Monitoring groundwater quality with real-time data, stable water isotopes, and microbial community analysis : A comparison with conventional methods

Lyons, Kevin J.

2023-03-15

Lyons , K J , Ikonen , J , Hokajärvi , A-M , Räsänen , T , Pitkänen , T , Kauppinen , A , Kujala , K , Rossi , P M & Miettinen , I T 2023 , ' Monitoring groundwater quality with real-time data, stable water isotopes, and microbial community analysis : A comparison with conventional methods ' , Science of the Total Environment , vol. 864 , 161199 . https://doi.org/10.1016/j.scitotenv.2022.161199

http://hdl.handle.net/10138/356347
https://doi.org/10.1016/j.scitotenv.2022.161199

Science of the Total Environment 864 (2023) 161199

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Monitoring groundwater quality with real-time data, stable water isotopes, and microbial community analysis: A comparison with conventional methods

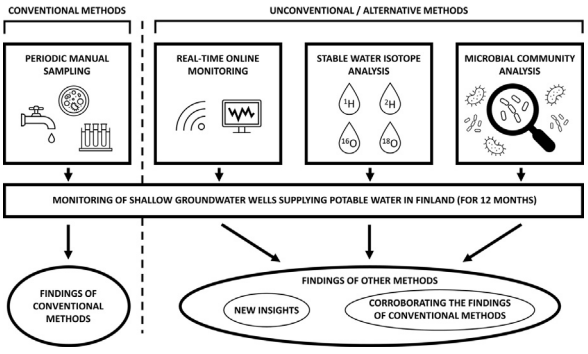
Kevin J. Lyons [a,*,1], Jenni Ikonen [b,1], Anna-Maria Hokajärvi [b], Teemu Räsänen [c,d], Tarja Pitkänen [b,e], Ari Kauppinen [b,f], Katharina Kujala [a], Pekka M. Rossi [a], Ilkka T. Miettinen [b]

[a] Water, Energy and Environmental Engineering Research Unit, University of Oulu, Oulu, Finland
[b] Expert Microbiology Unit, Finnish Institute for Health and Welfare, Kuopio, Finland
[c] Preventos Informatics Oy, Kuopio, Finland
[d] Department of Environmental Technology, Savonia University of Applied Sciences, Kuopio, Finland
[e] Department of Food Hygiene and Environmental Health, Faculty of Veterinary Medicine, University of Helsinki, Helsinki, Finland
[f] Animal Health Diagnostic Unit, Laboratory and Research Division, Finnish Food Authority, Helsinki, Finland

HIGHLIGHTS

- Groundwater is a vital and often threatened source of potable water worldwide.
- Current monitoring methods often overlook groundwater quality deficiencies.
- Here, three alternative methods were applied to monitoring groundwater wells.
- Surface water intrusion and faecal contamination were identified in the wells.
- Alternative methods can complement pre-existing methods and add new insights.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Christian Herrera

Keywords:
Water supply
Water quality
Potable water
Drinking water
Health risks
Microbiology

ABSTRACT

Groundwater provides much of the world's potable water. Nevertheless, groundwater quality monitoring programmes often rely on a sporadic, slow, and narrowly focused combination of periodic manual sampling and laboratory analyses, such that some water quality deficiencies go undetected, or are detected too late to prevent adverse consequences. In an effort to address this shortcoming, we conducted enhanced monitoring of untreated groundwater quality over 12 months (February 2019–February 2020) in four shallow wells supplying potable water in Finland. We supplemented periodic manual sampling and laboratory analyses with (i) real-time online monitoring of physicochemical and hydrological parameters, (ii) analysis of stable water isotopes from groundwater and nearby surface waters, and (iii) microbial community analysis of groundwater via amplicon sequencing of the 16S rRNA gene and 16S rRNA. We also developed an early warning system (EWS) for detecting water quality anomalies by automating real-time online monitoring data collection, transfer, and analysis – using electrical conductivity (EC) and turbidity as indirect water quality indicators. Real-time online monitoring measurements were largely in fair agreement with periodic manual measurements, demonstrating their usefulness for monitoring water quality; and the findings of conventional monitoring, stable water isotopes, and microbial community analysis revealed indications of surface water intrusion and faecal

* Corresponding author at: WE3 Research Unit, PO Box 4300, University of Oulu, 90014, Finland.
E-mail address: kevin.lyons@oulu.fi (K.J. Lyons).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.scitotenv.2022.161199
Received 18 October 2022; Received in revised form 20 December 2022; Accepted 22 December 2022
Available online 26 December 2022

contamination at some of the studied sites. With further advances in technology and affordability expected into the future, the supplementary methods used here could be more widely implemented to enhance groundwater quality monitoring – by contributing new insights and/or corroborating the findings of conventional analyses.

1. Introduction

Groundwater provides about 25–50 % of the world's potable water (Sampat, 2000; Zektser and Everett, 2004), and hence the supply and potability of groundwater are crucial for maintaining good public health and the proper functioning of society (Figueras and Borrego, 2010; Macler and Merkle, 2000). In many regions of the world, however, human activities such as agriculture, industry, and urbanisation are directly and indirectly threatening this valuable natural resource through depletion (Famiglietti, 2014; Konikow and Kendy, 2005; Wada et al., 2010) and water quality degradation (Burri et al., 2019). In addition, groundwater quality monitoring programmes – and potable water monitoring programmes in general – often rely on a sporadic, slow, and narrowly focused combination of periodic manual sampling and laboratory analyses, such that some water quality deficiencies go undetected, or are detected too late to prevent adverse consequences (Banna et al., 2014; Calderwood et al., 2020; Capodaglio and Callegari, 2009; Storey et al., 2011; Velasquez-Orta et al., 2017). Given these threats, there is a need to continuously improve groundwater monitoring strategies, and explore alternative analytical methods, to ensure safe supplies of potable water.

Continuous methods are increasingly being used to monitor groundwater bodies and other water sources (Banna et al., 2014; Capodaglio and Callegari, 2009; Lee et al., 2007; Lee and Kwon, 2016; Storey et al., 2011). These methods allow water supply managers to rapidly and remotely view and evaluate measurements of water quality parameters taken at much more frequent time intervals than possible by conventional periodic manual sampling and laboratory analyses – potentially enabling more reliable detection of water quality deficiencies and a faster response time (Favere et al., 2020; Ikonen et al., 2017; Storey et al., 2011). Some continuous methods produce new measurements with such rapidity (e.g. every few minutes) that they are also known as 'real-time methods' (Calderwood et al., 2020; Drage and Kennedy, 2020; Højris et al., 2018; Oppus et al., 2020). In October 2015, Annex II of the European Union's Drinking Water Directive (DWD) (98/83/EC) was amended (2015/1787) to allow for the use of continuous methods in drinking water monitoring programmes, as an alternative to the manual collection and analysis of discrete water samples. The current December 2020 recast of the DWD (2020/2184) retains this amendment, modified to also allow for the possibility of using both methods simultaneously.

However, given that not all parameters can be monitored by continuous methods in current use, there is also a need to explore and apply other analytical methods. One such method is stable water isotope analysis, which can be used to identify signs of surface water intrusion in groundwater wells, by comparing $^{2}H/^{1}H$ and $^{18}O/^{16}O$ isotope ratios in samples from groundwater and nearby surface water bodies (Hunt et al., 2005; Parlov et al., 2019). This method relies on the fact that evaporation of lighter isotopes (^{1}H and ^{16}O) typically occurs more from surface water than groundwater, leaving surface water more enriched in heavy isotopes (^{2}H and ^{18}O) (Gat, 2010, Gat, 1996). Another method is 16S rRNA amplicon sequencing (a.k.a. 16S metabarcoding), which can be used to investigate the composition and spatiotemporal variation of groundwater microbial communities (Chik et al., 2020; Clark et al., 2018; Kim et al., 2015). Like continuous methods, both of these approaches can provide insights which conventional periodic manual sampling and laboratory analyses cannot.

Therefore, the aim of this study was to examine the potential benefits of enhancing the conventional monitoring of four shallow groundwater wells with (i) real-time online monitoring, (ii) stable water isotope analysis, and (iii) 16S rRNA amplicon sequencing. The idea was that by combining conventional and supplementary methods, a more complete and reliable understanding of the factors influencing each well would emerge.

