In-situ fluorescence spectroscopy is a more rapid and resilient indicator of faecal contamination risk in drinking water than faecal indicator organisms

James P.R. Sorensen [a,b,*], Jacintha Nayebare [c], Andrew F. Carr [b], Robert Lyness [d], Luiza C. Campos [d], Lena Ciric [d], Timothy Goodall [e], Robinah Kulabako [f], Catherine M. Rushworth Curran [g], Alan M. MacDonald [h], Michael Owor [c], Daniel S. Read [e], Richard G. Taylor [b]

[a] British Geological Survey, Maclean Building, Wallingford, OX10 8BB, United Kingdom of Great Britain and Northern Ireland UK
[b] Department of Geography, University College London, London WC1E 6BT, United Kingdom of Great Britain and Northern Ireland UK
[c] Department of Geology and Petroleum Studies, Makerere University, Uganda
[d] Department of Civil, Environmental & Geomatic Engineering, University College London, London WC1E 6BT, United Kingdom of Great Britain and Northern Ireland UK
[e] UK Centre for Ecology & Hydrology (UKCEH), Maclean Building, Wallingford, OX10 8BB, United Kingdom of Great Britain and Northern Ireland UK
[f] Department of Civil and Environmental Engineering, Makerere University, Uganda
[g] Catherine M Rushworth Curran Ltd., 27 Silverhall Street, Isleworth, TW7 6RF, United Kingdom of Great Britain and Northern Ireland UK
[h] British Geological Survey, Lyell Centre, Research Avenue South, Edinburgh EH14 4AP, United Kingdom of Great Britain and Northern Ireland UK

A R T I C L E I N F O

A B S T R A C T

Faecal indicator organisms (FIOs) are limited in their ability to protect public health from the microbial contamination of drinking water because of their transience and time required to deliver a result. We evaluated alternative rapid, and potentially more resilient, approaches against a benchmark FIO of thermotolerant coliforms (TTCs) to characterise faecal contamination over 14 months at 40 groundwater sources in a Ugandan town. Rapid approaches included: in-situ tryptophan-like fluorescence (TLF), humic-like fluorescence (HLF), turbidity; sanitary inspections; and total bacterial cells by flow cytometry. TTCs varied widely in six sampling visits: a third of sources tested both positive and negative, 50% of sources had a range of at least 720 cfu/100 mL, and a two-day heavy rainfall event increased median TTCs five-fold. Using source medians, TLF was the best predictor in logistic regression models of TTCs $\geq$10 cfu/100 mL (AUC 0.88) and best correlated to TTC enumeration (ρ_s 0.81), with HLF performing similarly. Relationships between TLF or HLF and TTCs were stronger in the wet season than the dry season, when TLF and HLF were instead more associated with total bacterial cells. Source rank-order between sampling rounds was considerably more consistent, according to cross-correlations, using TLF or HLF (min ρ_s 0.81) than TTCs (min ρ_s 0.34). Furthermore, dry season TLF and HLF cross-correlated more strongly (ρ_s 0.68) than dry season TTCs (ρ_s 0.50) with wet season TTCs, when TTCs were elevated. In-situ TLF or HLF are more rapid and resilient indicators of faecal contamination risk than TTCs.

1. Introduction

Microbial contamination of drinking water remains a primary water quality concern in low-, middle-, and high-income countries (Hunter et al., 2010; WHO 2017b). The greatest public health risk relates to the consumption of drinking water contaminated with human and animal faeces to which at least two billion people are currently exposed worldwide (WHO 2019). Faecal contamination of drinking water sources has traditionally been assessed by overnight culturing of surrogate faecal indicator organisms (FIOs) to infer the potential presence of enteric pathogens. However, drinking water compliance monitoring using FIOs provides ineffective protection of public health (Stelma Jr and Wymer 2012; WHO 2017a) and waterborne outbreaks remain common, even in high-income countries (Collier et al., 2021). The main concerns relating to FIOs are that microbial contamination is highly variable temporally, which is not characterised by infrequent (e.g. quarterly/annual) FIO sampling in many circumstances (Hrudey and Hrudey 2004), and results are delivered after exposure has occurred.

Furthermore, FIO analysis requires well-trained personnel, restricting the extent of nationally representative surveys, and because no result is provided in-situ at the source, communication of risks and behavioural change is also inhibited (UNICEF/WHO 2017).

To address some of these limitations with FIO monitoring, the World Health Organisation (WHO) recommends a risk-based management approach to ensure water safety (WHO 2017b). A risk-based approach often includes sanitary inspections of the source (Kelly et al., 2020) and operational monitoring of parameters that can be quantified rapidly to indicate changes in source water quality (WHO 2017b), notably turbidity (WHO 2017d), in addition to FIO culturing. There is also a current drive by UNICEF/WHO (2017) for the development of new water quality approaches for the more rapid detection of faecal contamination risk.

Fluorescence spectroscopy is a rapid, reagentless technique used to characterise fluorescent natural organic matter (NOM) in water (Bier-oza et al., 2009; Carstea et al., 2010; Fellman et al., 2010; Hudson et al., 2007). There is substantial evidence that natural waters contaminated with wastewater display enhanced fluorescent NOM (Baker 2001; Baker and Inverarity 2004; Carstea et al., 2016; Goldman et al., 2012; Lapworth et al., 2008; Reynolds and Ahmad 1997; Zhou et al., 2016). Of particular interest has been a fluorescence peak at an excitation (λ_{ex})/emission (λ_{em}) wavelength pair around 280/350 nm, termed tryptophan-like fluorescence (TLF). TLF has long been considered an indicator of biological activity in water (Cammack et al., 2004; Elliott et al., 2006; Sorensen et al., 2020b) and occurs in high concentrations in human and animal wastes (Baker 2001, 2002). This latter observation led to the suggestion that fluorescence spectroscopy could be a useful early-warning indicator for the wastewater contamination of drinking water (Dalterio et al., 1986; Fox et al., 2017; Stedmon et al., 2011). Fluorescence can be quantified either instantaneously in-situ using portable sensors (Carstea et al., 2020) or in real-time during online deployment at piped water sources (Sorensen et al., 2018a).

There is now growing evidence fluorescence spectroscopy is an instantaneous indicator of faecally contaminated drinking water, as determined by the relationship between TLF and the FIO thermotolerant (faecal) coliforms (TTCs) (Sorensen et al., 2015a; Sorensen et al., 2016; Sorensen et al., 2018b; Ward et al., 2020), or specifically *Escherichia coli* (Baker et al., 2015; Fox et al., 2017; Frank et al., 2017; Mendoza et al., 2020; Nowicki et al., 2019; Sorensen et al., 2018a). Furthermore, laboratory studies have shown that *E. coli* directly produce TLF and excrete compounds that fluoresce in the TLF region (Dalterio et al., 1986; Fox et al., 2017). In a collation of groundwater and surface data across four countries ($n = 564$), a TLF threshold of 1.3 ppb dissolved tryptophan could classify TTC and *E. coli* presence-absence with false-negative and false-positive error rates of 4 and 15%, respectively (Sorensen et al., 2018b). There was also a very strong correlation (ρ_s 0.80) between TLF intensity and TTC and *E. coli* enumeration. Importantly, there is also provisional evidence that TLF is a more resilient indicator of faecal contamination risk than TTCs in groundwater (Sorensen et al., 2015a). We have demonstrated that modelled faecal contamination risk using TLF remained perennially elevated in some contaminated sources, whilst risks suggested by TTCs were only seasonally elevated (Sorensen et al., 2015a).

Some studies have also demonstrated strong relationships between humic-like fluorescence (HLF) (λ_{em} of 400–480 nm) and *E. coli* either in the laboratory (Fox et al., 2017) or in groundwater (Frank et al., 2017; Sorensen et al., 2018a), although any relationship between FIOs and HLF is less well documented than FIOs and TLF. HLF has typically been considered of terrestrial origin in freshwater (Coble et al., 2014), but it is elevated in wastewater (Hur et al., 2010; Sihan et al., 2021) and can also be produced in-situ by bacteria, including *E. coli* (Fox et al., 2017; Kida et al., 2019).

In this study, we evaluate the utility of both TLF and HLF as instantaneous, in-situ indicators of faecal contamination risk in groundwater, the world's largest store of freshwater and the primary source of drinking water for up to two billion people (Gleeson et al., 2010). We repeatedly sampled 40 groundwater sources in a community in Uganda across a period of fourteen months for TTCs and alternative rapid approaches that could be used to indicate faecal contamination. The rapid indicators included standard approaches of sanitary inspections, turbidity and electrical conductivity, alongside the more novel indicators of in-situ fluorescence spectroscopy and total (planktonic) bacterial cells (TBCs) by flow cytometry. We aim to demonstrate: (1) in-situ TLF and HLF are the superior rapid indicators of TTCs; and (2) the seasonal nature of the associations amongst TLF, HLF, TTCs, and TBCs.

