the logistic model is written:

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \sum_{j=1}^{q} \beta_j x_{ij}, \tag{1}$$

where x_{ij} is the value of predictor j at water supply i and the β_j are coefficients which, in this study, were estimated by least squares (Draper and Smith, 1981). A stepwise regression algorithm (Draper and Smith, 1981) was used to decide which of the available predictors should be included in the model. This approach guards against too many predictors being included. In such circumstances the model is said to be overfitted, meaning that it is overly suited to the intricacies of the data set to which it has been fitted but will not achieve the same accuracy on independent validation data.

The stepwise algorithm adds one predictor to the model at a time. It considers all of the available predictors not yet included in the model in turn and selects the one which causes the largest decrease in the mean squared residuals. If this decrease is deemed significant for a p-value of 0.05, according to a F-test (Draper and Smith, 1981), this parameter is added to the model. The iterative process continues with the remaining parameters until none of them lead to a significant improvement to the model. The stepwise algorithm was implemented using the MATLAB function 'stepwiseglm'. The observed values of TLF were highly skewed (Figure S1). Therefore these data were log transformed prior to the estimation of regression models. Note that a positive shift of 0.1 was required to ensure that the logarithm was always defined.

A classifier of TTC contamination was then formed by designating all water supplies with p_i greater than some threshold p^* as contaminated. There is uncertainty associated with such a classification. False positive errors occur when the classifier erroneously suggests that the supply is contaminated whereas false negative errors occur when the supply is erroneously classified as uncontaminated. Similarly, true positive and true negative results refer to correctly classified contaminated and uncontaminated supplies.

The effectiveness of a classifier can be assessed in terms of the receiver operating characteristic curve (Hanley and McNeil, 1983). This is a plot of the proportion of true positive results against the proportion of false positive results as p^* is varied. The area under this curve is a measure of the effectiveness of the classifier. It will be 1 for a perfect classifier and 0.5 if the classifier is performing no better than a random choice.

2.4.3. Predicting the number of TTCs

The TTC counts were represented by a linear regression model:

$$y_i = \beta_0 + \sum_{j=1}^{q} \beta_j x_{ij} + \varepsilon_i, \tag{2}$$

where the ε_i were independent realizations of Gaussian random variable with zero mean and variance σ^2. The β_j and σ^2 were the model parameters which in this study were estimated by least squares. Again, the optimal predictors for the model were selected by stepwise regression. When a linear regression model was

estimated for TTC counts, the residuals were highly skewed, contrary to the assumptions required by the model. Therefore a model was instead estimated for the logarithm of the TTC counts.

A regression model such as Equation (2) determines the entire probability density function for the TTC count at a water supply from the values of the predictor variables from that supply. Therefore, the model can be used to determine the probability that the TTC count falls within particular intervals. In this case, the five risk classes outlined by the WHO (1997) to prioritise interventions in areas where microbiological contamination of water supplies is widespread were used.

3. Results

3.1. Characteristics of TTC contaminated water supplies

The results of the Mann–Whitney rank Sum tests suggested that the mean TLF concentration was significantly ($p < 0.001$) larger in water supplies contaminated with TTCs as opposed to water supplies where they were absent (Fig. 3). The relationship remained significant at the $p < 0.001$ level when the supplies from each season were considered separately and when only boreholes were considered. No significant difference was evident when only shallow wells were considered but, since there were only five shallow wells where TTCs were absent, this finding is likely to be the result of insufficient data. Overall, the median TLF concentrations were 0.9 and 9.2 µg/L in uncontaminated and contaminated supplies, respectively.

Other variables showing a highly significant difference ($p < 0.001$) included SRS, turbidity, depth to water, Cl and SO_4 (Fig. 3). Phosphate species or DOC were not appreciably elevated in contaminated water supplies, although NH_4 was significant ($p = 0.019$). Temperature, pH and turbidity typically ranged between 23.8 and 27.7 °C, 6.4–7.4, and 1–20 NTU, respectively, across the dataset. All variables correlated with faecal contamination were also significantly correlated with TLF concentration (notably SRS, NO_3, and NH_4), with no relationship observed with DOC, pH, DO, or temperature (Figure S2).