2. Materials and methods

2.1. Study sites and sampling programme

In this study, we monitored four groundwater wells supplying potable water in northern and central Finland. Wells 1, 2 and 3 are in the North Ostrobothnia region of northern Finland – wells 1 and 2 in the same inland municipality, and well 3 in a different municipality ~3 km from the sea – and well 4 is in the North Savo region of central Finland. Relevant characteristics of the four wells are shown in Table 1 and schematic maps of the study sites are shown in Fig. S1. These wells were selected for monitoring based on the findings of a previous study (Lyons et al., 2021), which raised concerns about potential risks to water quality at these sites, either from suspected surface water intrusion or impacts from nearby land use.

Four different approaches were used to monitor the groundwater wells in this study: (i) manual on-site measurements and manual sampling for laboratory analyses (to periodically assess physicochemical and microbiological parameters of untreated groundwater), (ii) real-time online monitoring (to continuously assess physicochemical and hydrological parameters of untreated groundwater), (iii) periodic analysis of stable water isotopes from samples of untreated groundwater and samples from nearby surface water sources (to investigate groundwater–surface water interactions), and (iv) springtime and autumntime amplicon sequencing of the 16S rRNA gene and 16S rRNA from untreated groundwater samples (to assess the composition of groundwater microbial communities). Each of these approaches is elaborated in greater detail in the following sections. (For a summary of all sampling and monitoring conducted during this study, see Table S1.)

2.2. Periodic manual on-site measurements and manual sampling for laboratory analyses

To assess physicochemical and microbiological groundwater quality, on-site groundwater measurements and untreated groundwater samples for laboratory analyses were taken manually at the four groundwater

Table 1
General characteristics of the four groundwater wells monitored in this study.

Well	Type	Aquifer	Depth (m)	GW depth (m)	Users	Intake (m^3/day)	Treatment	Last changes to well structure	Nearby risk factors (within 1 km^2)
1	Dug	Sand/gravel	6	2	<100	13.7	none	1980s	M, SG, SW
2	Dug	Sand/gravel	7.5	3	4200[a]	650	UV, ALK	1961	SG, SA, S, R, SW
3	Tube	Sand/gravel	≥8	1.5	7000[a]	400	UV, ALK, CH	1993	A, SW, R, RA, S, C
4	Dug	Sand/gravel	9	5–10	20,000[a]	600–1000	UV, ALK	1969	R, SW, T

ALK = alkalisation, UV = ultraviolet disinfection, CH = chemical purification, GW = groundwater, A = agriculture, C = cemetery, M = marsh, R = roads, RA = recreational area, S = school, SA = swimming area, SG = sand or gravel pit, SW = surface water, T = town.

[a] Water served from several wells to the same network.

wells during 8 monthly sampling timepoints (March–November 2019; excluding July).

During manual groundwater sampling, untreated groundwater samples were collected at each of the study sites from a sampling tap specially designed for this purpose. In each case, the sampling tap was first flame-sterilised (by spraying 70 % ethanol onto the tap from a spray bottle and lighting the ethanol with a handheld lighter). The tap was then opened, and untreated groundwater was collected in (i) a clean plastic bucket, for on-site physicochemical measurements using handheld field meters, (ii) a 500 mL brown glass bottle, for physicochemical analyses performed in the laboratory, and (iii) a 1 L polypropylene (PP) bottle, for cultivation-based analyses of microbiological indicators. Between sampling rounds, the brown glass bottles were washed in acid (2 % HCl), rinsed with distilled water, dried, and incinerated at 550 °C to remove carbon traces; and the PP bottles were washed and steam-sterilised in an autoclave at 120 °C.

Groundwater temperature, pH, electrical conductivity (EC), redox potential, salinity, and dissolved oxygen (DO) were measured on site with the WTW Multi 350i (at wells 1, 2 and 3) and WTW Multi 3430 (at well 4) handheld meters and associated sensors (WTW, Weilheim, Germany). All redox potential values reported in this work have been converted to standard hydrogen electrode (ORP) values. Temperature, pH, EC, turbidity, UV absorbance at 254 nm, Fe, and Mn were measured from the groundwater samples in the laboratory. Fe and Mn were measured because high levels of Fe and Mn are a common groundwater quality concern in Finnish wells (Isomäki et al., 2006; Pitkänen et al., 2015). Cultivation methods were used to assess counts of coliform bacteria (SFS-EN ISO 9308-1; SFS 3016) (including *Escherichia coli*, a faecal indicator), spores of sulphite-reducing clostridia (SSRC) (ISO 6461-2) (an indicator of microbial persistence forms), and heterotrophic bacteria (a general indicator of changes in bacterial water quality) (Allen et al., 2004). Counts of heterotrophic bacteria were determined via spread-plating technique on R2A agar at 22 °C for 7 days (Greenberg et al., 1992; Reasoner and Geldreich, 1985). The Colilert-18® method (for the detection of coliforms and *E. coli*) was applied monthly to samples from well 4 only (Fricker et al., 1997).

In addition to the main sampling described above, two additional sampling timepoints for all wells were arranged to coincide with the snowmelt period in springtime (late April–early May 2019) and a period of rain in autumntime (late October 2019), as these were considered risk periods for surface water intrusion. During these additional sampling timepoints manual on-site measurements were taken as before. Moreover, dead-end ultrafiltration (DEUF) capsules (ASAHI Rexeed-25A, Asahi Kasei Medical Co., Ltd., Tokyo, Japan) were used to filter ~200 L of untreated groundwater at each site as described previously (Lyons et al., 2021). A third DEUF sampling timepoint was arranged for well 1 only (in early October 2019), after a period of heavy autumntime rain, but no manual measurements were performed at that time. All manually collected groundwater samples and DEUF capsules were transported from the study sites to the laboratory in cool boxes and either processed or frozen within 24 h.

Material caught in the DEUF capsules was eluted in the laboratory as described previously (Lyons et al., 2021); including secondary concentration by filtration through 0.22-μm Millipore Express PLUS membrane filters (Merck KGaA, Darmstadt, Germany). The 0.22-μm membrane filters themselves were used for nucleic acid extractions (see Section 2.6), and the flow-through from the 0.22-μm filtration was concentrated again via polyethylene glycol (PEG)-precipitation (Kauppinen et al., 2019), and used to assess counts of somatic coliphages (USEPA Method 1601; USEPA Method 1602) and F-specific coliphages (USEPA Method 1601). Counts of intestinal enterococci were analysed from raw DEUF eluates using 0.45-μm GN-6 Metricel® MCE filters (Pall Corporation, New York, USA) (SFS-EN ISO 7899-2).

2.3. Real-time online groundwater monitoring

A real-time online monitoring system was installed in each of the four groundwater wells, to continuously assess physicochemical and hydrological parameters of the four groundwater wells. At wells 1 and 4, this system was comprised of a YSI 6920 V2-2 multiparameter water quality sonde (YSI Inc., Yellow Springs, OH, USA) and a DL-12 data logger (EHP Environment Oy, Oulu, Finland); and, at wells 2 and 3, a YSI 600 OMS V2-1 multiparameter optical monitoring sonde (YSI Inc., Yellow Springs, OH, USA) and a DL-12 data logger (EHP Environment Oy, Oulu, Finland). YSI sondes were mounted on flow-through columns and installed at the wells on untreated groundwater sampling taps, from which untreated groundwater came directly from the well pump at a flow rate of about 1–3 L/min (Fig. S2A). Groundwater temperature (°C), EC (μS/cm), turbidity (NTU), and groundwater level were monitored at all wells. pH, redox potential (mV), and DO (mg/L) were monitored at wells 1 and 4; groundwater volume was monitored at sites 1 and 3; and groundwater flow was monitored at wells 3 and 4. Measurements of all physicochemical parameters were taken by the YSI sonde, groundwater level was monitored by a pressure sensor, and groundwater flow and volume readings were obtained from pre-existing well meters. In all cases, measurements were taken every 5 min during the monitoring period.

Real-time monitoring proceeded for 12 months (February 2019–February 2020). However, not all parameters were measured for this entire duration (Table S1). At each site, measurements from the YSI sonde were transferred to a DL-12 data logger (EHP Environment Oy, Oulu, Finland) via the SDI-12 protocol (Fig. S2B). From the four data loggers, the measurement data were transferred to a server via the global system for mobile communications (GSM) network, and then to a cloud-based data repository (EHP Environment Oy, Oulu, Finland) using general packet radio service (GPRS) data transfer technology. The measurement data were stored in a receiving database from which they could be retrieved for processing and analytics. The data were available through a web-based interface and through a representational state transfer application programming interface (REST API), which enables data to be shared with third-party systems. In this study, measurement data were directed automatically, using the open REST API, to a cloud-based data processing and outlier detection service (Preventos Informatics Oy, Kuopio, Finland).