2. Methods

2.1. Study area

Lukaya is a town in central Uganda around 100 km southwest of the capital city Kampala and close to the shores of Lake Victoria (Fig. 1A). The town's population was 24,000 in the last census (UBOS 2014), with a density of *c.* 640 inhabitants per km^2 within the built-up area that is growing at 3% per year (PDP 2017). To the east of town is the Lweera Swamp where commercial rice farming is practised (Fig. 1B). The
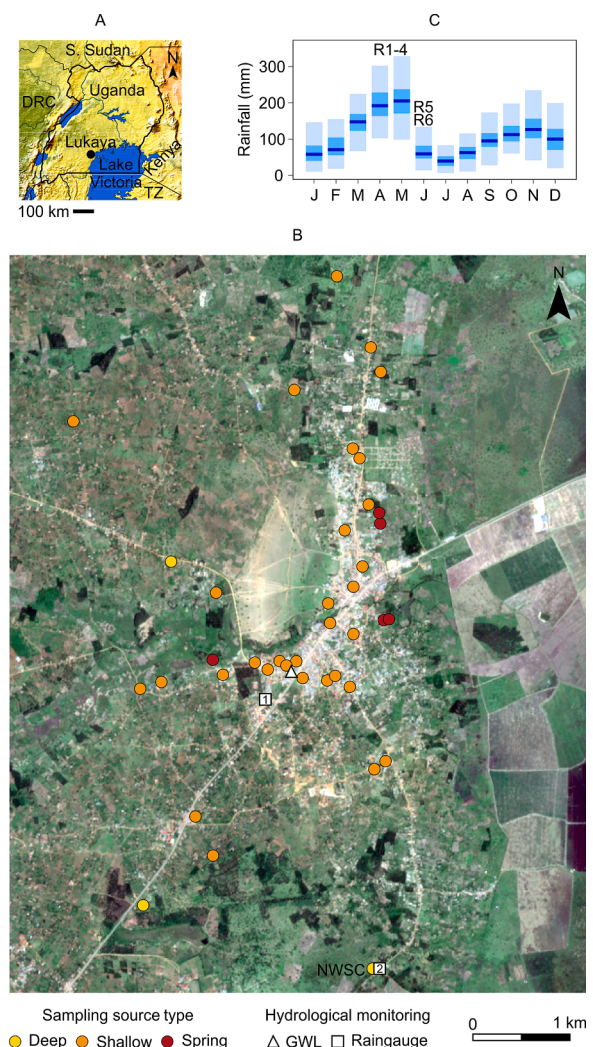

Fig. 1. (A) Location of Lukaya within Uganda (TZ = Tanzania); (B) Sampling sources and hydrological monitoring in Lukaya mapped on Copernicus Sentinel data (2020), with the piped water source labelled NWSC; (C) Tukey boxplot, excluding outliers, of CRU monthly rainfall data (1900–2019) for grid cell 0.25 S, 31.75 E (Harris et al., 2020) indicating the timing of sampling rounds R1–6.

climate is humid with a mean annual rainfall of 890 mm (Nayebare et al., 2020) that is bimodal and focussed within the rainy seasons of March to May and September to November (Fig. 1C).

The town predominantly sits on Precambrian basement rocks with aquifers developed within the weathered overburden and fractured bedrock with a shallow water table between 0.5 and 9 m below ground level (bgl). Groundwater is the primary source of water for the town, with the majority obtained from hand pumped wells and springs. Piped water is used by <1% of households (Nayebare 2021), which is obtained from a borehole operated by the National Water Sewerage Corporation (NWSC) to the south of the town (Fig. 1B) (Nayebare et al., 2020).

The town possesses neither a sewer network nor a wastewater treatment facility. On-site sanitation facilities number around 2100 and predominantly comprise partially lined pit latrines that are elevated because of the shallow water table. The pits are not emptied: when full, faecal matter is moved from one pit to another, or a new pit is dug. Many pits also have overflow outlets in case of inundation during the rainy seasons (Nayebare et al., 2020).

2.2. Hydrological monitoring

Tipping bucket rainfall gauges, Lambrecht meteo model 15189 (Lambrecht meteo GmbH, Germany), were installed in two locations (1 and 2, Fig. 1B) and data were aggregated to daily sums. A rainfall timeseries for the study period was produced using the data from station 1, in the centre of Lukaya, unless records were absent or failed quality checks, in which case data were replaced using records from station 2. Groundwater levels in the weathered overburden were monitored in three boreholes screened at the following depth intervals: 10.2–16.1, 11.4–17.3, and 23.5–29.4 m bgl. All boreholes are located within 20 m of each other at the surface (Fig. 1B). Levels were monitored using Rugged TROLL 100 data loggers (In-Situ, USA).

2.3. Water sampling and analysis

2.3.1. Water sources and sampling rounds

An exhaustive survey of all water sources in the town was previously undertaken by Nayebare et al. (2020) and identified 56 shallow hand-dug wells equipped with hand pumps (shallow), 4 boreholes (deep) and 7 unprotected springs. The shallow sources vary between 3 and 8 m depth, and boreholes are at least 30 m deep, including the NWSC water supply well drilled to 61 m bgl (Nayebare et al., 2020). All shallow and deep sources are protected and considered *improved* water sources, and the springs are *unimproved* (WHO 2017c).

A stratified sampling approach was implemented to sample 40 of the water sources (Fig. 1B). The selected sources included: three deep sources, with the fourth having no accessible sampling location, and five springs, with the other two springs being such gentle seepages that groundwater inputs were not visible and these springs were not heavily utilised by the community. Finally, 32 of the 56 shallow sources were selected to maximise the spatial spread of shallow sources across the town, whilst accounting for some sources which had become non-functional.

Sources were sampled in six rounds (R) across 14 months. R1-4 were undertaken in 2018 from late-April to late-May when monthly rainfall typically peaks (Fig. 1C); each round was separated by six to nine days. R5 and R6 were undertaken in 2019 from mid- to late-June when monthly rainfall is close to its annual minimum and the rounds were separated by four days. In R5 and R6, four of the hand pumps on the shallow sources had become non-functional and only 36 sources were sampled. Note that rainfall had progressed through two wet and dry seasons between R4 and R5.

2.3.2. Water sampling and analysis

All shallow and deep water sources were unlocked and in frequent use by the community or owner throughout daylight hours. Nevertheless, all sources were allowed to flow for an additional minute before sampling to ensure all pipework was adequately flushed. All unprotected springs were sampled from the surface water channel as close to the point of groundwater discharge as possible.

Each source was sampled for a range of possible rapid indicators of faecal contamination. TLF and HLF were quantified using separate UviLux fluorimeters (Chelsea Technologies Limited, UK) targeting excitation-emission values of 280/360 (λ_{ex} / λ_{em}) and 280/450 nm (λ_{ex}/ λ_{em}), respectively. Whilst the HLF λ_{em} targeted the established peak, the λ_{ex} was matched to that of TLF to monitor the extent of optical overlap between the two regions. The bandpass filters for λ_{ex} and λ_{em} were ± 15 and ± 27.5 nm, respectively, for both fluorimeters (Figure S1). The TLF fluorimeter was calibrated using eight standards (0, 1, 2, 5, 10, 50, and 100 ppb) of L-tryptophan dissolved in ultrapure water. The factory calibration was implemented for the HLF sensor, which expresses intensity in quinine sulphate units (QSU). This is a standardised unit relating the fluorescence intensity at λ_{ex} 347.5 nm and λ_{em} 450 nm from 1 ppb of quinine sulphate dissolved in 0.105 M perchloric acid to direct calibration of the HLF sensor with pyrene tetrasulphonic acid in deionised water. The TLF ppb dissolved tryptophan data can be converted to QSU by division by 2.5037 or 2.3696 in rounds 1–4 and 5–6, respectively, to allow calculation of TLF:HLF ratios.

Fluorescence spectroscopy measurements were taken by submerging the fluorimeter in 150 mL of groundwater contained in a polypropylene beaker. Each measurement was taken in the dark by placing the beaker and fluorimeter within a covered stainless steel container. Given the sensitivity of the fluorimeters, all measurements were taken in duplicate, or repeated further to obtain reproducible data. Field repeatability (σ) of TLF and HLF measurements were calculated as 0.4 ppb and 0.1 QSU, respectively, across all data in R5 and 6. Specific electrical conductivity (SEC), pH and temperature were monitored using a multiparameter Manta-2 sonde (Eureka Waterprobes, USA). Turbidity was measured using a DR/890 portable colorimeter (HACH, USA), including blank correction with deionised water before each measurement, except during R4 when the Manta-2 was used. To account for absolute differences between the turbidimeters, the turbidity data were min-max normalised in each round. Fluorescence data did not require linear correction for temperature quenching (Khamis et al., 2015), with only a range between 22.4–25.6 °C. The pH of the samples was 4.6–6.7, hence pH would not appreciably have impacted the fluorescence (Reynolds 2003).

Sanitary risk inspections were undertaken at each source by the same assessor during sampling in R5 (WHO 2020). The surveys consisted of a list of nine yes-no questions (Table S1): to identify sources of contamination observable at the surface, pathways for contaminants to enter the source, and breakdowns in barriers to contamination (Kelly et al., 2020). The questions differed for the shallow/deep sources and springs because of different potential pathways leading to contamination. The total number of positive responses to the questions equates to the sanitary risk score (SRS).

Flow cytometry analysis for total (planktonic) bacteria cells (TBCs) was conducted in the laboratory on preserved samples, but the analysis can also be undertaken rapidly and online at a water source (Safford and Bischel 2019). Samples (2 mL) were collected in 4.5 mL polypropylene cryovials (STARLAB, UK) that were pre-loaded with the preservative glutaraldehyde (Sigma-Aldrich, UK) and the surfactant Pluronic F68 (Gibco, USA) (Marie et al., 2014) at final concentrations of 1% and 0.01%, respectively. The samples were kept in a cool box for up to 8 h, then frozen at −18 °C, defrosted overnight during transit to the UK in a cool box, and then analysed the following morning. Analysis was conducted using a BD Accuri C6 flow cytometer utilising a 488 nm solid state laser (Becton Dickinson UK Ltd., UK). Water samples (500 μL) were stained with a 1:50 v/v solution of SYBR Green I (Sigma-Aldrich, UK) to a final concentration of 1:10,000 v/v for 20 min in the dark at room temperature. Samples were run at a slow flow rate (14 mL/min, 10 mm core) for 5 min and a detection threshold of 1500 on channel FL1. A

single manually drawn gate was created to discriminate bacterial cells from particulate background, and cells per mL were calculated using the total cell count in 5 min divided by the reported volume run in μL (Sorensen et al., 2018a).

Thermotolerant (faecal) coliforms (TTCs) were selected as the FIO of contamination. TTCs include the preferred FIO *E. coli* (WHO 2017b), in addition to other genera such as *Klebsiella* spp. that are less likely to originate from a faecal source (Leclerc et al., 2001). Nevertheless, TTCs are considered acceptable FIO alternatives to *E. coli* by the WHO (2017b), as the majority of TTCs comprise *E. coli* in most circumstances. Indeed, 99% of TTCs were confirmed as *E. coli* in shallow groundwater contaminated by on-site sanitation in a similar climatological and hydrogeological setting in Kampala, Uganda (Howard et al., 2003). TTC samples were collected in sterile 250 mL polypropylene bottles and stored in a cool box (up to 8 h) before analysis. TTCs were isolated and enumerated using the membrane filtration method with Membrane Lauryl Sulphate Broth (MLSB, Oxoid Ltd, UK) as the selective medium (Sorensen et al., 2015a). Typically, 100 mL of the water sample was passed through a 0.45 μm cellulose nitrate filter (GE Whatman, UK). However, a smaller filtrate volume (1–50 mL) was used for a minority of samples to ensure colonies were not too numerous to count (TNTC), with the volume selected according to the corresponding TLF measurement and previous TTC analyses at the source. The filter was placed on an absorbent pad (Pall Gelman, Germany) saturated with MLSB broth in a plate and incubated at 44 °C for 18 –24 h in a Paqualab® 50 (ELE International, UK). Plates were inspected within 15 mins of removal from the incubator and all cream to yellow colonies greater than 1 mm considered TTCs. Where plates were TNTC the analysis was repeated the following day using a smaller volume of the remaining sample that had been kept refrigerated at 4 °C.