3.2. Predicting the presence/absence of TTCs

Transformed (and raw) TLF was a significant predictor ($p < 0.001$) of the presence/absence of TTC according to the logistic regression model (Equation (1)). Other indicators were significantly correlated with the probability of TTC contamination (Table 1), including Cl and the SRS. However, transformed TLF was the first predictor to be included in the logistic regression model and the results of the F-tests demonstrate that the addition of other predictors did not significantly improve the model (Table S2). The findings were the same when data from the wet and dry seasons were considered individually, as well as when boreholes and shallow wells were considered separately. Therefore, for the predictors investigated in this study, transformed TLF appears to be the best indicator of the presence/absence of TTCs.

When the logistic regression model with only TLF as a predictor was used to form a classifier of TTC contamination, the area under the receiver operator characteristic curve was 0.92 (Fig. 4a). This is much closer to the perfect classifier value of 1 than the random selection value of 0.5. The corresponding areas for other single predictor models ranged from 0.63 to 0.90 (Table 1). There is a trade-off between false positive and false negative errors as the classifier's threshold on TLF readings is varied (Fig. 4b). When the threshold is 3.5 µg/L both the false positive and false negative error rates are approximately 0.14. However, there are a small number of sites where the contamination status inferred from TLF and TTC
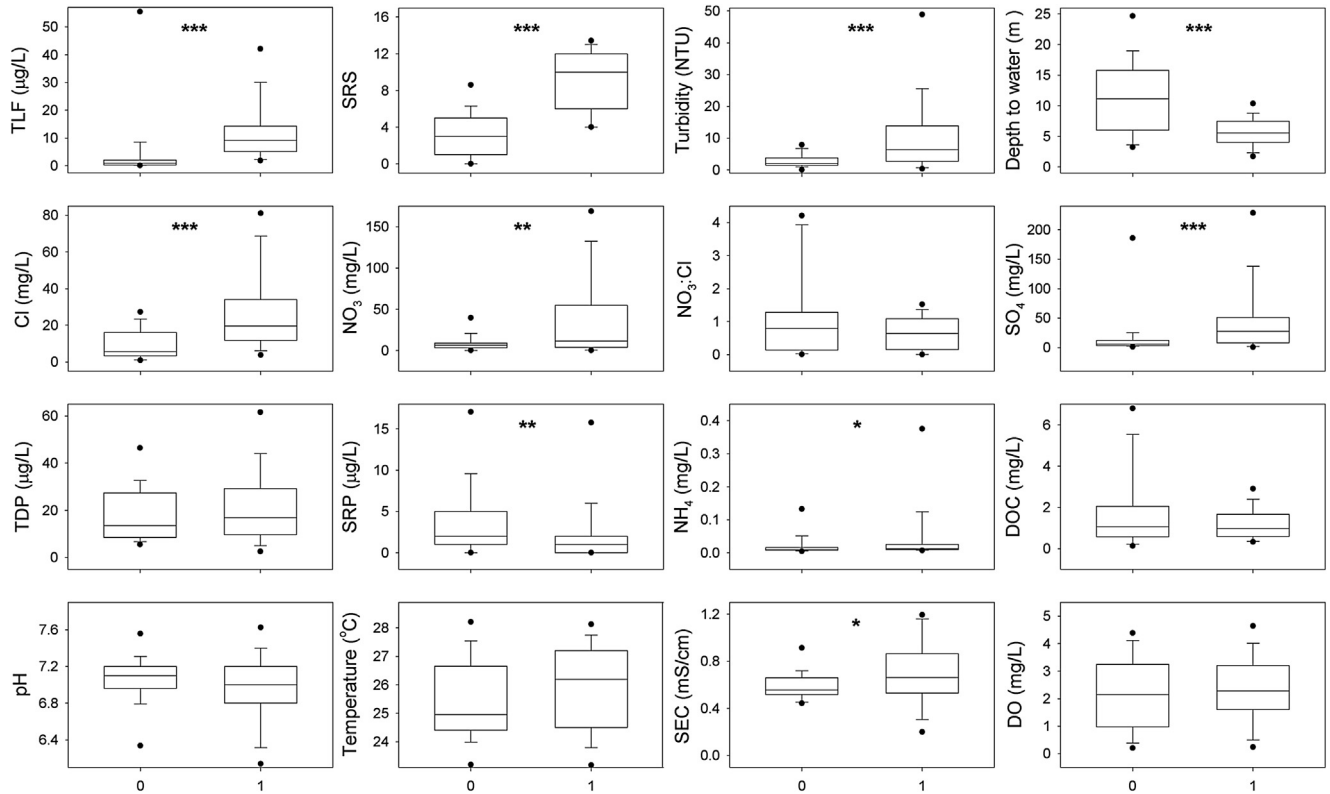
Fig. 3. Box plots of transformed tryptophan-like fluorescence and other variables for the presence or absence of thermotolerant coliforms. P-values are denoted as * <0.05, ** <0.01, and *** <0.001. Dots indicate 5th and 95th percentiles.