In addition to measurements obtained directly via the real-time online monitoring systems, data approximating precipitation, snow depth, and air temperature at the study sites were obtained by downloading publicly available timeseries data on these parameters from the nearest observation stations of the Finnish Meteorological Institute (FMI) (https://en.ilmatieteenlaitos.fi/download-observations). As wells 1 and 2 are in the same inland municipality, a single FMI observation station proved to be the nearest to both wells.

2.4. Development of an EWS for detecting deviations in groundwater quality

An early warning system (EWS) for the detection of groundwater quality deviations was developed by automating the real-time collection, transfer, and analysis of EC and turbidity measurements. These parameters were chosen for three reasons: (i) they are known to be effective indirect water quality indicators (Isomäki et al., 2008; Turunen et al., 2020; WHO, 2017), (ii) they are rather easy and cost-effective to monitor on a continuous or real-time basis, and (iii) the sensors used for measuring these parameters do not typically require much maintenance – although, optical methods of measuring turbidity can sometimes be impeded by sensor fouling or air bubbles (Anderson, 2005). In addition to detecting groundwater quality deviations, the system was designed to report results to users in real time. The steps of the system can be summarised as follows: (i) pre-processing of measurement data, (ii) creating a baseline model based on historical data, (iii) comparing the latest measured value with the model, (iv) sending an alarm upon detecting several consecutive anomalous values (but no alarm for single anomalous values).

2.5. Stable water isotope analyses

To enable analyses of stable water isotopes in this study, untreated groundwater samples were taken monthly from all wells, and surface water samples were taken during summertime and autumntime from

surface water sources near wells 1 and 2 (within 300 m of the well in each case). Dual isotope ratios ($^{2}H/^{1}H$ and $^{18}O/^{16}O$) were determined from water samples using cavity ring-down spectroscopy with a Picarro L2130-i analyser (Picarro, Inc., Santa Clara, CA, USA). All isotope ratios are expressed in δ notation relative to Vienna Standard Mean Ocean Water 2 (VSMOW2) with precision for $\delta^{2}H$ and $\delta^{18}O$ values of δ0.1 ‰ and δ0.025 ‰, respectively. Regional isotope data were collected from previous studies conducted in Finland for comparison and analysis: namely, a local meteoric water line (LMWL) from Oulanka National Park indicating variance in the isotopic signatures of local meteoric water (i.e. snow and rain) (Rossi et al., 2015), and local evaporation lines (LEL) from Rokua (Isokangas et al., 2015) and Posio (Nora et al., 2019), indicating variance in the isotopic signatures of local surface water sources.

2.6. Nucleic acid extractions, microbial source tracking markers, and gram-negative gene copy number counts

Total nucleic acids were extracted from the 0.22-μm membrane filters through which DEUF eluates had been filtered, using the Chemagic DNA Plant Kit (Perkin Elmer, Waltham, MA, USA) as described previously (Brester et al., 2020). Purified RNA was obtained by processing the total nucleic acids with the Ambion Turbo DNA-free DNase kit (Life Technologies, Carlsbad, CA, USA). cDNA was synthesized using Invitrogen Superscript IV VILO system (Thermo Fisher Scientific, Waltham, MA, USA). Purified RNA was stored at −75 °C or colder until use, while total nucleic acids (serving as a DNA template) and cDNA were stored at −20 °C. Gene copy number counts of microbial source tracking (MST) markers GenBac3 (targeting Bacteroidales bacteria as general indicators of faecal contamination) (Dick and Field, 2004; Siefring et al., 2008), and HF183 (targeting human-associated *Bacteroides* bacteria as indicators of human-derived faecal contamination) (Bernhard and Field, 2000a, 2000b; Converse et al., 2009; Haugland et al., 2010), and Gram-negative bacteria (Kärkkäinen et al., 2010) were measured from DNA and cDNA with TaqMan chemistry as described previously (Pitkänen et al., 2013). Primers and probes used in this study are shown in Table S2. RT-qPCR performance features and detection and quantification limits are shown in Table S3.

2.7. Amplicon sequencing of the V3–V4 region of the 16S rRNA and rRNA gene

Extracted nucleic acids were sent to Macrogen, Inc. (Seoul, South Korea) for amplicon generation and subsequent paired-end sequencing. The primers Bakt_341F (CCTACGGGNGGCWGCAG) and Bakt_805R (GACTACHVGGGTATCTAATCC), which target the V3–V4 variable region of the 16S rRNA gene (Herlemann et al., 2011), were used to generate amplicons from DNA (theoretically targeting all microbes with 16S rRNA genes) and cDNA (theoretically targeting only metabolically active microbes with 16S rRNA genes). Amplicons were sequenced as 2 × 300 bp paired-end reads using the Illumina MiSeq platform. Negative controls were included in the high-throughput amplicon sequencing analysis (an 'extraction control' to test for potential contamination from the DNA extraction kit, and several 'elution controls' to test for potential contamination arising from the DEUF processing steps). Nucleic acid templates for springtime DEUF samples from well 4 were lost in transit to the sequencing company and therefore not sequenced.

2.8. Sequencing data processing and analysis

The 16S rRNA amplicon data for DNA and cDNA libraries were processed and analysed via the QIIME 2 pipeline (version 2021.2) (Bolyen et al., 2019). The 'dada2 denoise-paired' QIIME 2 command was used, with the parameters –p-trim-left-f 17, –p-trim-left-r 21, –p-trunc-len-f 294, and –p-trunc-len-r 216, to trim sequences (to remove primer remnants and bad quality reads with quality scores of <20) and to denoise and merge trimmed reads to produce a table of amplicon sequence variants (ASVs) (Callahan et al., 2016). Taxonomic classification of ASVs was performed by using the 'q2-feature-classifier' plugin of QIIME 2 to train a naïve Bayes classifier on the V3–V4 variable region of pre-formatted representative 16S rRNA sequences derived from the SILVA rRNA database (release 138) using RESCRIPt (Bokulich et al., 2018; Quast et al., 2012; Robeson et al., 2020). Alpha and beta diversity metrics and principal coordinates analysis (PCoA) of the Bray-Curtis dissimilarity matrix were calculated and derived via the 'diversity core-metrics-phylogenetic' command, using a sampling depth of 18,332 (Halko et al., 2011; Sørensen, 1948). Non-metric multidimensional scaling (nMDS) of the QIIME2 ASV table was performed via the 'metaMDS' function of the 'vegan' R package using the Bray-Curtis dissimilarity metric (Oksanen, 2020). Metadata variables associated with water quality (e.g. turbidity) and site-specific environmental features (e.g. roads, fields, marshes, surface waters) were fitted to the nMDS plots using the 'envfit' function. Only variables that were significantly correlated with microbial community composition (i.e. those identified in 'envfit' with a P value of <0.05) were included in the figures.

2.9. Sequencing data availability

The 16S amplicon sequencing data for this study – with sequencing primer remnants removed using 'cutadapt' (Martin, 2011) – have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under primary accession number PRJEB52434 (https://www.ebi.ac.uk/ena/browser/view/PRJEB52434).

3. Results

3.1. Periodic manual sampling and measurements

3.1.1. Physicochemical and hydrological parameters

Finnish groundwater used for potable water supply is often cool (2.3–8.9 °C), slightly acidic (pH 6.3–6.5), and oxic (1–12 mg/L) (Isomäki et al., 2008). By these measures, the wells studied here were rather typical. However, some extreme physicochemical values and unexpected patterns of variation were observed throughout the monitoring period. Measured physicochemical and hydrological parameters of the four groundwater wells are summarised in Table 2. (A more detailed comparative summary of on-site, online, and lab measurements can be found in Table S4.) In manual measurements, well 1 had the lowest median EC values and the highest median redox potential values; well 3 had the highest median temperature, turbidity, UV_{254} absorbance, Fe and Mn values, and the lowest DO values; and well 4 had the highest median EC and pH values.