2.4. Statistical analysis and modelling

Rapid approaches to assess faecal contamination were tested against the benchmark FIO of TTCs using R v4.0.3 (R Core Team 2020) and base commands unless otherwise stated. Logistic regression models were developed for each rapid approach as a predictor of ≥10 cfu/100 mL TTCs. There were insufficient data ($n = 1$) where TTCs <1 cfu/100 mL to develop models for TTC presence-absence. Model performance was assessed using the area under the receive operating curve (AUC) (Mandrekar 2010), which is a plot of the proportion of true positive results against the proportion of false positive results as the threshold of the predictor is varied. A perfect classifier has an AUC of 1 and a random classifier has a value of 0.5. Furthermore, we consider AUC values of 0.7 to 0.8, 0.8 to 0.9, and 0.9 and greater as acceptable, excellent, and outstanding, respectively (Hosmer et al., 2013). Rank correlations between rapid approaches and TTCs were estimated using the non-parametric Spearman's rank (ρ_s) (Spearman 1904), given the non-Gaussian distribution of many of the variables. Coefficients of 0.80–1.00, 0.60–0.79, 0.40–0.59, 0.20–0.39, 0.00–0.19 were considered very strong, strong, moderate, weak, and very weak, respectively.

Multiple linear regression was applied to investigate what combination of rapid approaches was optimal for the prediction of TTC enumeration. A forward stepwise algorithm was used using 10-fold cross validation within the R package *car* (Fox and Weisberg 2018). One predictor is added to the model at a time to achieve the largest decrease in the root mean square error (RMSE), until no further reduction can be yielded. The normality of model residuals was evaluated using Q-Q plots. Initial models produced non-Gaussian residuals, in violation of the assumptions, so all variables with a skewness >1 were natural log transformed. An addition of 1 was made to TTCs to ensure the logarithm could be defined.

Differences in both rapid approaches and TTCs between sampling rounds were explored using the Friedman test in the R package *PMCMR* (Pohlert 2014), with post-hoc Nemenyi tests (Demšar 2006). The Friedman test is a non-parametric alternative to the repeated-measures

ANOVA and tests the null hypothesis that at least one group does not belong to the same population. If the Friedman test is significant ($p < 0.05$), the subsequent multiple comparison Nemenyi tests report significant ($p < 0.05$) differences between each pair of groups, if their corresponding mean ranks differ by at least the critical difference (Demšar 2006; Pohlert 2014).

The comparative resilience of rapid approaches and TTCs were evaluated by cross-correlating each variable with itself between sampling rounds. Additionally, TLF and HLF were cross-correlated with TTCs and TBCs across the sampling rounds to explore the seasonal nature of any associations. Spearman's Rank was used because the variables were non-Gaussian and because we were most interested in the rank-order of the sources as an indicator of relative risk across the community.

Groundwater levels (GWLs) for the most complete record, BH ALP-3, were hindcasted by 24 days to contextualise groundwater conditions before and during R1 and 2 where GWL observations were not collected. Hindcasting was conducted using a forward model implementing the water table fluctuation method and assuming diffuse recharge from daily rainfall observations (Cuthbert et al., 2019). The model was parameterised using a linear rainfall-recharge relationship with a rainfall threshold of 10 mm, an exponential recession coefficient of 1.1×10^{-3} day^{-1} to a base of 1149 m asl, and a specific yield of 5%. The model effectively captures which rainfall events result in groundwater recharge, the timing of GWL responses, and the rate of recession, with an r^2 of 0.85 and RMSE of 0.23 m (Figure S2).

3. Results

3.1. Widespread prevalence and high variability of TTCs

All sources show evidence of at least intermittent faecal contamination, inferred through the presence of TTCs (Fig. 2A&B). Fifty percent of the sources have median TTCs of at least 88 cfu/100 mL (Fig. 2A), with a range in median counts between <1 and 5101 cfu/100 mL. The shallow sources cover the entire range in median counts, whilst median
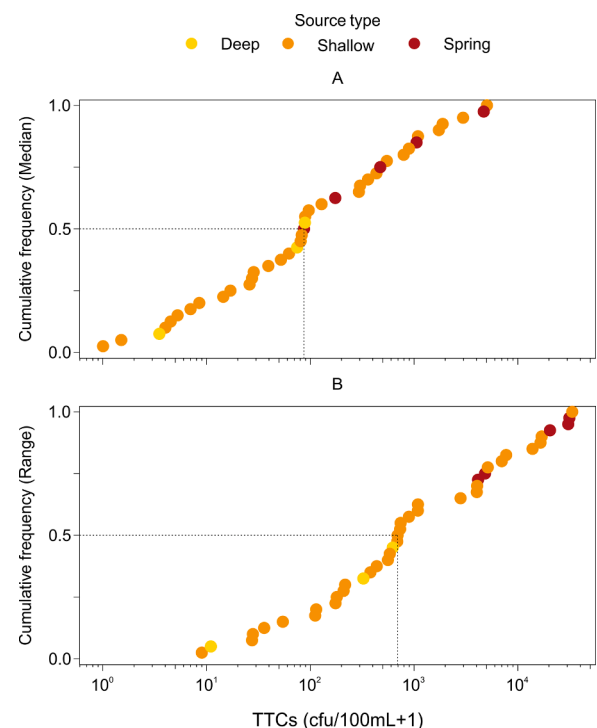

Fig. 2. Empirical cumulative distribution functions of (A) median TTCs and (B) range in TTCs for each water sources ($n >= 40$).

TTCs at springs and deep sources are in the upper 50% and lower 53% of all sources, respectively (Fig. 2A).

TTCs vary widely at each source with 50% of sources having a range of at least 720 cfu/100 mL (Fig. 2B). The range in TTCs at a source is at least 8 cfu/100 mL and up to 34,000 cfu/100 mL, with all but two sources varying between risk categories, based upon the order of magnitude of TTCs, previously defined by WHO (1997). A third of the sources transit between testing negative and positive for TTCs, including the only source with a median count of <1 cfu/100 mL (Fig. 3D). There is a tendency for spring sources to have greater ranges in TTCs, than the other types of source (Fig. 2B).

3.2. TLF and HLF are superior rapid approaches to indicate TTCs using source medians

Median TLF is the only significant predictor of median TTCs ≥10 cfu/100 mL according to logistic regression models (β = 1.09, p-value =

0.042) (Fig. 3A; Table S2). The AUC using TLF as a classifier is 0.88, which is closer to the perfect classifier value of 1 than the random selector of 0.5 (Fig. 3A) and considered "excellent". An optimal TLF threshold of 2.2 ppb can be defined to classify TTCs ≥10 cfu/100 mL, with associated false-negative and false-positive rates of 16% and 25%, respectively (Fig. 3C).

The AUC when classifying median TTCs ≥10 cfu/100 mL using HLF is 0.85 and considered "excellent", with the logistic regression model being borderline significant (β = 1.90, p-value 0.059) (Fig. 3A). A HLF threshold of 0.85 QSU can classify median TTCs of ≥10 cfu/100 mL with identical error rates to those as the proposed TLF threshold. In fact, if the TLF ppb threshold is converted into QSU then it is almost equivalent to the HLF threshold.

The AUC is "acceptable" for Sanitary risk scores (SRS) and median total bacterial cells (TBCs), and demonstrates that SEC and turbidity performed no better than a random classifier (Fig. 3A). Considering only the shallow sources (n = 32), SRS is a significant predictor (β = 0.90, p-value 0.038) with an "acceptable" AUC of 0.79. Only one individual sanitary inspection question, whether drainage was inadequate, is a significant predictor (p-value < 0.05) of median TTCs ≥10 cfu/100 mL for shallow and deep sources where sanitary inspection questions were identical. The AUC for inadequate drainage as a classifier is 0.75 and considered "acceptable" (F3).

Median TLF is very strongly correlated with median TTCs (ρ_s 0.81, Fig. 3D), being the most correlated rapid approach (Fig. 3B). All types of water source follow the same rising trend (Fig. 3D), with two notable outliers. One outlier is a shallow source, which has a median TLF of 65.9 ppb, more than three times the TLF intensity of any other source, although median TTCs are also high at 296 cfu/100 mL. The second outlier is a deep source with median TTCs of 89 cfu/100 mL, yet the lowest median TLF of 0.5 ppb, as well as being the only site with a zero SRS. Median HLF is similarly correlated with median TTCs (ρ_s 0.79) as TLF, with identical outliers. Median TBCs correlate moderately with median TTCs, but other rapid indicators are only weakly related to median TTCs (Fig. 3B).

No other rapid approaches provide additive performance to ln(TLF) for the prediction of median ln(TTCs) using the stepwise forward linear regression algorithm. The linear regression model has an r^2 of 0.51 and p-value <0.001 (Eq. (1)). Omitting ln(TLF), only ln(HLF) is included by the algorithm and the model has an r^2 of 0.48 and p-value <0.001 (Eq. (2)). Natural log transforms of TTCs, TLF, and HLF where required in the linear regression models to ensure the model residuals were Gaussian (see Figure S3 for Q-Q plots).

$$ln(TTC + 1) = 1.66 ln TLF + 2.19 \tag{1}$$

$$ln(TTC + 1) = 1.44 ln HLF + 3.73 \tag{2}$$

3.3. Relationships between rapid approaches and TTCs by sampling round

TTCs are significantly different between the wet season rounds of R1, R2 and R4 and the dry season rounds of R5 and R6 (Fig. 4A & B). Median TTCs were higher during than wet season (up to 382 cfu/100 mL, R1) than the dry season (as low as 13 cfu/100 mL, R5) (Fig. 4B). TTCs rapidly reduced in the absence of large rainfall events, for example, median TTCs reduced from 382 to 55 cfu/100 mL within 17–23 days between rounds R1 and R3. The two large successive daily rainfall events of 40 mm preceding round R4 resulted in substantial groundwater recharge, an almost five-fold increase in median TTCs to 262 cfu/100 mL, and increases in TTCs at 73% sources.