Table 1
Estimated coefficients and p-values for the hypothesis that the coefficients are zero for single-predictor models of probability of the presence and number of thermotolerant coliforms. Areas under the receiver operator characteristic curves for single predictor classifiers are also shown. Classifiers are only formed for predictors that are significantly ($p < 0.05$) related to contamination.

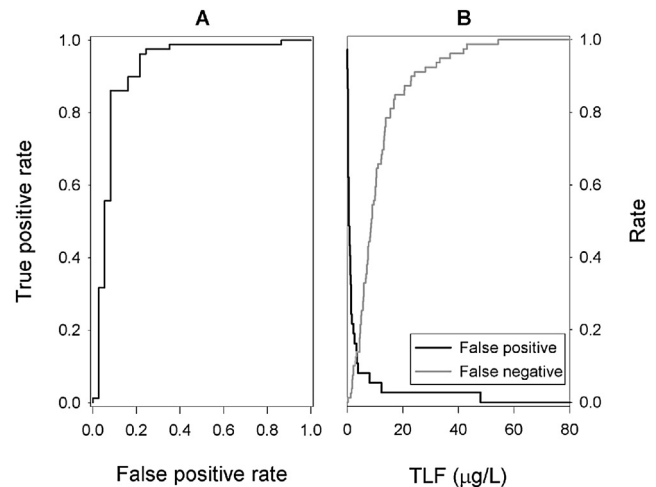| Predictor	Logistic regression			Linear regression	
	Area	β	p-value	β	p-value
Transformed TLF	0.92	2.12	<0.001	1.78	<0.001
SRS	0.90	0.61	<0.001	0.63	<0.001
Turbidity	0.76	0.48	0.019	1.05	<0.001
Depth to water	0.80	−2.86	<0.001	−2.42	<0.001
Cl	0.86	1.63	<0.001	1.61	<0.001
NO₃	0.73	0.35	0.019	0.86	<0.001
NO3:Cl	NA	−0.34	0.221	0.36	0.331
SO₄	0.71	0.31	0.050	0.33	0.128
TP	NA	0.25	0.382	−0.09	0.820
SRP	0.63	−0.41	0.024	−0.47	0.039
NH₄	NA	2.88	0.174	2.04	0.016
DOC	NA	−0.74	0.057	−0.59	0.286
SEC	NA	0.00	0.502	0.00	0.101
DO	NA	−0.17	0.423	0.21	0.489

Fig. 4. (A) Receiver operator characteristic curve for classifier of the presence or absence of TTC using transformed TLF as the only predictor; (B) error rates for this classifier.

differ. For example there are two water supplies where the TLF concentration is less than 1.5 μg/L but the TTC count is 2 and 10 c.f.u/100 mL. Note that since TLF is the only predictor and the relationship between these readings and the probability of contamination is monotonic, putting a threshold on these readings is equivalent to putting a threshold on the probability of contamination.

3.3. Predicting the number of TTCs

Transformed TLF was a significant predictor ($p < 0.001$) of the number of TTCs (Table 1). Other similarly significant indicators included Cl, NO₃, SRS, and turbidity. However, the stepwise regression algorithm included transformed TLF first, signifying it is the most important predictor. When transformed TLF is the only predictor the estimated model has an R^2 of 0.57 and is written:

$$\ln y_i = 1.36 + 1.78 \ln(\text{TLF} + 0.1). \qquad (3)$$

NO₃ and SRS were added at the second and third stages of the stepwise algorithm. These predictors cause the R^2 to increase to

0.67 and the final model is written:

$$\ln y_i = -0.34 + 1.11 \ln(\mathrm{TLF} + 0.1) + 0.35 \, \mathrm{NO_3^-} + 0.25 \, \mathrm{SRS}. \tag{4}$$

The addition of other predictors did not significantly improve the model (Table S2). This might indicate that the predictors included in Equation (4) accounted for the other observed significant relationships.

Additionally, there is an increase in the median transformed TLF concentration with elevated WHO risk class, although the levels of scatter produce overlap between all classes (Fig. 5a). Nevertheless, the single-predictor linear regression model using TLF can be used to estimate the probability of a water supply belonging to each risk class (Fig. 5b). This highlights that as TLF concentration increases there is a greater probability of the water supply being within higher risk groups.