3.1.2. Microbiological water quality indicators

Detected microbiological water quality indicators are summarised in Fig. 1. (For raw data and summary statistics, see Table S5.) Heterotrophic bacteria, coliform bacteria (including *E. coli*), somatic coliphages, and spores of sulphite-reducing clostridia (SSRC) were measured during 8 monthly sampling timepoints (March–November 2019; excluding July). The highest counts of heterotrophic bacteria and coliform bacteria observed in each well ranged from 600 to 3100 CFU/mL and from 3 to 1300 CFU/L, respectively (Fig. 1A and B). All of these values were high in the sense of being greater than at least one standard deviation above their respective means and medians (Table S5). *E. coli* bacteria (indicative of faecal contamination) were not detected in any groundwater samples, the Colilert-18® method applied to samples from well 4 also returned non-detects at every timepoint, and somatic coliphages were detected only once in the study (1 PFU/1.1 L, well 3, May). Low levels of SSRC (1 CFU/L) were detected in well 2 in May, and higher levels in well 3 at almost all timepoints, the highest being in August and September (both 20 CFU/L) (Fig. 1C).

Gene copy numbers of Gram-negative bacteria and the GenBac3 (general indicator of faecal contamination) and HF183 (indicator of human-derived faecal contamination) microbial source tracking (MST) markers were measured from DNA and cDNA in springtime (late April–early May 2019) and autumntime (late October 2019). The highest gene copy numbers of Gram-negative bacteria in DNA were observed in well 3 during

Table 2
Summary of physicochemical and hydrological parameters at the four groundwater wells.

	T (°C)	pH	EC (µS/cm)	DO (mg/L)	ORP (mV)	turb (NTU)	UV_{254} abs	Fe (mg/L)	Mn (mg/L)	level* (m)	flow (L/s)
	On-site	On-site	On-site	On-site	On-site	Lab	Lab	Lab	Lab	Online	Online
Well 1											
Median	6.25	6.17	27	9.5	357	0.03	0.047	<0.02	<0.006	1.789	NA
Mean	6.02	6.05	108	8.96	374	0.04	0.066	<0.02	<0.006	1.877	NA
SD	0.59	0.32	6.54	2.71	130.23	0.03	0.05	0.005	0.002	0.288	NA
Max	6.5	6.24	40	12.65	530	0.19	0.181	0.03	0.01	2.734	NA
Min	5.1	5.48	25	5.8	212	0.01	0.037	<0.02	<0.006	1.561	NA
Well 2											
Median	5.8	6.57	90	13.4	266	0.08	0.021	<0.02	<0.006	1.238	NA
Mean	5.8	6.2	104	12.9	300	0.12	0.024	<0.02	<0.006	1.314	NA
SD	0.24	0.17	68.4	2.61	79.5	0.16	0.01	0.014	0.001	0.337	NA
Max	6.1	6.9	235	16.5	420	0.51	0.055	0.05	0.007	2.153	NA
Min	5.4	6.46	32	10	235	0.01	0.017	<0.02	<0.006	0.869	NA
Well 3											
Median	8.7	6.2	93	3	95.75	11.9	1.650	2.76	0.15	4.920	4.7
Mean	8.6	6.2	92	4.3	95.75	14.5	1.632	2.85	0.154	4.870	4.7
SD	1.20	0.12	7.75	3.59	19.45	7.60	0.14	0.38	0.01	0.166	0.254
Max	10.2	6.36	101	8.6	109.5	26.26	1.82	3.7	0.18	5.234	5.8
Min	6.9	6.1	82	0.95	82	5.64	1.416	2.45	0.141	4.550	3.2
Well 4											
Median	6.65	7.10	381	14.12	NA	0.05	0.06	<0.02	0.013	4.275	6.2
Mean	6.51	7.06	398	13.89	NA	0.12	0.06	<0.02	0.012	4.316	6.0
SD	0.38	0.30	36.6	0.54	NA	0.14	0.01	0	0.003	0.135	1.81
Max	6.9	7.41	467	14.51	NA	0.32	0.084	0.02	0.017	4.579	14.6
Min	5.9	6.61	364	12.86	NA	0.008	0.054	<0.02	<0.006	3.433	0.8

T = temperature, EC = electrical conductivity, DO = dissolved oxygen, ORP = redox potential, turb = turbidity, UV_{254} abs = UV_{254} absorbance, Fe = iron, Mn = manganese; * = groundwater level units are not standardised (e.g. 'metres above sea level'), rather 'metres above some arbitrary site-specific level'; SD = standard deviation; NA = this parameter was not measured at all for the well in question.

springtime (3.2×10^5 GC/100 mL) and autumntime (2.7×10^5 GC/100 mL), and the highest level in cDNA was observed in well 2 during springtime (9.9×10^6 GC/100 mL) (Fig. 1D and E). The GenBac3 marker was detected in cDNA at about 150 GC/100 mL in well 2 during springtime, at a level below the limit of quantification in well 3 during springtime, and at <10 GC/100 mL in well 4 during autumntime (Fig. 1F). Assays aiming to detect the GenBac3 marker in DNA, and the HF183 marker in DNA and cDNA yielded non-detects for all wells at both timepoints. Similarly, tests for intestinal enterococci and F-specific coliphages in springtime and autumntime samples also yielded non-detects for all wells at both timepoints.

3.2. Real-time online monitoring

Real-time online monitoring measurements were largely in fair agreement with periodic manual measurements. In real-time measurements, well 1 had the lowest EC values and the highest redox potential values; well 3 had the highest groundwater temperature and turbidity values; and well 4 had the highest EC values. These observations matched exactly with those of the manual methods. However, various short-term and long-term relationships between monitored parameters were also apparent in the real-time data, either in plots of the original measurements taken at 5-min intervals, or in plots of daily mean values. Some of these appear to reflect true variation in measured parameters, whereas others are the result of artificial variation caused by maintenance procedures. (For a big-picture summary of all raw real-time online monitoring data and Finnish Meteorological Institute (FMI) data, see Fig. S3.)

3.2.1. Springtime snowmelt is the main source of groundwater recharge in the studied wells

At all wells, an annual peak in groundwater level was observed shortly after the springtime decline in snow depth, indicating snowmelt to be the primary source of annual recharge in these wells (Fig. 2A). In well 1, there were also several noticeable increases in groundwater level after heavy rains during the summer, with dissolved oxygen (DO) levels appearing to peak between groundwater level peaks (Fig. S4). In most wells, groundwater level appeared to rise again in autumn, presumably indicating groundwater recharge from rain events prior to the beginning of another period of snow accumulation.

3.2.2. Operational and maintenance-related variation

The groundwater level in well 2 is shown in Fig. 2B. This well is situated near a river, and there is a known risk of surface water intrusion there during the springtime snowmelt period every year, when the level of the surface water rises above the level of the groundwater. Hence, the well is taken out of operation during springtime each year to avoid surface water intruding into the well. In 2019, normal pumping of groundwater was halted between late April and early June. During this time, however, 10 % of the normal amount of groundwater was pumped from the well via a bypass system so that the interaction between surface water and groundwater could be examined as part of this study. At all other times during the monitoring period, the pumping at well 2 was on a daily on/off cycle, which was reflected in daily variation in groundwater level measurements. In fact, during the monitoring period, cyclical daily patterns were observed at times in the real-time online monitoring data of all wells due to daily well pumping schemes (Fig. S5). These included parameters such as groundwater level and temperature, EC, DO, turbidity, pH, ORP, flow and volume, although not all of these parameters exhibited cyclical patterns in all wells at all times. Such cyclicality is notable, as it may influence the nature of any models devised to detect deviations from 'normal state' (i.e. baseline) behaviour.

The electrical conductivity (EC) and turbidity in well 3 are shown in Fig. 2C. Although the real-time online monitoring largely proceeded in a satisfactory manner in all wells, the high iron concentration of the groundwater in well 3 (median of 2.8 mg/L according to the periodic manual measurements) posed a challenge for the recording of accurate turbidity measurements at this well. All turbidity sensors used in this study were fitted with an automatic wiper. However, at well 3, an iron precipitate