TLF shows a similar trend to TTCs across sampling rounds (Fig. 4B). Significant differences exist between wet and dry season rounds, with median TLF being highest in round R1 and lowest in rounds R5 and R6. HLF and TBCs also show significant differences only between wet and dry season rounds, with both at minima in the dry season. Turbidity and SEC show the least variability by sampling round, with fewest significant
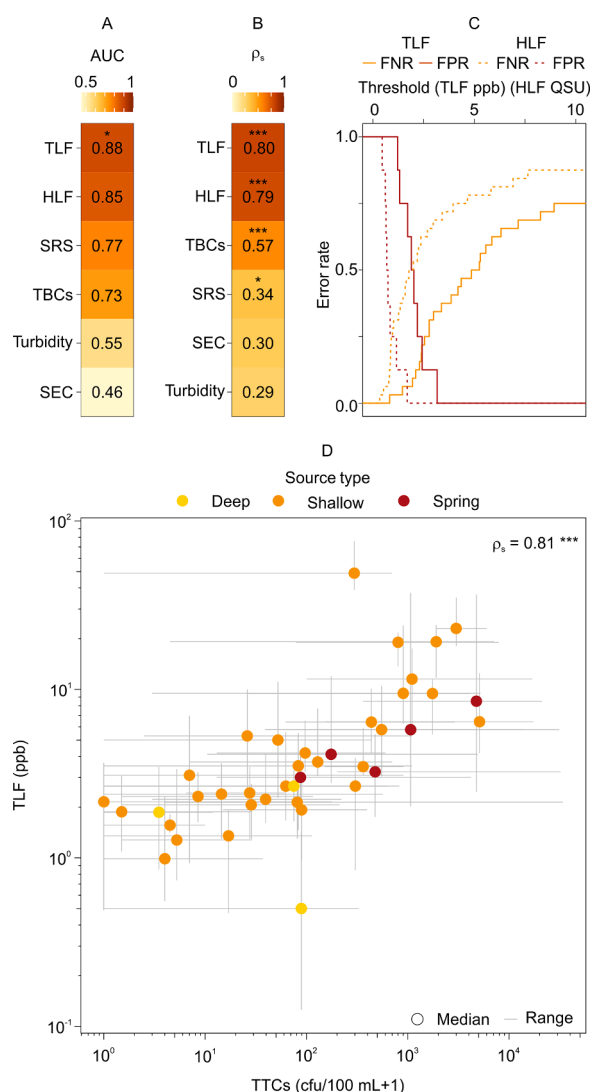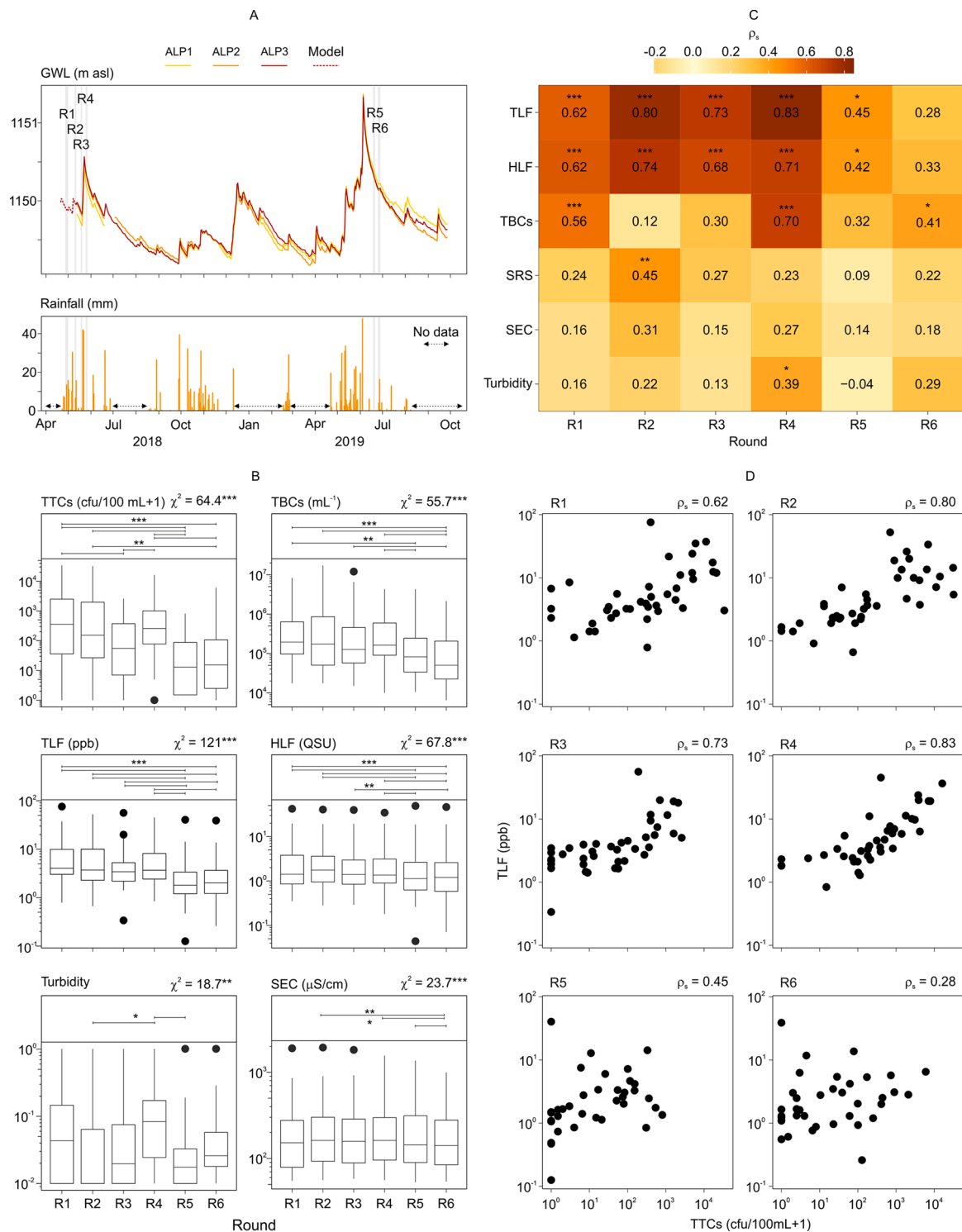
Fig. 3. (A) Area under curve (AUC) and significance of the logistic regression models for each rapid approach as a classifier of TTCs ≥10 cfu/100 mL; (B) Spearman's Rank correlation coefficients and significance for TTCs and each rapid approach; (C) False-negative (FNR) and false-positive (FPR) rates for TLF/HLF thresholds as classifiers of TTCs ≥10 cfu/100 mL; (D) Scatterplot of median TTCs and TLF for each source, illustrating ranges in both variables. The medians of rapid approaches and TTCs at each source (n = 40) are used in all statistics. p-values of <0.05, <0.01 and <0.001 are denoted by '*','**', and '***', respectively.

Fig. 4. (A) Relationship between groundwater levels (GWLs) and rainfall in Lukaya illustrating timing of all sampling rounds in grey; (B) Tukey boxplots of TTCs and rapid approaches by sampling round with χ^2 and significance of Friedman tests above each subplot and significant differences between rounds from post-hoc Nemenyi tests marked by ends of horizontal lines; (C) Spearman's Rank correlation coefficients and significance between TTCs and each rapid approach for all sampling rounds; (D) Scatterplots of TTCs and TLF for each sampling round with corresponding Spearman's Rank correlation coefficients shown. p-values of <0.05, <0.01 and <0.001 are denoted by '*','**', and '***', respectively.

differences between rounds.

TLF is generally the most strongly correlated rapid approach with TTCs in each sampling round (Fig. 4C). Positive correlations are very strong or strong during the wet season rounds R1–4, but only moderate or weak in dry season rounds R5 and R6, respectively (Fig. 4D). The strongest coefficient is during round R4, following the two large

successive rainfall events. Correlation coefficients between HLF and TTCs are similar or marginally lower, notably in round R4, than between TLF and TTCs, with significant correlations in all rounds (Fig. 4C), apart from R6 where significance is borderline (p = 0.051). TLF and HLF are also better predictors of TTCs ≥ 10 cfu/100 mL in the wet season rounds (mean AUC 0.84 and 0.73, respectively) than dry season rounds (mean AUC

AUC 0.68 for both).

TBCs are intermittently significantly correlated with TTCs, with strong (ρ_s 0.70) and moderate (ρ_s 0.56) correlations during rounds R4 and R1, respectively. Other rapid approaches are rarely significantly correlated with TTCs and coefficients are typically weak or very weak (Fig. 4C). Only TBCs in the dry season and SRS in the wet season have an AUC > 0.70 for classifying TTCs ≥10 cfu/100 mL (both mean seasonal AUC 0.71).

There are also notable associations between rapid approaches. During rounds R5 and R6 there is an almost perfect positive correlation between TLF and HLF (mean ρ_s 0.97); ρ_s remains very strong, but is lower in rounds R1–4 (mean r^2 0.88) (Figure S4). The TLF:HLF ratio is higher in rounds R1–4 (median 0.95) than R5–6 (median 0.70), with the percentage of samples having a ratio >1 also decreasing from 45 to 7%. The lower TLF:HLF ratio in R5–6 is a result of a greater reduction in TLF relative to HLF (Fig. 4C). In rounds R5 and R6, when the relationships between TLF/HLF and TTCs weaken, TLF and HLF are both strongly positively correlated with TBCs (mean ρ_s 0.62).