3.4. Risk classes for water supplies across Kabwe estimated by TTC count or modelled TLF

Deeper boreholes are considered less at-risk than shallow wells across the city using either modelled TLF or TTC count (Fig. 6). Boreholes are most at-risk within the central districts, with these risks elevated in the wet season, and higher when allocated using modelled TLF compared to the TTC count. There are also notable seasonal changes in risk in the city's centralised peri-urban water supply boreholes (K1-12) using modelled TLF which are not apparent from the TTC data (Fig. 6). In the dry season, there are no risks by TTC count or modelled TLF in any of the boreholes, with the exception of K3 and K11 being considered as low risk by modelled TLF. Following the onset of the rains, modelled TLF suggests increases in risk at the majority of these boreholes, including some located beyond the city limits (K2, K6, and K12). Furthermore, high risks are indicated at those sites closest to the urban area (K8, K9 and K11). On the other hand, TTC data only suggest elevated medium and low wet season risks at K8 and K9, respectively.
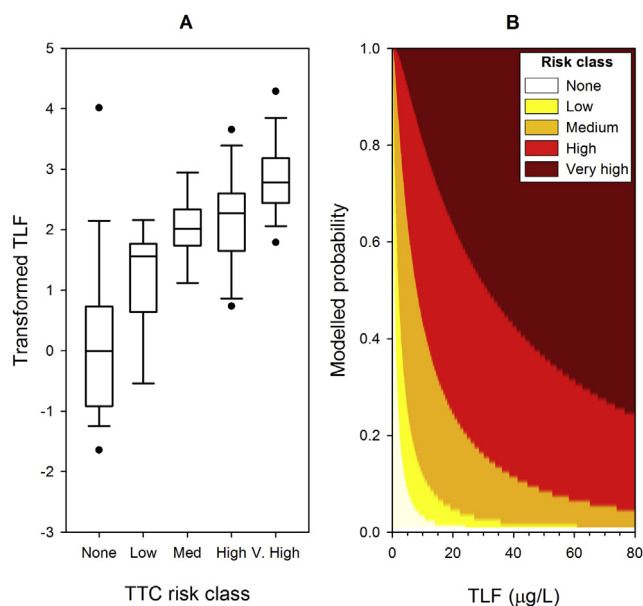
Shallow wells are typically at medium or greater risk perennially in the central districts of Kabwe by modelled TLF (Fig. 6). However, TTC counts suggest many of these wells have no risk during the dry season, but medium or greater during the wet season. Nevertheless wells within Makululu generally remained at least high risk during both seasons using either modelled TLF or TTC count. Elsewhere, other wells also considered at a minimum of high risk by both TTC count and modelled TLF in the dry season, predominantly remain so in the wet season.

4. Discussion

4.1. Uncertainty in TTC counts and TLF concentrations

Doubts have been expressed about the reliability of TTCs to indicate faecal contamination in water. Although the TTC group includes the species *E. coli*, which is generally considered to be specific for faecal contamination, it also includes other genera such as *Klebsiella* and *Citrobacter* which are not necessarily of faecal origin and can emanate from alternative organic sources such as decaying plant materials and soils (WHO, 2011). Nevertheless, studies have shown that greater than 90% of thermotolerant coliforms are *E. coli* (Dufour, 1997 cited in Leclerc et al. (2001)) and as high as 99% in groundwater impacted by poor environmental sanitation in Africa (Howard et al., 2003). Therefore this is not considered a significant source of uncertainty.

The uncertainty of measurement of the microbiological analysis was calculated using the RSD$_{RC}$, which is the average relative standard deviation of reproducibility of a series of duplicate analyses. The RSD$_{RC}$ calculated for the duration of the sampling programme was 0.1433. Using a coverage factor of 2.056, which was determined from the number of duplicate analyses used to calculate RSD$_{RC}$, the 95% confidence intervals for any result could be calculated. For example, a result of 100 c.f.u/100 ml would have an upper and lower limit of 197 and 51 c.f.u/100 ml respectively. This figure is slightly higher than might be anticipated in a laboratory accredited to International Standard ISO17025:2005, but is not unusual for microbiological analysis.

False-negatives and underestimates of TTCs may occur where the environment conditions within certain groundwater supplies are unfavourable for the survival of bacteria, but TLF is less affected. There were three supplies where TTCs were absent but TLF exceeded 10 μg/L. At two of these sites the water differed from a typically neutral pH; being 5.1 and 10.5 in supplies K40 and K55, respectively. Moreover, K40 also contained high concentrations of heavy metals which have a toxic influence on bacteria (Foppen and Schijven, 2006), yet also quench TLF fluorescence to a degree (Tabak et al., 1989). These two sites actually account for two of the five shallow wells where TTCs were absent.