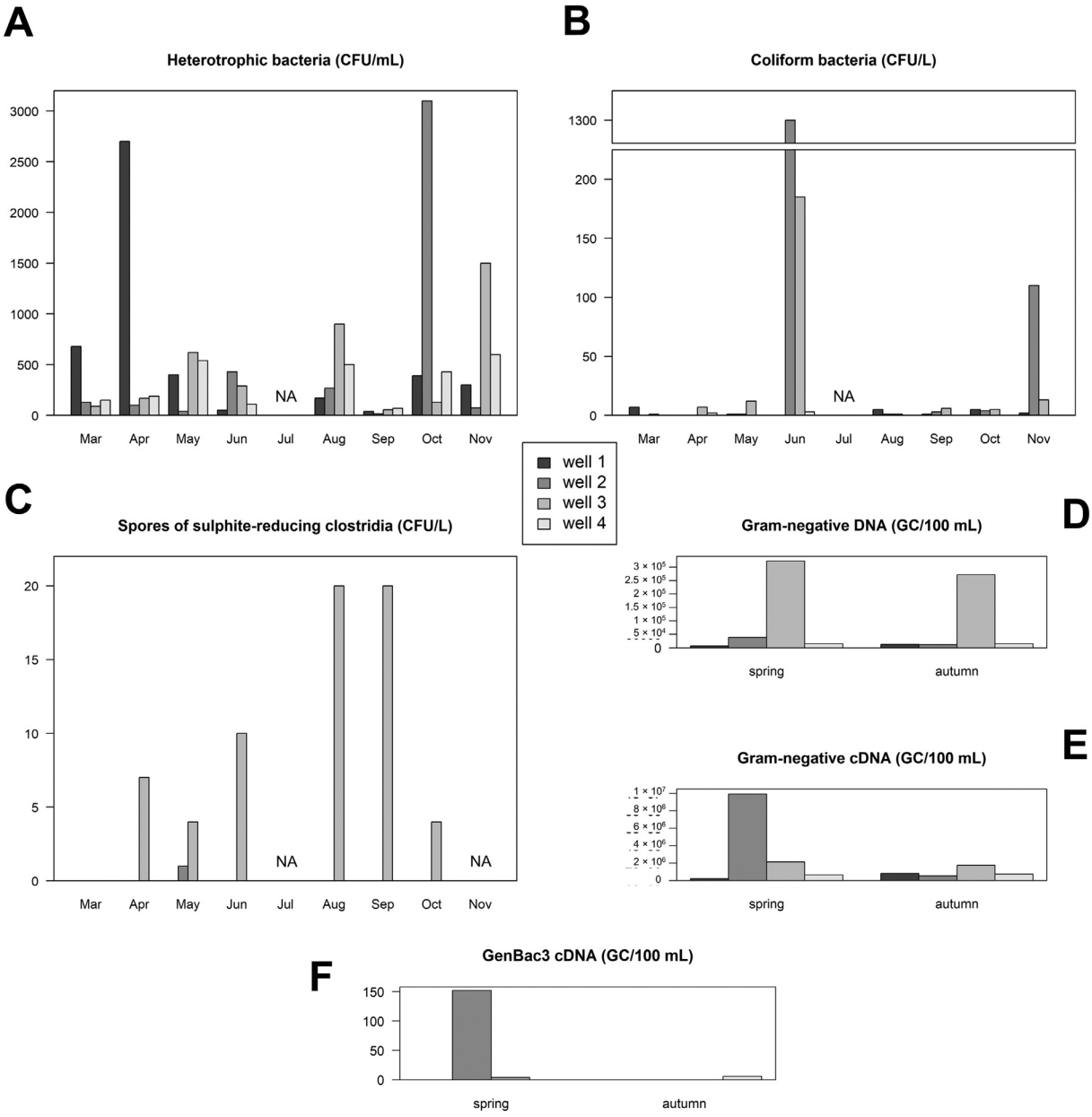


Fig. 1. Microbiological water quality indicators in untreated groundwater samples. A: Counts of heterotrophic bacteria (CFU/mL), determined by culturing on Reasoner's 2A (R2A) agar at 22 °C for 7 days. No samples were taken in July. B: Counts of coliform bacteria (CFU/L). Both SFS 3016 and ISO 9308-1 methods were conducted at each major sampling timepoint for samples from each well (Mar to Nov, excluding July). In some cases, no coliforms were detected by either method. In cases where one method detected coliforms and the other did not, the values shown above are from a single method. In cases where both methods detected coliforms, the values shown above are averages of the results obtained by both methods. Two sets of samples were taken during the snowmelt period in April (results are averaged here). C: Counts of spores of sulphite-reducing clostridia (SSRC) (CFU/L), determined by the ISO 26461–2 method. D and E: Gene copy numbers of Gram-negative bacteria in DNA and cDNA (GC/100 mL). F: Gene copy numbers of the GenBac3 MST marker (general indicator of faecal contamination) in cDNA from all four wells. The GenBac3 marker was not detected in DNA. spring = late April–early May; autumn = October.

accumulated around the sensor during the monitoring period, leading to the occurrence, at times, of extremely high turbidity readings which did not represent the true turbidity of the groundwater. Due to this repeated fouling of the sensor, repeated manual cleaning of the sensor was required. These cleaning events were apparent in the measurement values. In addition to disruptions from sensor cleaning, turbidity values were also periodically disrupted by another maintenance procedure, namely the flushing of a pipe to remove iron precipitate (Fig. 2C). This flushing procedure caused a momentary turbidity peak (probably caused by precipitate that became detached from the pipe surfaces), and also increased the EC (probably bringing water from a different layer of the groundwater body to the well).

3.3. Development of the early warning system (EWS)

3.3.1. Overview and data pre-processing

As part of this study, an early warning system (EWS) was developed that can (i) automatically detect deviations from normal (baseline) groundwater quality, based on the interpretation of real-time online monitoring data, and (ii) when a deviation is detected, send an alarm to the relevant personnel in real time, in the form of an easily understandable text-based report. Although this work provides the theoretical proof of concept for the design, development, and operation of an EWS, this EWS was not extensively tested in practice due to time and budget constraints.