3.4. Cross-correlations between rapid approaches and TTCs across sampling rounds

There are very strong positive rank cross-correlations for both HLF and TLF between sampling rounds at the 36 sources, but rank cross-correlations are weaker and more varied for TTCs (Fig. 5A, B, D). The source rank-order by HLF is most consistent with a mean ρ_s of 0.91 (σ 0.04, all p-values <0.001) and, remarkably, a ρ_s of 0.95 between rounds R1 and R6 (Fig. 5B), separated by 14 months. The mean ρ_s for TLF is 0.86, with consistently very strong correlations between all rounds (σ 0.03, all p-values <0.001). The rank-order of sources by TTCs is inconsistent, moderately correlated on average (ρ_s mean 0.57, σ 0.11), but with only a weak correlation between rounds R1 and R6 (Fig. 5D). Bulk hydrochemistry rank-order of sources, as indicated by SEC, is also consistent between rounds (ρ_s mean 0.90, σ 0.07) (Fig. 5D), but SEC is unrelated to TTCs (Fig. 4C).

A survey of TLF or HLF across the community in either the dry or the wet season relates to TTCs during the wet season when TTCs are elevated. Ranking the sources based on HLF intensity during any sampling round correlates well (ρ_s mean 0.68, σ 0.06, all p-values <0.001) with the rank-order of the sources by TTCs during the wet season rounds R1–4 (Fig. 5F). TLF cross-correlates similarly to HLF with TTCs over the same time period (ρ_s mean 0.67, σ 0.09, 92% p-values <0.001), although some coefficients for round R1 with TTCs are weaker (Fig. 5E). Note, because of the very strong rank-correlation between HLF and TLF in the dry season rounds (mean ρ_s 0.97), both rank-correlate near-identically and strongly with TTCs during the wet season rounds. Importantly, dry season TLF and HLF (both mean ρ_s 0.68, σ 0.05, all p-values <0.001) both correlate more strongly than dry season TTCs (mean ρ_s 0.50, σ 0.10, 38% p-values <0.001) with wet season TTCs.

4. Discussion

4.1. In-situ TLF/HLF as rapid approaches to indicate faecal contamination

In our study, TLF/HLF are the superior rapid approaches to indicate faecal contamination of groundwater sources, as determined by TTCs. To set these results in a wider context, we re-analysed existing published datasets from contrasting hydrogeological settings following the same statistical approach (Fig. 6). The datasets were collated from: i) boreholes drilled to a consistent depth in an alluvial aquifer in Bihar, India (*n*
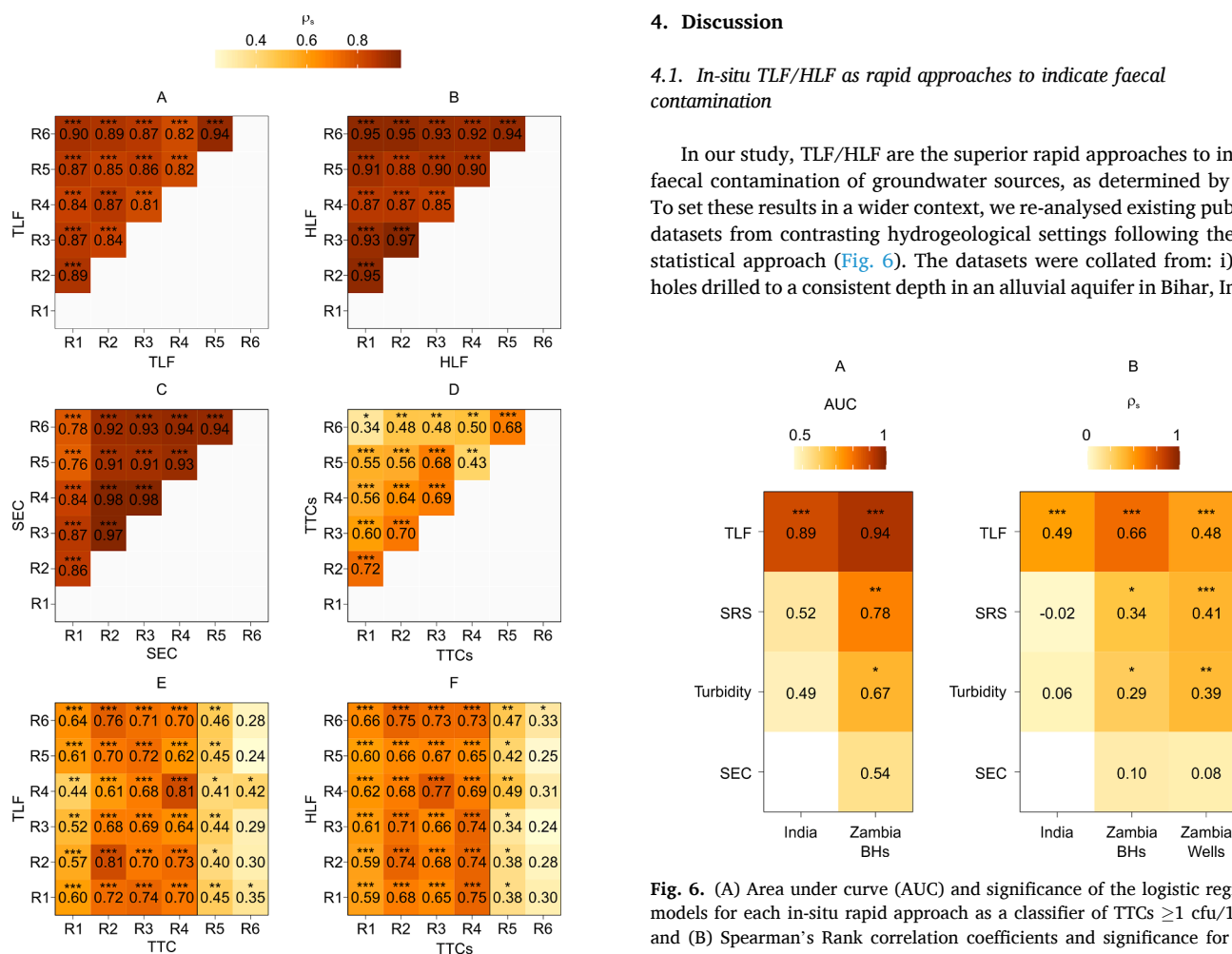

Fig. 5. Cross-correlations between variables in each sampling round illustrated by Spearman's Rank correlation coefficients for (A) TLF; (B) HLF; (C) SEC (D) TTCs; (E) TLF and TTCs; (F) HLF and TTCs (*n* = 36). p-values of <0.05, <0.01 and <0.001 are denoted by '*','**', and '***', respectively.
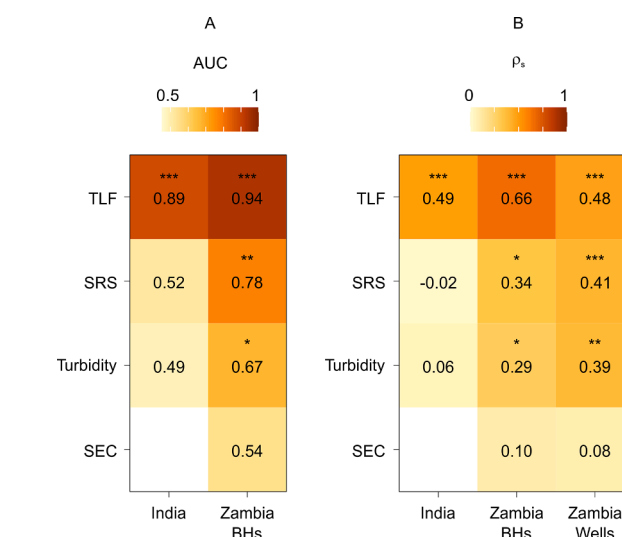

Fig. 6. (A) Area under curve (AUC) and significance of the logistic regression models for each in-situ rapid approach as a classifier of TTCs ≥1 cfu/100 mL and (B) Spearman's Rank correlation coefficients and significance for in-situ rapid approaches against TTCs. Data are from previous studies in India (Sorensen et al., 2016), and Zambia split by source type: borehole (BH) and Wells (Shallow well) (Sorensen et al., 2015a). AUC is not shown for Zambia Wells because of only five from 61 samples where TTCs <1 cfu/100 mL. p-values of <0.05, <0.01 and <0.001 are denoted by '*','**', and '***', respectively.

= 145) (Sorensen et al., 2016); and ii) boreholes ($n = 50$) and shallow hand-dug wells ($n = 61$) tapping either quartzite/dolomite or the overlying weathered saprolite/laterite, respectively, in Zambia (Sorensen et al., 2015a).

The re-analysis of these datasets demonstrates TLF is an effective significant predictor of the presence-absence of TTCs in a 100 mL sample in these other settings (Fig. 6A). Logistic regression models using TLF are significant ($p < 0.001$) and the AUC is 0.89–0.94. SRS and turbidity perform no better than a random classifier in India; whilst both are significant predictors ($p < 0.05$) in boreholes in Zambia, their AUCs are much lower than TLF (Fig. 6A). The shallow wells in Zambia were typically always contaminated with TTCs present in all but 5 of the 61 samples, so AUCs were not estimated.

TLF is the most correlated rapid indicator of the number of TTCs in our study, and the re-analysis of other published data (Fig. 6B). In India, there is a significant relationship between TLF and TTCs, but not between either SRS or turbidity and TTCs. In Zambia, TLF is strongly correlated with TTCs in boreholes, but only weak relationships exist between SRS or turbidity and TTCs. Correlation coefficients with TTCs also remain strongest for TLF, from the rapid approaches, in the shallow wells. An online application of TLF in groundwater-derived public water supplies in the UK has also demonstrated that TLF was better correlated (ρ_s 0.71) with *E. coli* than online turbidity (ρ_s 0.48) (Sorensen et al., 2018a).

There are alternative groundwater studies that have presented evidence that TLF has been unrelated to FIOs in groundwater. Nevertheless, TLF has still served as an effective in-situ indicator of contamination deriving from faecal matter in these studies. For example, Sorensen et al. (2020b) showed TLF was related to the density of on-site sanitation and associated nitrate but not TTCs beneath Dakar, Senegal. This study was also undertaken during the dry season and the results of our study suggests TLF/HLF relationships with FIOs are seasonal, and it is possible that during the wet season TLF/HLF could relate to FIOs in Dakar. Alternatively, the fluorophores in Dakar could relate to historic faecal contamination, as also observed at a source adjacent to an abandoned pit latrine in Malawi containing perennially high TLF, but sporadic and low TTC counts (Ward et al., 2021).