It was confirmed that none of the supplies had recently been chlorinated before sampling. However, a parallel study confirmed the presence of trihalomethanes (by-products of chlorination) at K16 and K30 in the dry season (Sorensen et al., 2015), although this discrepancy could indicate the leakage of chlorinated mains water into the shallow aquifer. In addition to sterilising the supply, chlorination is generally considered to quench TLF (Henderson et al., 2009). However, the results do not appear anomalous with K16 having a TLF of 1.5 μg/L and TTC count of <2 c.f.u/100 mL and K30 having a TLF of 13.4 μg/L and TTC count of 700 c.f.u/100 mL.

Other environmental factors that may influence the intensity of TLF in the environment did not vary appreciably. Thermal quenching may have been limited to approximately 10–20% over the typical temperature range based on previous results using bacteria cultures (Elliott et al., 2006b). This is further supported by a laboratory study using dissolved tryptophan, where Khamis et al.
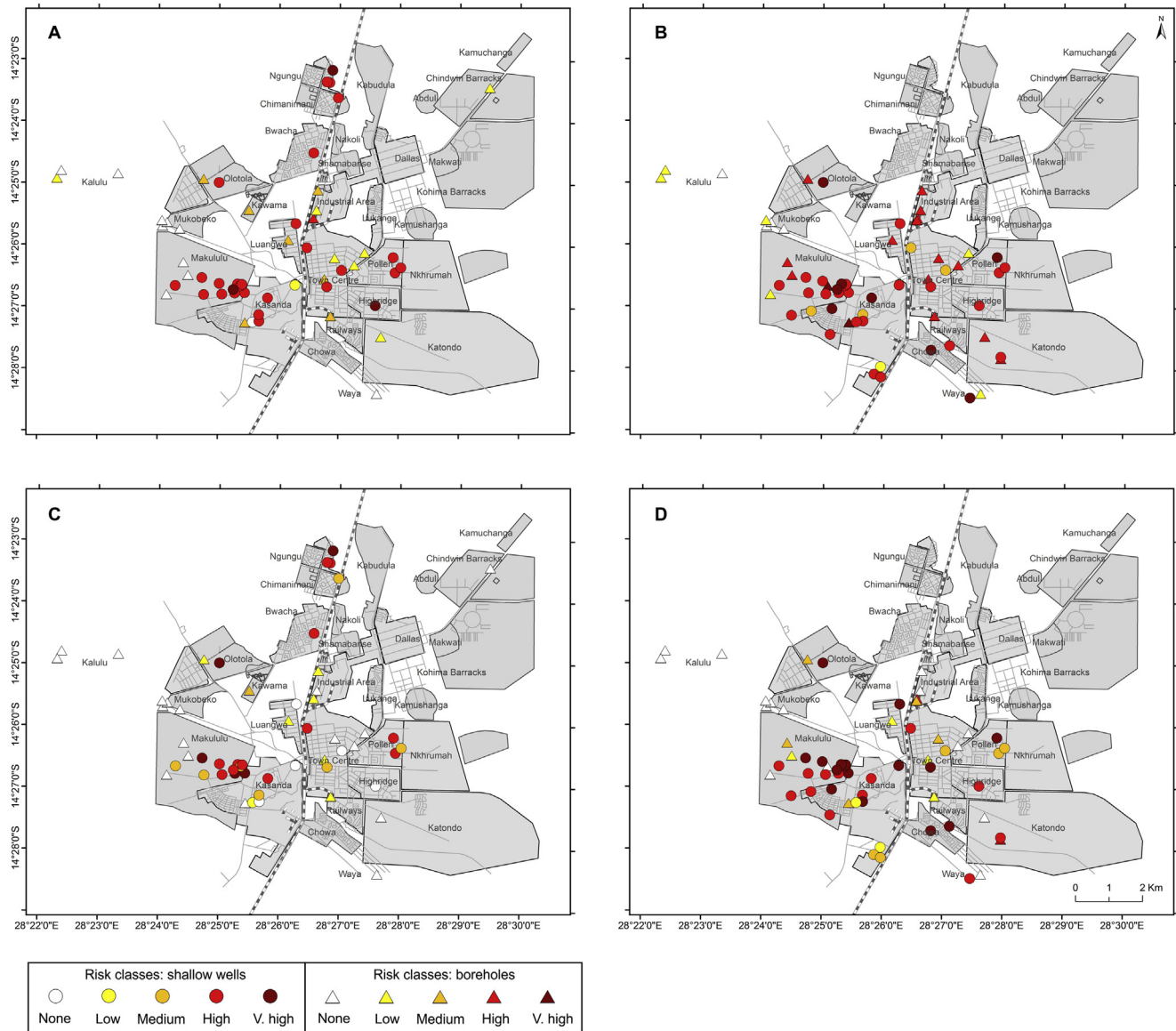