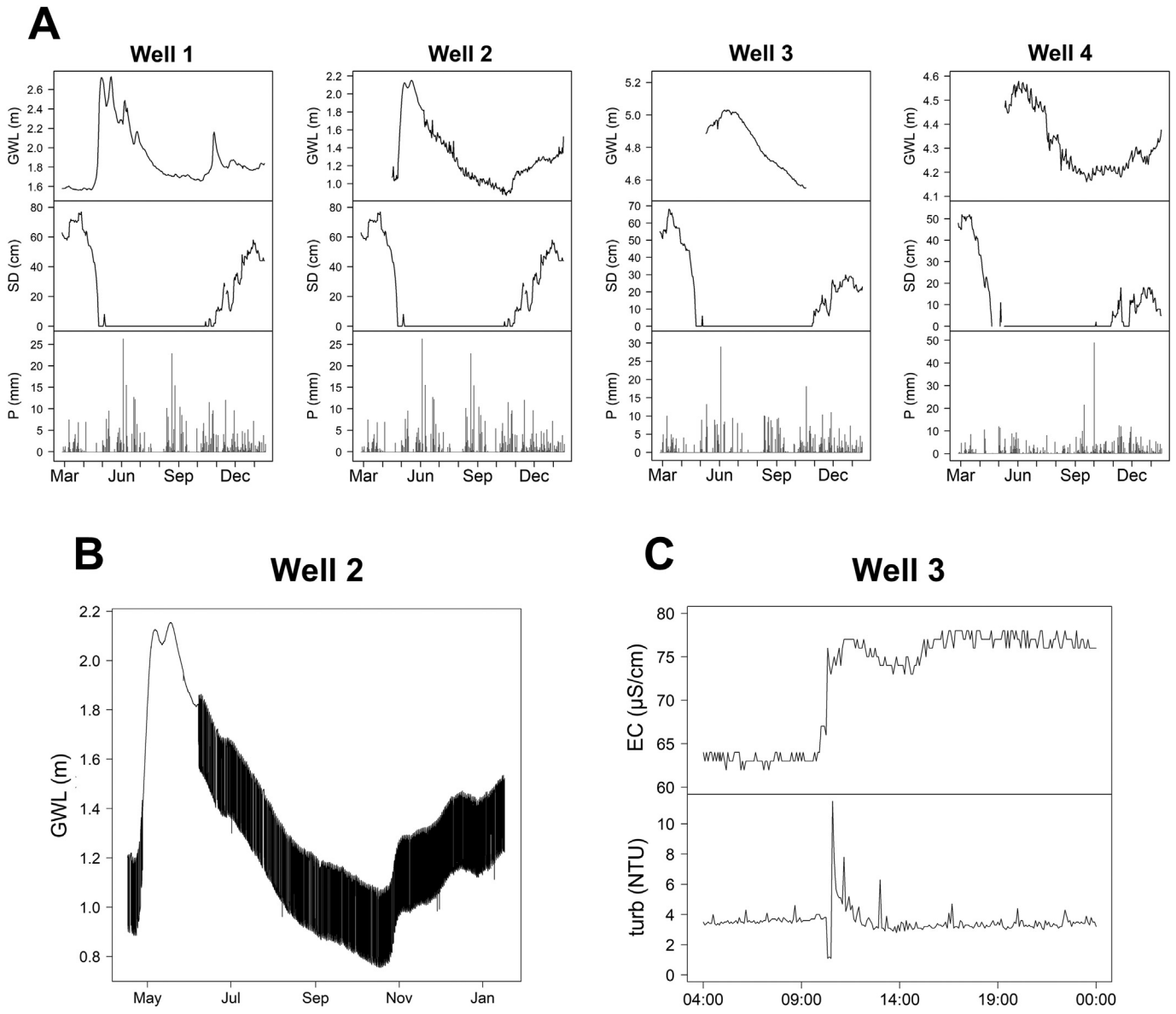


Fig. 2. True and artifical variation in the real-time online monitoring data. A: Daily data for groundwater level, snow depth and precipitation at the four studied wells. Snow depth and precipitation data are derived from the nearest observation station of the Finnish Meteorological Institute (FMI). As wells 1 and 2 are in the same inland municipality, a single FMI observation station proved to be the nearest to both wells, hence the snow depth and precipitation data are the same in both cases. Groundwater level data for well 3 from mid-October onwards have been omitted here, as well maintenance around this time impaired sensor function leading to unreliable values. GWL = groundwater level, SD = snow depth, P = precipitation. B: Groundwater level (GWL) variation in well 2. Normal pumping was halted between late April and early June. At other times, pumping follows a daily on/off cycle, according to consumption. Groundwater level units shown here are not standardised (e.g. 'metres above sea level'), rather 'metres above an arbitrary site-specific level'. C: Electrical conductivity (EC) and turbidity (turb) values for well 3 from 17 October 2019. Pipe flushing increased the EC and caused a momentary increase in the turbidity measurements.

The EWS developed during this study relies on measurements of the electrical conductivity (EC) and turbidity of groundwater taken at 5-minute intervals by the real-time online monitoring systems. The raw measurement data contained peaks and noise which obscured the true variation in EC and turbidity. A moving average filter (Smith, 2003) was used to remove noise from the raw measurement data. An example of the effects of this pre-processing is shown in Fig. S6.

3.3.2. Developing baseline models of groundwater conductivity and turbidity

Groundwater consumption can vary depending on 'day of the week' and 'time of day'. This variation is caused by the changing water use rates of consumers in the area (e.g. households, agriculture, industry, services). In this study, models describing the baseline behaviour of groundwater conductivity and turbidity in each well were created by sorting data from the previous four weeks by 'day of the week', and calculating averages for each timepoint (i.e. each 5-minute interval) of each day. Examples of daily variation in conductivity and turbidity values in well 4 are presented in Fig. 3 along with calculated models. Such models must be recalculated from time to time to account for changes in rates of water use – arising, for example, from a sudden increase in the local population (e.g. an influx of holidaymakers in summertime) or an increase in nearby industrial activity.

3.3.3. Sending an alarm in response to water quality anomalies

The detection of water quality anomalies is based on comparing the most recently measured value with the value given by the baseline model at the same timepoint. If the most recently measured value deviates significantly from the model – and what counts as a 'significant' deviation must

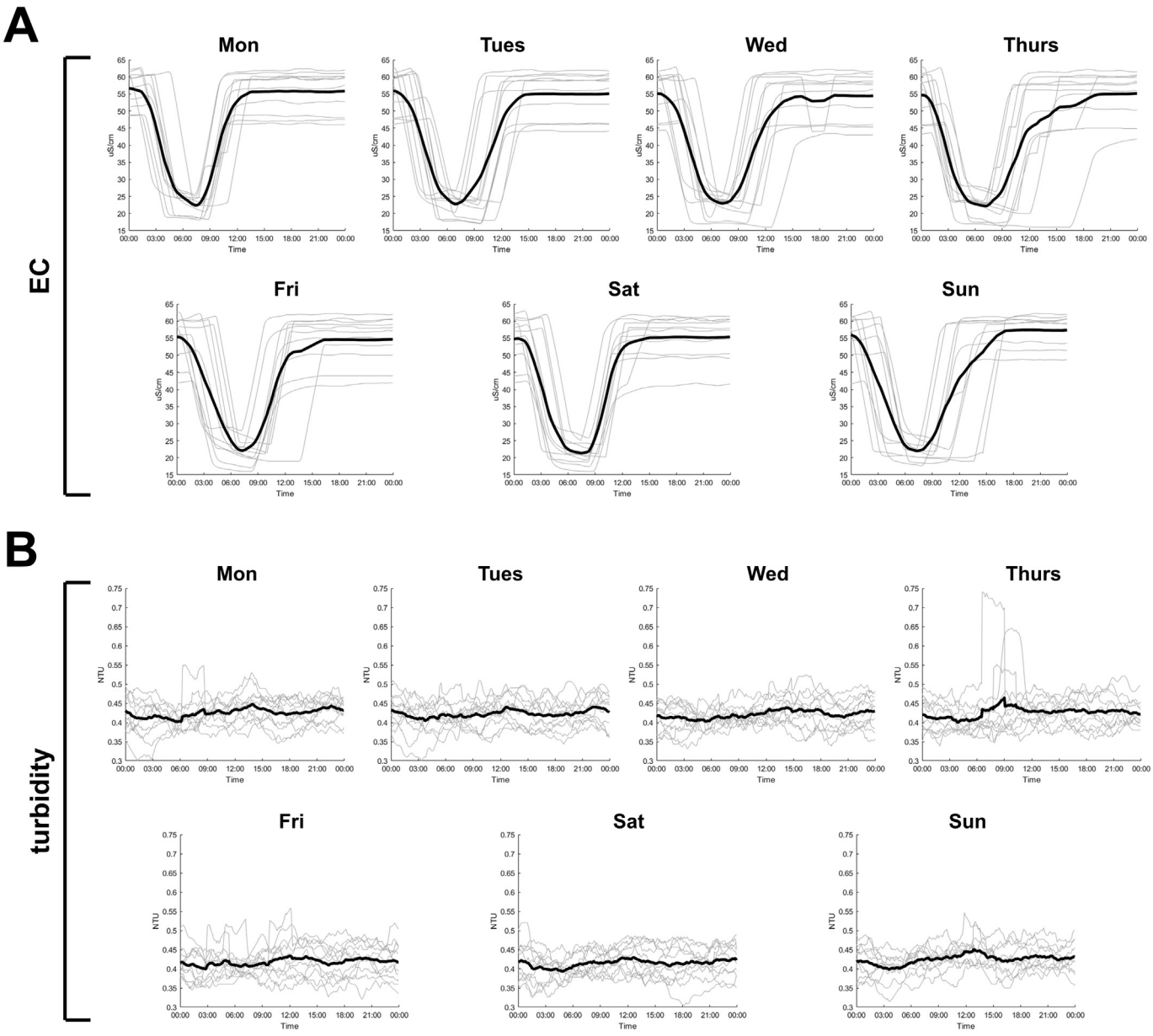


Fig. 3. Daily variation of electrical conductivity (A) and turbidity (B) in well 4 (grey) and calculated model (black).

be determined on a site-specific basis – it can be interpreted as an anomaly. However, individual anomalies can occur in the measurements for various benign reasons (e.g. air bubbles), and so to eliminate false alarms, the system only sends an alarm when several consecutive measured values have been interpreted as anomalies. The settings for detecting anomalies can be modified to suit the context. In this case, the measurement interval was 5 min, and the system only sends an alarm when five consecutive measured values have been interpreted as anomalies – i.e. five consecutive anomalies indicate that an abnormal situation has persisted for 25 min and is perhaps worthy of investigation. In this way, it can be ensured that single measurement errors do not affect the detection of true water quality anomalies.

If a true water quality anomaly is found, an alarm can be transmitted in three different ways: (i) an alarm can be visualized in the system's web browser-based map interface, (ii) an alarm message can be sent automatically as an SMS message to the supervisor's mobile phone, or (iii) by email to a wider group. Upon receiving the alarm message, the supervisor can immediately check the status of the water intake from the mobile application and consider whether precautions should be taken.

3.4. Stable water isotope analyses

Groundwater and surface water bodies are often connected (Winter et al., 1998). Based on field analyses of the study sites, previously collected data, and information obtained from the water utilities, all of the groundwater wells investigated here are suspected to have connections to nearby surface water bodies. In this study, stable water isotope analyses were used to investigate this possibility further (Fig. 4). In northern and central Finland, most groundwater recharge comes from melted snow, so the isotopic signature of a groundwater sample is typically expected to be similar to the isotopic signature of meteoric water (e.g. the local meteoric water line (LMWL) in Fig. 4). If the isotopic signature of a groundwater sample instead appears somewhat similar to the isotopic signature of surface water that has been subjected to evaporation (e.g. the local evaporation lines (LELs) in Fig. 4), this may be a sign that surface water is intruding into the groundwater well.