There remains inconsistent evidence regarding the use of turbidity and SEC (Buckerfield et al., 2019; Jung et al., 2014; Pronk et al., 2006; Pronk et al., 2009; Valenzuela et al., 2009), or sanitary inspections (Bain et al., 2014; Kelly et al., 2020; Misati et al., 2017) to determine faecal contamination risk in groundwater. Turbidity and SEC can derive from a variety of common sources, including the re-mobilisation of particles within the aquifer, and the relationship with faecal indicator bacteria in the literature is consequentially inconsistent (WHO 2017d). We consider that TLF/HLF are more appropriate indicators of variations in source water quality that relate to faecal contamination. A recent review by Kelly et al. (2020) suggested it was inappropriate to use sanitary inspections as indicators of microbial water quality. They argued that microbial samples from the same source are highly varied, whereas a sanitary inspection serves as a "lasting condition of the water source". Other limitations of SRS are they only represent conditions local to the source, whereas rapid subsurface transport of enteric pathogens can occur over large distances in fracture flow aquifers (Worthington and Smart 2017), and it is not possible to assess failure of the sanitary seal subsurface. Nevertheless, sanitary inspections are undoubtedly invaluable irrespective of whether they are indicative of microbial water quality, particularly as they provide information about potential risks and causes of contamination to inform interventions.

4.2. TLF and HLF are more resilient indicators of faecal contamination risk than TTCs

TLF and HLF are more resilient faecal contamination indicators in groundwater than TTCs within our study. We highlight comparable observations in Zambia where TLF remained elevated in several shallow sources over a period of four months, whereas TTCs were only elevated in the wet season (Sorensen et al., 2015a); this dynamic was also recently suggested at five water sources in Malawi by Ward et al. (2021). Despite TLF remaining seasonally elevated in several Zambian sources, there was an overall trend towards higher median TLF and TTCs in the wet season (7.1 ppb and 48 cfu/100 mL) relative to the dry season (2.8 ppb and 2 cfu/100 mL). *Re*-analysis of the Zambia data demonstrates that the relationship between TLF and TTCs is stronger in the wet (ρ 0.82) than the dry season (ρ 0.67). Moreover, there is also a stronger cross-correlation between dry season TLF and elevated wet season TTCs (ρ 0.80), than dry and wet season TTCs (ρ 0.60). In summary, both TLF and TTC vary seasonally in Uganda and Zambia, but ranking the sources within a community by faecal contamination risk using TLF is a more temporally robust approach than using TTCs.

Contrasting seasonal variations and relationships between TLF/HLF and TTCs suggest TLF/HLF differ from TTCs in one or more properties: (i) their source, (ii) their transport properties, and/or (iii) their persistence in the subsurface. The dominant source term for both types of faecal indicator is likely to be effluent from on-site sanitation in urban low-income settings, where present, except potentially where there are naturally high levels of sedimentary fluorescent NOM, or water is contaminated with fluorescent xenobiotic compounds, such as diesel (Carstea et al., 2010). Inputs from on-site sanitation are likely to be greatest during the wet season, particularly following large rainfall events, when latrines can be inundated and overflow (Nayebare et al., 2020), and accumulated faecal matter on the ground surface can be mobilised (Howard et al., 2003). There is also likely to be a continuous input function from on-site sanitation, as pit latrines and septic tanks, leak year-round.

The faecal indicators have different transport properties. Frank et al. (2021) demonstrated that dissolved tryptophan was comparable in transit time and recovery to the conservative dye tracer uranine over short distance, <2 h tracer tests, with no evidence of retardation. Frank et al. also demonstrated similar recovery for a humic acid, although there was some evidence of retardation and the tracer peak was marginally delayed by five minutes, in comparison to uranine. It is unclear what proportion of TLF/HLF can be attributed to dissolved pure tryptophan or the humic acid used in any given setting, although TLF/HLF fluorophores are predominantly extracellular in groundwater (Sorensen et al., 2020a). Nevertheless, there will also be an element of sorption and desorption of dissolved OM between groundwater and the aquifer matrix and soils (Shen et al., 2015), particularly for more hydrophobic molecules, which may have a TLF/HLF component, as well as a minor component contained within cells. TTCs can be transported more rapidly than solutes, notably in heterogeneous media such as weather crystalline rocks, but are subject to appreciable attenuation (Taylor et al., 2004). TTCs tend to accumulate and be transported laterally when flow velocities increase (WHO 2017b), such as during a rainfall event generating groundwater recharge. Therefore, TLF/HLF fluorophores are likely to be more readily and continuously transported than TTCs in groundwater.

The persistence of TLF/HLF fluorophores and TTCs are likely to differ in groundwater. HLF is expected to be the most persistent indicator, demonstrating the strongest rank-order cross-correlation between sampling rounds. Furthermore, although HLF decreases in the dry season, there is a proportionally greater loss in TLF indicating either preferential breakdown or more efficient lateral transport of TLF fluorophores. HLF has been demonstrated to be more recalcitrant, resistant to breakdown, than TLF in surface water and wastewater (Cory and Kaplan 2012; Ignatev and Tuhkanen 2019) and a greater proportion of fluorophores are like to be recalcitrant in groundwater where NOM is typically less bioavailable (Chapelle 2021; Shen et al., 2015). The refractory nature of some HLF fluorophores led Zheng et al. (2020) to suggest that HLF is an effective tracer of wastewater in groundwater. There is also potential for the in-situ production of TLF or HLF from NOM entering an aquifer system (Fox et al., 2017; Yang et al., 2020).

Therefore, fluorophore persistence in the subsurface could be a result of the continuous recycling and microbial transformation of NOM arriving in the system as opposed to the accumulation of recalcitrant molecules (Benk et al., 2019; Roth et al., 2019). The dry season relationships between TLF/HLF and TBCs when faecal inputs are more limited, as also observed in Senegal (Sorensen et al., 2020b), suggest bacteria are using the NOM as a substrate and potentially generating fluorophores in-situ. Irrespective of the relative persistence of either TLF or HLF, either wavelength pair would provide a similar indicator of faecal contamination risk given their co-correlation in our study and the optical overlap between the peaks. TTCs are generally only considered indicative of recent contamination with die-off within 16–45 days (Taylor et al., 2004), in contrast to the more persistent fluorescence indicators.

The more efficient transport of TLF/HLF fluorophores and their greater persistence in the subsurface in comparison to TTCs could explain why dry season TLF/HLF relates to wet season TTCs. Firstly, more efficient transport could facilitate the perennial transport of TLF/HLF fluorophores from a faecal source to a water source, whereas TTCs are predominantly mobilised following rainfall in the wet season. It should also be re-iterated that TLF/HLF does also respond to rainfall with higher intensity in the wet season, indicating higher seasonal risks. Secondly, faecal contamination events at a water source would remain detectable for a longer period by fluorescing more persistent TLF/HLF fluorophores than TTCs. If these events are focussed in the wet season, as observed here, then the proportion of TLF/HLF persisting into the dry season may relate to wet season TTCs, given the two types of indicator correlate very strongly after heavy rainfall (e.g. TLF, ρ_s 0.83, R4).

4.3. Remaining uncertainties, instrumentation improvements, and future work

There are a range of potential interferents with in-situ fluorescence measurements that are discussed in a review by Carstea et al. (2020) but these have not adversely impacted previous TLF-FIO studies (Sorensen et al., 2018b) or this study across a range of settings. Corrections for temperature, turbidity, and absorbance of light by the sample matrix (the inner-filtering effect) are not likely to be necessary in the majority of groundwater settings (Khamis et al., 2015; Sorensen et al., 2015a). Moreover, the next generation of commercially available portable fluorimeters are now capable of automatic corrections. pH does not have an appreciable impact on TLF/HLF between values of 5 and 8 (Reynolds 2003; Spencer et al., 2007), and groundwater outside this range is unlikely to be suitable for drinking. High concentrations of metal ions could quench fluorescence (Yang et al., 2018), which is most likely where water is contaminated by mining and industry. Certain water treatments, including chlorination, also quench fluorescence (Henderson et al., 2009) so the well owner or other informed individuals should be interviewed to assess if the water has been treated prior to testing, as would be undertaken before FIO sampling. Alternatively, a chlorine residual test could be performed.

There is the potential for TLF or HLF fluorophores to originate from contamination unrelated to faecal sources such as diesel and fuel derivatives, food waste, paper mills, and pesticides (Carstea et al., 2016). In these instances, a source displaying high TLF/HLF should still be considered a higher faecal contamination risk than one displaying low TLF/HLF, as there would be evidence that a pathway is present to a source of anthropogenic waste. Sedimentary fluorescent NOM contained within the aquifer could also potentially be problematic when comparing faecal contamination risks determined by TLF/HLF, particularly between study areas. In which case, deviation from baseline fluorescence intensity in uncontaminated sources would be more important than the absolute value for determining risk.

Relatively high upfront costs undoubtedly constrain widespread adoption of in-situ fluorescence spectroscopy. The present generation of single peak fluorimeters cost in the region of US$5000–7000, before considering accessories that can cost a further US$2000–3000

(Sorensen et al., 2018b). However, there is substantial scope to reduce these costs through the development of lower-cost portable fluorimeters, engineered specifically to provide an in-situ indication of faecal contamination risk at a water source. Multiple researchers have developed prototype fluorimeters with various benefits over commercial alternatives (Bedell et al., 2020; Bridgeman et al., 2015; Simões et al., 2021), but field validation and a discussion of indicative costs are absent or limited. As part of our study, we successfully developed and demonstrated the efficacy of a lower-cost prototype portable multi-wavelength LED-based fluorimeter on duplicate samples in rounds R5 and R6 (Figure S5). The prototype provided comparable results to the UviLux sensors in both the laboratory (Table S5) and field. For example, the prototype derived HLF data from R5 and R6 both correlate strongly with TTCs in R4 (mean ρ_s 0.69). Therefore, a low-cost, portable fluorimeter to indicate faecal contamination risk could be produced for a total component cost of $1100. Further details are provided in the supplementary information (S1). In addition to reducing costs, future development should investigate the production of low-cost sealed long-life containers of TLF/HLF standards. These containers would enable calibration checks, ideally annually, and negative controls to be performed by the end-user without return to the manufacturer or access to a well-equipped laboratory with reagents and high quality deionised water (Sorensen et al., 2018b).