Fig. 5. (A) Box plot of transformed TLF for each TTC risk class with dots indicating 5th and 95th percentiles where the sample number exceeds nine; (B) modelled probability that a water supply belongs to each risk class for a given TLF concentration. Risk classes are defined by TTC count as <2 (none), 2 to <10 (low), 10 to <100 (medium), 100 to <1000 (high), >1000 c.f.u/100 mL (very high).

Fig. 6. Water supply risk class across Kabwe allocated using modelled TLF (A) dry season, (B) wet season; and TTC data (C) dry season, (D) wet season. TTC risk classes are defined as <2 (none), 2 to <10 (low), 10 to <100 (medium), 100 to <1000 (high), >1000 c.f.u/100 mL (very high).

(2015) demonstrated a 0.65 and 1.14 μg/L reduction in TLF for every 1 °C increase in temperature in 10 and 20 μg/L solutions, respectively. This equates to uncertainty of 22—25% across the typical temperature, and it should be re-iterated that there was no observed inverse relationship between TLF and temperature (Figure S2). pH values are within the 5—8 bracket of minimal impact, with the exception of one supply (Patel-Sorrentino et al., 2002; Reynolds, 2003). Turbidity was less than 200 NTU, hence attenuation of TLF was considered negligible, although low-levels of signal amplification due to light scattering by particles could have been possible (Khamis et al., 2015). Finally, no correction of the data due to inner-filtering was necessary (Ohno, 2002), as the mean absorbance of all samples at 254 nm was only 0.04 ± 0.21 (3σ).

4.2. Is TLF a better indicator of enteric pathogens than TTCs?

Bacteria can be effectively removed during the infiltration and percolation of water due to natural filtering and adsorption. This is also likely to remove a proportion of TLF, given its direct association with bacteria (Elliott et al., 2006a). However, a fraction of the fluorescence is likely to be in a free dissolved form (Baker et al., 2007), which could mean it is transported more easily and rapidly through porous media. There is evidence for this in Kabwe where increases in modelled TLF risks were greater than TTC risks in the deeper boreholes following the onset of the seasonal rains. This suggests TLF can migrate through the superficials more effectively. Furthermore, once into the karstic aquifer, increases in modelled TLF risks were identified in boreholes well outside the urban limits where TTCs were absent. A more mobile indicator of enteric pathogens is favourable because viruses are transported more efficiently than bacteria through the subsurface due to their smaller sizes (Borchardt et al., 2007; Hunt et al., 2014).

The perennial persistence of high risks by modelled TLF for shallow wells in central Kabwe, as opposed to the seasonal risks by TTCs suggest TLF is a less transient faecal indicator in groundwater. The presence of TTCs in groundwater is generally considered evidence of recent faecal contamination, with *E. coli* remaining active

for 16—45 days (Taylor et al., 2004). These rates of inactivation are similar for many pathogens but enteric viruses generally inactivate at a slower rate (John and Rose, 2005). Therefore, elevated TLF could be a better indicator of pathogens which are more long-lived in the environment, thus providing a more precautionary estimate of pathogen risk.

It is unsurprising that TLF is more resilient than TTCs in groundwater as it is contained within organisms. Therefore, as TTCs and other organisms die-off following their release into groundwater (an unfavourable habitat) many cells will rupture and release TLF in dissolved and more complex phases into the environment. It may remain present in groundwater in these forms as groundwater ecosystems are less active in comparison to surface waters, or be taken up by other microorganisms. Either way, it will continue to contribute to a TLF signal. Furthermore, it has been demonstrated that under nutrient limited conditions *E. coli* release dissolved tryptophan, as they transit from a culturable to a dormant viable but nonculturable state (Arana et al., 2004).