At well 1, the isotopic signatures of groundwater samples varied along the LMWL, and no clear signs of evaporation (i.e. surface water intrusion) were observed (Fig. 4). The isotopic signatures of samples collected from

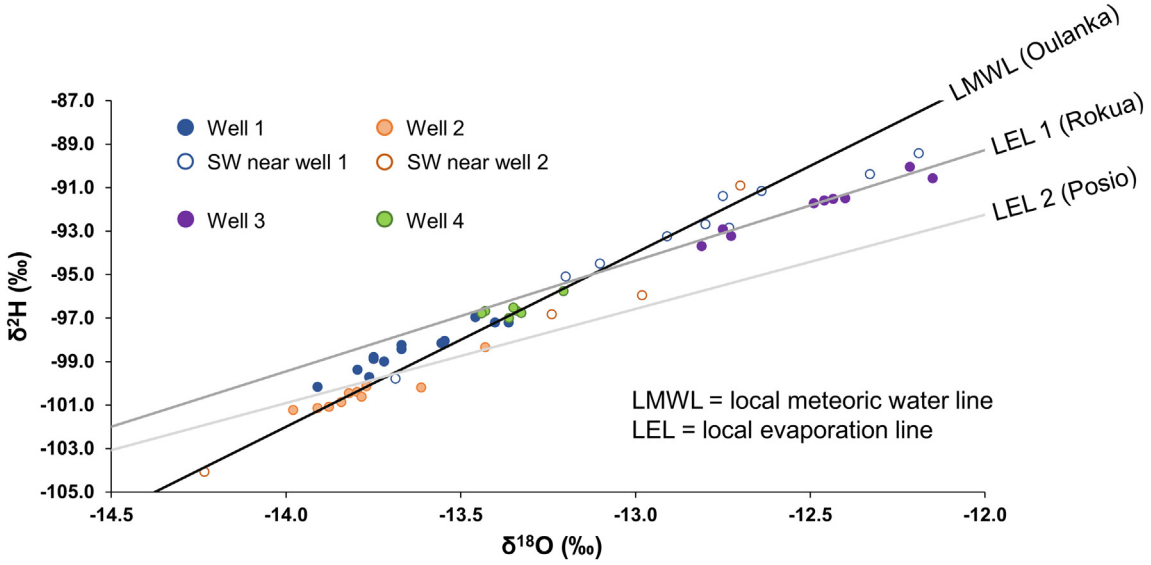


Fig. 4. Stable water isotope results from the wells compared to surface water samples from well 1 and 2 surroundings and to rainfall. Data for Oulanka local meteoric water line (LMWL) and Rokua and Posio local evaporation lines (LEL) were taken from previous studies (Isokangas et al., 2015; Nora et al., 2019; Rossi et al., 2015). Samples near LMWL indicate similarity to the isotopic signature of meteoric water, and samples near the LELs indicate similarity to the isotopic signature of surface water that has been subjected to evaporation (i.e. enriched in the heavy isotopes ^{2}H and ^{18}O). SW = surface water.

nearby surface waters – ditches, a small stream, and a river – mostly also varied along the LMWL line (indicating possible groundwater sources for these surface waters), with only two or three surface water samples showing signs of evaporation. Hence, the possibility of surface water intrusion at well 1 can be neither confirmed nor refuted. Given that the isotopic signatures of the surface water samples mostly did not show strong signs of evaporation, further investigation and/or other methods may be necessary to assess the possibility of surface water intrusion at this site.

At well 2, the isotopic signatures of most groundwater samples varied along the LMWL, as at well 1, with two samples deviating slightly more from the LMWL than the others. The isotopic signatures of two of the surface water samples followed the LMWL, but two others deviated somewhat. At this well, it was also not possible to confirm nor refute the possibility of surface water intrusion using isotopes, because during springtime flooding the isotopic signatures of the groundwater and the nearby surface water (river) can both be influenced by snowmelt. Hence, similarly to well 1, further investigation and/or other methods may be necessary to assess the possibility of surface water intrusion at this site.

At well 3, clear signs of evaporation were observed in the isotopic signatures of most of the groundwater samples, including all samples taken between late summer and autumn (late August–late November 2019). This strongly suggests that surface water which has been subjected to evaporation in summer has entered the well.

At well 4, the isotopic signatures of groundwater samples varied along the LMWL (with less spread than other wells), and no clear signs of evaporation were observed. The slight variation in the signal and the placement of the results on more negative part of the oxygen-hydrogen axis suggests that the main source of groundwater is snowmelt and/or cool period precipitation.

3.5. 16S rRNA amplicon sequencing data

3.5.1. High-level observations

Phylum- and class-level 16S taxonomic profiles varied between wells (Fig. 5A and B). DNA-derived and RNA-derived 16S rRNA amplicons varied at each well, indicating a difference between dormant and metabolically active communities. The DNA- and RNA-derived amplicon profiles of some wells remained rather similar across seasons, whereas others varied more. Overall, at the phylum level, DNA-derived 16S rRNA amplicons had high relative abundances of Patescibacteria (median relative abundance, MRA = 39.6 %), whereas RNA-derived 16S rRNA amplicons had high relative abundances of Proteobacteria (MRA = 30.1 %); these comprised mostly Parcubacteria (MRA = 29.4 %) and Gammaproteobacteria (MRA = 26.2 %) at the class level, respectively. Alpha diversity metrics for all samples and controls are shown in Table S7. Alpha diversity metrics for well 3 were relatively low compared to the other wells, especially when considering cDNA libraries. nMDS plots also showed well 3 to be an outlier among the studied wells, and revealed correlations between microbial communities and metadata variables (Fig. S7).

3.5.2. Specific taxa of interest

High relative abundances of specific taxa of interest were detected in several wells and timepoints (Table S6). For well 1, all DNA-derived profiles were quite similar at the phylum level, as were the RNA-derived profiles, regardless of season (Fig. 5). Well 2, however, had more noticeable differences between springtime and autumntime samplings. For the springtime samples, relatively high relative abundances of chloroplast sequences from the freshwater microalgal genus *Epipyxis* were observed in DNA- (7.9 %) and RNA-derived (26.9 %) amplicon libraries (responsible for the bulk of the Cyanobacteria blocks shown in Fig. 5), as well as lower relative abundances of chloroplast sequences from the freshwater algal species *Neotessella volvocina* (5.4 % and 2.5 %, respectively). For well 2, the springtime RNA-derived amplicon library also had moderate to high relative abundances of the Gram-negative bacterial genera *Polaromonas* (13.4 %), *Polynucleobacter* (5.5 %), and *Zoogloea* (5.3 %); and the autumntime RNA-derived amplicon library had a high relative abundance of the nitrite-oxidizing bacterial genus, *Nitrospira* (11.2 %). The springtime DNA-derived amplicon library for well 3 had relatively high levels of the class *Clostridia* (23.2 %), of which a considerable portion was the *Blautia* genus (10.9 %), a mammalian gut bacterium which may indicate faecal contamination. The springtime and autumntime RNA-derived libraries from well 3 were dominated very strongly by Proteobacteria (spring: 75.1 %, autumn: 79.8 %). All amplicon libraries from well 3, both DNA- and RNA-derived, had high relative abundances of the *Gallionella* genus of iron-oxidizing bacteria (range: 16.5–72.9 %). The autumntime DNA-derived library from well 4 had a high relative abundance of the Parcubacteria superphylum (31.7 %), which has previously been shown to dominate DNA-derived 16S rRNA gene amplicons in groundwater environments (Bruno et al., 2017;