It remains unclear how TLF/HLF relate to the presence of enteric pathogens or risks posed to human health. There is one published study showing a relationship between TLF and DNA markers of enteric pathogens, although this study is limited to 22 sources in one town (Sorensen et al., 2015b). Future work should explore the potential link between TLF/HLF and enteric pathogens using molecular approaches, as well as exploring the viability of pathogens where possible. Furthermore, studies should investigate if and how TLF/HLF could effectively be used for on-site risk communication to induce behavioural change in communities and reduce the disease burden relating to the consumption of faecally contamination drinking water.

5. Conclusions

In-situ fluorescence spectroscopy provides an instantaneous assessment of water source quality that relates to faecal contamination risk determined by faecal indicator organisms (FIOs). Consequently, faecal contamination risks can be assessed immediately, including in real-time, and could be communicated on-site to consumers to reduce exposure to contamination, whilst confirmative regulatory FIO analysis is undertaken. Furthermore, in-situ fluorescence can extend FIO sampling programs because data can be collected rapidly by users who require minimal training; nor are there consumable costs for additional samples.

TLF and HLF are more resilient indicators of faecal contamination risk than FIOs. Both types of indicator respond to rainfall and contamination events, with the strongest relationships between the indicators observed in the wet season, notably immediately after heavy rainfall (e.g. TLF-TTC ρ_s 0.83). However, ranking the sources across a community by risk using FIOs is more variable (cross-correlations ρ_s 0.34–72) between sampling rounds, than using TLF or HLF (cross-correlations ρ_s 0.81–97). This ranking of sources using TLF/HLF at any point in time relates to TTCs during the wet season, when TTCs are significantly elevated and risks to human health would consequently also be expected to be greatest. Furthermore, the source rank-orders in the dry season using TLF/HLF cross-correlate more strongly (both mean ρ_s 0.68) with wet season TTCs than dry season TTCs (mean ρ_s 0.50). Therefore, the comparative faecal contamination risks between sources generated by a dry season survey of TLF/HLF would be more accurate than using highly transient FIOs to indicate the comparative risks occurring in the wet season when risks are elevated. This characteristic is advantageous given water quality surveys are infrequent for private water supplies globally, as well as across low-income countries. TLF/HLF provide a more repeatable and temporally robust approach than FIOs to ranking

sources by faecal contamination risk across a community to strategise prioritisation of sources for drinking or interventions.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2021.117734.

References

Bain, R., Cronk, R., Wright, J., Yang, H., Slaymaker, T., Bartram, J., 2014. Fecal contamination of drinking-water in low-and middle-income countries: a systematic review and meta-analysis. PLoS Med. 11 (5), e1001644.

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

Baker, A., 2002. Fluorescence properties of some farm wastes: implications for water quality monitoring. Water Res. 36 (1), 189–195.

Baker, A., Cumberland, S.A., Bradley, C., Buckley, C., Bridgeman, J., 2015. To what extent can portable fluorescence spectroscopy be used in the real-time assessment of microbial water quality? Sci. Total Environ. 532, 14–19.

Baker, A., Inverarity, R., 2004. Protein-like fluorescence intensity as a possible tool for determining river water quality. Hydrol. Process. 18 (15), 2927–2945.

Bedell, E., Sharpe, T., Purvis, T., Brown, J., Thomas, E., 2020. Demonstration of tryptophan-like fluorescence sensor concepts for fecal exposure detection in drinking water in remote and resource constrained settings. Sustainability 12 (9), 3768.

Benk, S.A., Yan, L., Lehmann, R., Roth, V.-N., Schwab, V.F., Totsche, K.U., Küsel, K., Gleixner, G., 2019. Fueling diversity in the subsurface: composition and age of dissolved organic matter in the critical zone. Front. Earth Sci. 7, 296.

Bieroza, M., Baker, A., Bridgeman, J., 2009. Relating freshwater organic matter fluorescence to organic carbon removal efficiency in drinking water treatment. Sci. Total Environ. 407 (5), 1765–1774.

Bridgeman, J., Baker, A., Brown, D., Boxall, J., 2015. Portable LED fluorescence instrumentation for the rapid assessment of potable water quality. Sci. Total Environ. 524, 338–346.

Buckerfield, S.J., Quilliam, R.S., Waldron, S., Naylor, L.A., Li, S., Oliver, D.M., 2019. Rainfall-driven E. coli transfer to the stream-conduit network observed through increasing spatial scales in mixed land-use paddy farming karst terrain. Water research X 5, 100038.

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.

Carstea, E.M., Baker, A., Bieroza, M., Reynolds, D, 2010. Continuous fluorescence excitation–emission matrix monitoring of river organic matter. Water Res. 44 (18), 5356–5366.

Carstea, E.M., Bridgeman, J., Baker, A., Reynolds, D.M., 2016. Fluorescence spectroscopy for wastewater monitoring: a review. Water Res. 95, 205–219.

Carstea, E.M., Popa, C.L., Baker, A., Bridgeman, J., 2020. In situ fluorescence measurements of dissolved organic matter: a review. Sci. Total Environ. 699, 134361.

Chapelle, F.H., 2021. The bioavailability of dissolved, particulate, and adsorbed organic carbon in groundwater systems. Groundwater 59 (2), 226–235.

Coble, P.G., Lead, J., Baker, A., Reynolds, D.M., Spencer, R.G., 2014. Aquatic Organic Matter Fluorescence. Cambridge University Press.

Collier, S.A., Deng, L., Adam, E.A., Benedict, K.M., Beshearse, E.M., Blackstock, A.J., Bruce, B.B., Derado, G., Edens, C., Fullerton, K.E., 2021. Estimate of burden and direct healthcare cost of infectious waterborne disease in the United States. Emerging Infect. Dis. 27 (1), 140.

Copernicus Sentinel data (2020) Scene S2A_MSIL2A_20200209T080051_N0214_R035_T36MUE_20200209T110919. Available from: https://scihub.copernicus.eu/dhus/#/home.

Cory, R.M., Kaplan, L.A., 2012. Biological lability of streamwater fluorescent dissolved organic matter. Limnol. Oceanogr. 57 (5), 1347–1360.

Cuthbert, M.O., Taylor, R.G., Favreau, G., Todd, M.C., Shamsudduha, M., Villholth, K.G., MacDonald, A.M., Scanlon, B.R., Kotchoni, D.V., Vouillamoz, J.-.M., 2019. Observed controls on resilience of groundwater to climate variability in sub-Saharan Africa. Nature 572 (7768), 230–234.

Dalterio, R., Nelson, W., Britt, D., Sperry, J., Psaras, D., Tanguay, J., Suib, S., 1986. Steady-state and decay characteristics of protein tryptophan fluorescence from bacteria. Appl. Spectrosc. 40 (1), 86–90.

Demšar, J., 2006. Statistical comparisons of classifiers over multiple data sets. J. Mach. Learn. Res. 7, 1–30.

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

Fellman, J.B., Hood, E., Spencer, R.G., 2010. Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: a review. Limnol. Oceanogr. 55 (6), 2452–2462.

Fox, B., Thorn, R., Anesio, A., Reynolds, D, 2017. The in situ bacterial production of fluorescent organic matter; an investigation at a species level. Water Res.

Fox, J., Weisberg, S., 2018. An R companion to Applied Regression. Sage publications.

Frank, S., Goeppert, N., Goldscheider, N., 2017. Fluorescence-based multi-parameter approach to characterize dynamics of organic carbon, faecal bacteria and particles at alpine karst springs. Sci. Total Environ.

Frank, S., Goeppert, N., Goldscheider, N., 2021. Field tracer tests to evaluate transport properties of tryptophan and humic acid in karst. Groundwater 59 (1), 59–70.

Gleeson, T., VanderSteen, J., Sophocleous, M.A., Taniguchi, M., Alley, W.M., Allen, D.M., Zhou, Y., 2010. Groundwater sustainability strategies. Nat. Geosci. 3 (6), 378–379.

Goldman, J.H., Rounds, S.A., Needoba, J.A., 2012. Applications of fluorescence spectroscopy for predicting percent wastewater in an urban stream. Environ. Sci. Technol. 46 (8), 4374–4381.

Harris, I., Osborn, T.J., Jones, P., Lister, D, 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. Data 7 (1), 1–18.

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.

Hosmer Jr, D.W., Lemeshow, S., Sturdivant, R.X, 2013. Applied Logistic Regression. John Wiley & Sons.

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.

Hrudey, S.E., Hrudey, E.J., 2004. Safe Drinking Water. IWA publishing.

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.

Hunter, P.R., MacDonald, A.M., Carter, R.C., 2010. Water supply and health. PLoS Med. 7 (11), e1000361.

Hur, J., Lee, B.-.M., Lee, T.-.H., Park, D.-.H., 2010. Estimation of biological oxygen demand and chemical oxygen demand for combined sewer systems using synchronous fluorescence spectra. Sensors 10 (4), 2460–2471.

Ignatev, A., Tuhkanen, T., 2019. Monitoring WWTP performance using size-exclusion chromatography with simultaneous UV and fluorescence detection to track recalcitrant wastewater fractions. Chemosphere 214, 587–597.

Jung, A.-.V., Le Cann, P., Roig, B., Thomas, O., Baurès, E., Thomas, M.-.F., 2014. Microbial contamination detection in water resources: interest of current optical methods, trends and needs in the context of climate change. Int. J. Environ. Res. Public Health 11 (4), 4292–4310.

Kelly, E.R., Cronk, R., Kumpel, E., Howard, G., Bartram, J., 2020. How we assess water safety: a critical review of sanitary inspection and water quality analysis. Sci. Total Environ. 718, 137237.

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

Kida, M., Kojima, T., Tanabe, Y., Hayashi, K., Kudoh, S., Maie, N., Fujitake, N., 2019. Origin, distributions, and environmental significance of ubiquitous humic-like fluorophores in Antarctic lakes and streams. Water Res. 163, 114901.