5. Conclusions

This is the first study to investigate the use of in-situ tryptophan-like fluorescence (TLF) for the rapid assessment of the biological quality of drinking water supplies. TLF was significantly elevated in supplies where thermotolerant coliforms (TTCs) were present, alongside many other traditional indicators, but was demonstrated to be the most effective indicator of TTC presence/absence. TLF was also the most significant indicator of the number of TTCs, although the linear regression model could be improved through the inclusion of NO_3 and sanitary risk scores (SRS). A single-predictor linear regression model using TLF was used to estimate the probability of a water supply belonging to each WHO risk class. This highlights that as TLF concentration increases there is a greater probability of a water supply being within the higher risk groups.

The use of in-situ TLF sensors has multiple methodology advantages over bacterial indicators for inferring the presence of enteric pathogens. TLF is likely to be more mobile and resilient in groundwater and, as such a more precautionary indicator of enteric pathogens in groundwater, including where bacterial indicators are absent. Importantly, TLF has the added advantage of potentially detecting the presence of TTC in viable but nonculturable states. Furthermore, the sensors require no reagents and provide instantaneous readings. This could facilitate their inclusion in real-time pollution alert systems for drinking water supplies throughout the world, for the rapid mapping of enteric pathogen risks in developing regions, and as an initial screening tool to inform and complement further water quality investigations.

Acknowledgements

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.watres.2015.05.035.

References

Arana, I., Seco, C., Epelde, K., Muela, A., Fernández-Astorga, A., Barcina, I., 2004. Relationships between Escherichia coli cells and the surrounding medium during survival processes. Antonie Leeuwenhoek 86 (2), 189—199.

Ashbolt, N.J., 2004. Microbial contamination of drinking water and disease outcomes in developing regions. Toxicology 198 (1), 229—238.

Baker, A., 2001. Fluorescence excitation-emission matrix characterization of some sewage-impacted rivers. Environ. Sci. Technol. 35 (5), 948—953.

Baker, A., 2005. Thermal fluorescence quenching properties of dissolved organic matter. Water Res. 39 (18), 4405—4412.

Baker, A., Elliott, S., Lead, J.R., 2007. Effects of filtration and pH perturbation on freshwater organic matter fluorescence. Chemosphere 67 (10), 2035—2043.

Borchardt, M.A., Bradbury, K.R., Gotkowitz, M.B., Cherry, J.A., Parker, B.L., 2007. Human enteric viruses in groundwater from a confined bedrock aquifer. Environ. Sci. Technol. 41 (18), 6606—6612.

Cammack, W.L., Kalff, J., Prairie, Y.T., Smith, E.M., 2004. Fluorescent dissolved organic matter in lakes: relationships with heterotrophic metabolism. Limnol. Oceanogr. 49 (6), 2034—2045.

Cumberland, S., Bridgeman, J., Baker, A., Sterling, M., Ward, D., 2012. Fluorescence spectroscopy as a tool for determining microbial quality in potable water applications. Environ. Technol. 33 (6), 687—693.

Dobson, A.J., 2001. An Introduction to Generalized Linear Models. CRC Press.

Downing, B.D., Pellerin, B.A., Bergamaschi, B.A., Saraceno, J.F., Kraus, T.E., 2012. Seeing the light: the effects of particles, dissolved materials, and temperature on in situ measurements of DOM fluorescence in rivers and streams. Limnol. Oceanogr. Methods 10, 767—775.

Draper, N.R., Smith, H., 1981. Applied Regression Analysis, second ed.

Eisenreich, S., Bannerman, R., Armstrong, D., 1975. A simplified phosphorus analysis technique. Environ. Lett. 9 (1), 43—53.

Elliott, S., Lead, J., Baker, A., 2006a. Characterisation of the fluorescence from freshwater, planktonic bacteria. Water Res. 40 (10), 2075—2083.

Elliott, S., Lead, J., Baker, A., 2006b. Thermal quenching of fluorescence of freshwater, planktonic bacteria. Anal. Chim. Acta 564 (2), 219—225.

Foppen, J., Schijven, J., 2006. Evaluation of data from the literature on the transport and survival of Escherichia coli and thermotolerant coliforms in aquifers under saturated conditions. Water Res. 40 (3), 401—426.

Hanley, J.A., McNeil, B.J., 1983. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 148 (3), 839—843.

Henderson, R., Baker, A., Murphy, K., Hambly, A., Stuetz, R., Khan, S., 2009. Fluorescence as a potential monitoring tool for recycled water systems: a review. Water Res. 43 (4), 863—881.

Houston, J., 1982. Rainfall and recharge to a dolomite aquifer in a semi-arid climate at Kabwe, Zambia. J. Hydrol. 59 (1), 173—187.

Howard, G., Pedley, S., Barrett, M., Nalubega, M., Johal, K., 2003. Risk factors contributing to microbiological contamination of shallow groundwater in Kampala, Uganda. Water Res. 37 (14), 3421—3429.

Hudson, N., Baker, A., Reynolds, D., 2007. Fluorescence analysis of dissolved organic matter in natural, waste and polluted waters—a review. River Res. Appl. 23 (6), 631—649.

Hudson, N., Baker, A., Ward, D., Reynolds, D.M., Brunsdon, C., Carliell-Marquet, C., Browning, S., 2008. Can fluorescence spectrometry be used as a surrogate for the Biochemical Oxygen Demand (BOD) test in water quality assessment? An example from South West England. Sci. Total Environ. 391 (1), 149—158.

Hunt, R.J., Borchardt, M.A., Bradbury, K.R., 2014. Viruses as groundwater tracers: using ecohydrology to characterize short travel times in aquifers. Groundwater 52 (2), 187—193.

John, D.E., Rose, J.B., 2005. Review of factors affecting microbial survival in groundwater. Environ. Sci. Technol. 39 (19), 7345—7356.

Khamis, K., Sorensen, J., Bradley, C., Hannah, D., Lapworth, D.J., Stevens, R., 2015. In situ tryptophan-like fluorometers: assessing turbidity and temperature effects for freshwater applications. Environ. Sci. Process. Impacts 17 (4), 740—752.

Lapworth, D., Gooddy, D., Morris, B., Butcher, A., 2007. Investigating Groundwater Movement and Pathogen Transport in Sandstone Aquifers Using Intrinsic Fluorescence Spectroscopy.

Lapworth, D.J., Gooddy, D., Butcher, A., Morris, B., 2008. Tracing groundwater flow and sources of organic carbon in sandstone aquifers using fluorescence properties of dissolved organic matter (DOM). Appl. Geochem. 23 (12), 3384—3390.

Leclerc, H., Mossel, D., Edberg, S., Struijk, C., 2001. Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. Annu. Rev. Microbiol. 55 (1), 201—234.

LgWSC, 2014. Concept Paper for Kabwe Sanitation Interventions, p. 10.

Macler, B.A., Merkle, J.C., 2000. Current knowledge on groundwater microbial pathogens and their control. Hydrogeol. J. 8 (1), 29—40.

Mann, H.B., Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Statistics 50—60.

Nkhuwa, D., Xu, Y., Usher, B., 2006. Groundwater quality assessments in the John Laing and Misisi areas of Lusaka. Groundw. Pollut. Afr. 239—251.

Ohno, T., 2002. Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. Environ. Sci. Technol. 36 (4), 742—746.

Patel-Sorrentino, N., Mounier, S., Benaim, J., 2002. Excitation—emission fluorescence matrix to study pH influence on organic matter fluorescence in the Amazon

basin rivers. Water Res. 36 (10), 2571–2581.

Reynolds, D.M., 2003. Rapid and direct determination of tryptophan in water using synchronous fluorescence spectroscopy. Water Res. 37 (13), 3055–3060.

Savichtcheva, O., Okabe, S., 2006. Alternative indicators of fecal pollution: relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. Water Res. 40 (13), 2463–2476.

Sorensen, J.P.R., Lapworth, D.J., Nkhuwa, D.C.W., Stuart, M.E., Gooddy, D.C., Bell, R.A., Chirwa, M., Kabika, J., Liemisa, M., Chibesa, M., Pedley, S., 2015. Emerging contaminants in urban groundwater sources in Africa. Water Res. 72, 51–63.

Tabak, M., Sartor, G., Cavatorta, P., 1989. On the interactions of metal ions with tryptophan and glyciltryptophan: a fluorescence study. J. Luminescence 43 (6), 355–361.

Taylor, R., Cronin, A., Pedley, S., Barker, J., Atkinson, T., 2004. The implications of groundwater velocity variations on microbial transport and wellhead protection–review of field evidence. FEMS Microbiol. Ecol. 49 (1), 17–26.

WHO, 1997. Guidelines for Drinking-water Quality, second ed. WHO Press, Geneva, Switzerland.

WHO, 2002. The World Health Report: 2002: Reducing the Risks, Promoting Healthy Life.

WHO, 2011. Guidelines for Drinking-water Quality, fourth ed. WHO Press, Geneva, Switzerland.