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.

Mandrekar, J.N., 2010. Receiver operating characteristic curve in diagnostic test assessment. J. Thorac. Oncol. 5 (9), 1315–1316.

Marie, D., Rigaut-Jalabert, F., Vaulot, D, 2014. An improved protocol for flow cytometry analysis of phytoplankton cultures and natural samples. Cytom. Part A 85 (11), 962–968.

Mendoza, L.M., Mladenov, N., Kinoshita, A.M., Pinongcos, F., Verbyla, M.E., Gersberg, R., 2020. Fluorescence-based monitoring of anthropogenic pollutant

inputs to an urban stream in Southern California. USA. Sci. Total Environ. 718, 137206.

Misati, A.G., Ogendi, G., Peletz, R., Khush, R., Kumpel, E., 2017. Can sanitary surveys replace water quality testing? Evidence from Kisii, Kenya. Int. J. Environ. Res. Public Health 14 (2), 152.

Nayebare, J., 2021. Loading and Mobility of Faecal and Chemical Contaminants to Shallow groundwater: Evidence from Lukaya Town. Central Uganda, Makerere University.

Nayebare, J., Owor, M., Kulabako, R., Campos, L., Fottrell, E., Taylor, R., 2020. WASH conditions in a small town in Uganda: how safe are on-site facilities? J. Water Sanit. Hyg. Dev. 10 (1), 96–110.

Nowicki, S., Lapworth, D.J., Ward, J.S., Thomson, P., Charles, K., 2019. Tryptophan-like fluorescence as a measure of microbial contamination risk in groundwater. Sci. Total Environ. 646, 782–791.

PDP, 2017. Kalungu District. Final Physical Development Plan Report. In: Kalungu District. Final Physical Development Plan Report, 2017–2027, p. 256.

Pohlert, T., 2014. The pairwise multiple comparison of mean ranks package (PMCMR). R package 27 (2019), 9.

Pronk, M., Goldscheider, N., Zopfi, J., 2006. Dynamics and interaction of organic carbon, turbidity and bacteria in a karst aquifer system. Hydrogeol. J. 14 (4), 473–484.

Pronk, M., Goldscheider, N., Zopfi, J., Zwahlen, F., 2009. Percolation and particle transport in the unsaturated zone of a karst aquifer. Groundwater 47 (3), 361–369.

R Core Team, 2020. R: A language and Environment For Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL. https://www.R-project.org.

Reynolds, D., Ahmad, S., 1997. Rapid and direct determination of wastewater BOD values using a fluorescence technique. Water Res. 31 (8), 2012–2018.

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

Roth, V.-.N., Lange, M., Simon, C., Hertkorn, N., Bucher, S., Goodall, T., Griffiths, R.I., Mellado-Vázquez, P.G., Mommer, L., Oram, N.J., 2019. Persistence of dissolved organic matter explained by molecular changes during its passage through soil. Nat. Geosci. 12 (9), 755–761.

Safford, H.R., Bischel, H.N., 2019. Flow cytometry applications in water treatment, distribution, and reuse: a review. Water Res. 151, 110–133.

Shen, Y., Chapelle, F.H., Strom, E.W., Benner, R., 2015. Origins and bioavailability of dissolved organic matter in groundwater. Biogeochemistry 122 (1), 61–78.

Sihan, P., Xiaowen, C., Chenyue, C., Zhuo, C., NGO, H.H., Qi, S., Wenshan, G., Hong-Ying, H., 2021. Fluorescence analysis of centralized water supply systems: indications for rapid cross-connection detection and water quality safety guarantee. Chemosphere, 130290.

Simões, J., Yang, Z., Dong, T., 2021. An ultrasensitive fluorimetric sensor for pre-screening of water microbial contamination risk. Spectrochim. Acta A Mol. Biomol. Spectrosc. 258, 119805.

Sorensen, J., Lapworth, D., Marchant, B., Nkhuwa, D., Pedley, S., Stuart, M., Bell, R., Chirwa, M., Kabika, J., Liemisa, M., 2015a. In-situ tryptophan-like fluorescence: a real-time indicator of faecal contamination in drinking water supplies. Water Res. 81, 38–46.

Sorensen, J., Lapworth, D., Read, D., Nkhuwa, D., Bell, R., Chibesa, M., Chirwa, M., Kabika, J., Liemisa, M., Pedley, S., 2015b. Tracing enteric pathogen contamination in sub-Saharan African groundwater. Sci. Total Environ. 538, 888–895.

Sorensen, J., Sadhu, A., Sampath, G., Sugden, S., Gupta, S.D., Lapworth, D., Marchant, B., Pedley, S., 2016. Are sanitation interventions a threat to drinking water supplies in rural India? An application of tryptophan-like fluorescence. Water Res. 88, 923–932.

Sorensen, J., Vivanco, A., Ascott, M., Gooddy, D., Lapworth, D., Read, D., Rushworth, C., Bucknall, J., Herbert, K., Karapanos, I., 2018a. Online fluorescence spectroscopy for the real-time evaluation of the microbial quality of drinking water. Water Res. 137, 301–309.

Sorensen, J.P., Baker, A., Cumberland, S.A., Lapworth, D.J., MacDonald, A.M., Pedley, S., Taylor, R.G., Ward, J.S., 2018b. Real-time detection of faecally contaminated drinking water with tryptophan-like fluorescence: defining threshold values. Sci. Total Environ. 622, 1250–1257.

Sorensen, J.P., Carr, A.F., Nayebare, J., Diongue, D.M., Pouye, A., Roffo, R., Gwengweya, G., Ward, J.S., Kanoti, J., Okotto-Okotto, J., 2020a. Tryptophan-like

and humic-like fluorophores are extracellular in groundwater: implications as real-time faecal indicators. Sci. Rep. 10 (1), 1–9.

Sorensen, J.P., Diaw, M.T., Pouye, A., Roffo, R., Diongue, D.M., Faye, S.C., Gaye, C.B., Fox, B.G., Goodall, T., Lapworth, D.J., 2020b. In-situ fluorescence spectroscopy indicates total bacterial abundance and dissolved organic carbon. Sci. Total Environ., 139419

Spearman, C., 1904. The proof and measurement of association between two things. Am. J. Psychol. 15 (1), 72–101.

Spencer, R.G., Bolton, L., Baker, A., 2007. Freeze/thaw and pH effects on freshwater dissolved organic matter fluorescence and absorbance properties from a number of UK locations. Water Res. 41 (13), 2941–2950.

Stedmon, C.A., Seredyńska-Sobecka, B., Boe-Hansen, R., Le Tallec, N., Waul, C.K., Arvin, E., 2011. A potential approach for monitoring drinking water quality from groundwater systems using organic matter fluorescence as an early warning for contamination events. Water Res. 45 (18), 6030–6038.

Stelma Jr,, G.N., Wymer, L.J, 2012. Research considerations for more effective groundwater monitoring. J. Water Health 10 (4), 511–521.

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.

UBOS (2014) Central Region - Parish Level Profiles (Census 2014). Available from: https://www.ubos.org/explore-statistics/20/.

UNICEF/WHO (2017) Target product profile: rapid E. coli detection. Version 2.0.

Valenzuela, M., Lagos, B., Claret, M., Mondaca, M.A., Pérez, C., Parra, O., 2009. Fecal contamination of groundwater in a small rural dryland watershed in central Chile. Chil. J. Agric. Res. 69 (2), 235–243.

Ward, J.S., Lapworth, D.J., Read, D.S., Pedley, S., Banda, S.T., Monjerezi, M., Gwengweya, G., MacDonald, A.M., 2020. Large-scale survey of seasonal drinking water quality in Malawi using in situ tryptophan-like fluorescence and conventional water quality indicators. Sci. Total Environ. 744, 140674.

Ward, J.S., Lapworth, D.J., Read, D.S., Pedley, S., Banda, S.T., Monjerezi, M., Gwengweya, G., MacDonald, A.M., 2021. Tryptophan-like fluorescence as a high-level screening tool for detecting microbial contamination in drinking water. Sci. Total Environ. 750, 141284.

WHO, 1997. Guidelines For Drinking-Water Quality, 2nd Edition. WHO Press, Geneva, Switzerland.

WHO, 2017a. Drinking Water Parameter Cooperation Project: Support to the Revision of Annex I Council Directive 98/83/EC On the Quality of Water Intended For Human Consumption (Drinking Water Directive). WHO, Geneva, Switzerland.

WHO (2017b) Guidelines for drinking-water quality (4th edition), WHO Press, Geneva, Switzerland.

WHO (2017c) Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines.

WHO (2017d) Water quality and health-review of turbidity: information for regulators and water suppliers.

WHO (2019) Drinking-water. https://www.who.int/news-room/fact-sheets/detail/drinking-water.

WHO (2020) Sanitary Inspection Form. Available from https://www.who.int/water_sanitation_health/water-quality/safety-planning/sanitary-inspection-packages-for-drinking-water/en/.

Worthington, S.R., Smart, C.C., 2017. Transient bacterial contamination of the dual-porosity aquifer at Walkerton. Hydrogeol. J. 25 (4), 1003–1016.

Yang, Y., Ma, X., Yang, X., Xu, H., 2018. Influence of heavy metal ions on the spectra and charge characteristics of DOM of municipal sewage secondary effluent. Water Sci. Technol. 77 (4), 1098–1106.

Yang, Y., Yuan, X., Deng, Y., Xie, X., Gan, Y., Wang, Y., 2020. Seasonal dynamics of dissolved organic matter in high arsenic shallow groundwater systems. J. Hydrol. (Amst) 589, 125120.

Zheng, Y., He, W., Li, B., Hur, J., Guo, H., Li, X., 2020. Refractory humic-like substances: tracking environmental impacts of anthropogenic groundwater recharge. Environ. Sci. Technol. 54 (24), 15778–15788.

Zhou, Y., Jeppesen, E., Zhang, Y., Shi, K., Liu, X., Zhu, G., 2016. Dissolved organic matter fluorescence at wavelength 275/342 nm as a key indicator for detection of point-source contamination in a large Chinese drinking water lake. Chemosphere 144, 503–509.