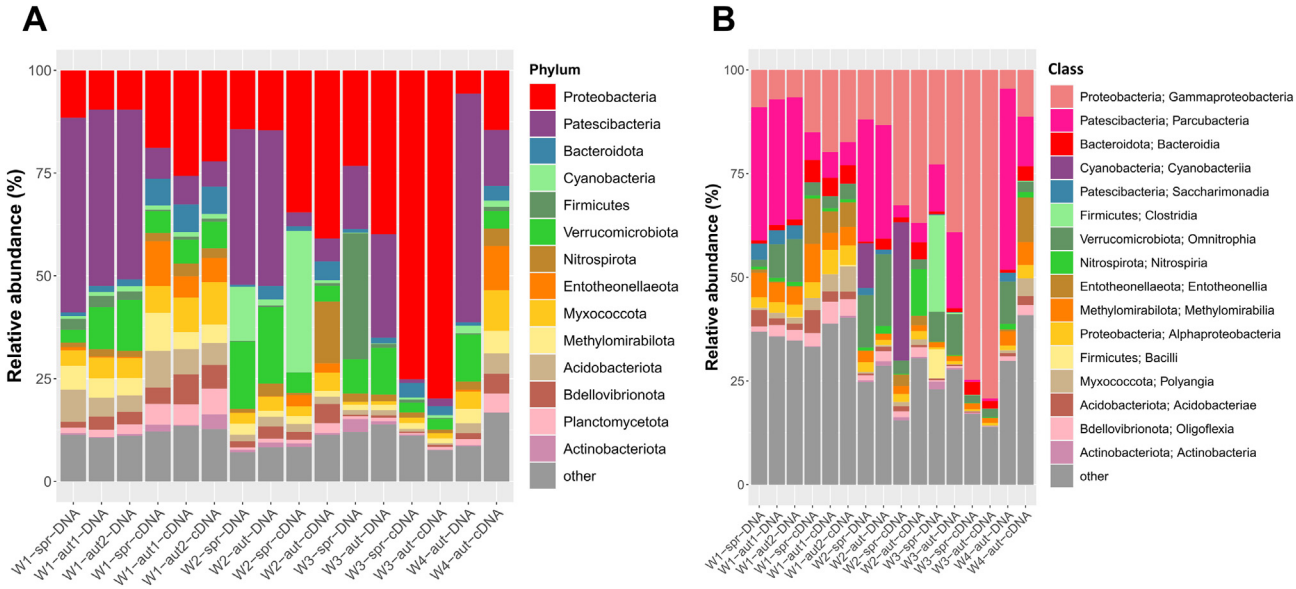


Fig. 5. Phylum- (A) and class-level (B) taxonomic classifications for 16S rRNA amplicon sequencing data. Only the most abundant microbial taxa are shown (those which were present at ≥5 % in at least one true sample or control sample). Abbreviations: W1, W2, W3, W4 = well 1, well 2, well 3, well 4; spr = springtime (sample taken in late April or early May); aut = autumntime; aut1 = sample taken in early October; aut or aut2 = sample taken in late October.

Herrmann et al., 2019; Kumar et al., 2017; Schwab et al., 2017). No springtime libraries were sequenced for well 4, and so a seasonal comparison could not be made.

3.5.3. Control samples

Compared with true samples, 16S rRNA libraries for negative control samples had lower read counts (Table S8) and noticeably different taxonomic profiles including higher relative abundances of groups such as Bacteroidia, Saccharimonadia, Alphaproteobacteria and Actinobacteria (Fig. S8). Taxa which were relatively abundant in controls but not in the true samples are shown in Table S9 to the highest possible taxonomic resolution. The abundant presence of these taxa in all controls, and their near-total absence in all true samples, as well as the overall taxonomic similarity between negative controls, suggests a considerable difference in taxonomic profile between controls and samples, indicating that bacterial carry-over from the large volume sampling and sample processing protocols was not a major issue in this study. Moreover, principal component analysis (PCoA) of the Bray-Curtis dissimilarity matrix revealed clear differences between true samples and controls (Fig. S9).

3.6. Well-specific summaries and potential risks

A summary of site-specific load sources, observations, and potential risks is given in Table 3.

3.6.1. Well 1

Findings from periodic manual measurements and sampling for well 1 were largely within expected ranges, with the possible exception of a peak of heterotrophic bacteria (2700 CFU/mL) in April. In the real-time data, there were several unexpected increases in groundwater level throughout the summer, which appeared to occur soon after heavy precipitation events, and these may suggest periodic intrusion of surface water into the well. There are many forestry ditches near this well, and a small stream discharging from a nearby peatland also flows within about 70 m

Table 3
Summary of site-specific load sources, observations, and potential risks.

Well	Load sources	Observations	Potential risks
1	● Surface water intrusion from nearby peatland	● Unusual variation in DO and groundwater level during summer ● Surface water intrusion leading to changes in microbiological and chemical quality of the groundwater ● Heterotrophic bacteria peak in April	● Surface water intrusion impacting chemical and microbiological groundwater quality
2	● Surface water intrusion from nearby river	● Surface water intrusion leading to changes in microbiological quality of the groundwater ● Peaks of SSRC, coliform bacteria, Gram-negative cDNA and GenBac3 in springtime ● Chloroplast sequences in 16S data could indicate surface water intrusion in springtime ● Peaks of heterotrophic bacteria and coliform bacteria in autumn	● Surface water intrusion impacting chemical and microbiological groundwater quality
3	● Surface water intrusion from nearby lakes ● Potential runoff from nearby agriculture ● Recreational use of the area (e.g. horses, dogs) ● High iron content in the soil	● Very problematic site (chronic chemical and microbiological water quality problems) ● Isotopes indicate surface water intrusion ● High iron concentration and high turbidity ● SSRC often detected ● Coliform bacteria peak in June, heterotrophic bacteria peak in November ● Gram-negative DNA most abundant at this well ● High relative abundances of an iron-oxidizing bacterium and a mammalian gut bacterium detected in 16S data (the latter could indicate faecal contamination)	● Surface water intrusion impacting chemical and microbiological groundwater quality ● Naturally high levels of iron and turbidity in the groundwater
4	● Built-up area ● Surface water runoff ● Nearby road	● Surface water not likely to be a source of contamination, based on isotope measurements	● Surface water intrusion impacting chemical and microbiological groundwater quality ● Cl^- in the water (unpublished data)

DO = dissolved oxygen, SSRC = spores of sulphite-reducing clostridia.

of the well. The level of the groundwater in the well appears to follow the variation in the level of this small stream. Thus, the risk posed by surface water intrusion at the well should not be discounted. However, with the exception of some unusual DO variability, no other abnormalities were observed in the real-time data. Isotopic signatures of untreated groundwater samples from this well did not show any obvious signs of evaporation. Taxonomic profiles of 16S amplicon libraries from DNA and RNA were different, but quite stable across seasons, with no clear indications of surface water intrusion or faecal contamination. Overall, the methods applied here could neither confirm nor refute the presence of surface water intrusion or faecal contamination at this site.

3.6.2. Well 2

It was known before this study that, at well 2, there is a risk of surface water intruding into the groundwater during the springtime snowmelt period every year. Microbiological indicator data from untreated groundwater samples appear to suggest some kind of change of conditions in the groundwater around this time: annual peaks were observed in springtime for coliform bacteria (1300 CFU/L in June), SSRC (1 CFU/L in May), Gram-negative cDNA gene copies (9.93×10^6 GC/100 mL in May), and general faecal indicator Bacteroidales bacteria (GenBac3) cDNA gene copies (152 GC/100 mL in May). In addition, high relative abundances of chloroplast sequences from the freshwater algae *Epipyxis* (7.8 % and 26.8 % in DNA and RNA, respectively) and *Neotessella volvocina* (5.4 % and 2.5 %) were observed in springtime 16S libraries (algae are known surface water indicators) (Gollnitz et al., 2003; Moulton-Hancock et al., 2000; Robertson and Edberg, 1997; USEPA, 1992). However, isotopic signatures of untreated groundwater samples from this well did not show any obvious signs of surface water intrusion. Peaks of heterotrophic bacteria (3100 CFU/mL in October) and coliform bacteria (110 CFU/L in November) in autumntime may indicate influence on the groundwater from autumntime rainfall. More support for this idea comes from the noticeable increase in groundwater level evident around this time (increasing about 20 cm between late-October and late-November, and then continuing to increase about another 20 cm before the end of the year). Overall, microbiological indicator data and 16S rRNA amplicon sequencing data appear to suggest some groundwater quality changes in springtime, possibly due to the known problem of springtime surface water intrusion, and rainfall may also be influencing groundwater quality in autumntime.

3.6.3. Well 3

Well 3 was problematic for a number of reasons. The groundwater in this well had the highest median temperature, turbidity, UV_{254} absorbance, Fe and Mn values, and the lowest DO values of all studied wells. The high Fe concentrations – which presented a considerable challenge to the interpretation of real-time online measurements – are presumably because the well is situated in a clay-rich coastal area of Finland, of a kind that is often associated with acid sulphate soils that can leach metals like Fe (Fältmarsch et al., 2008). The YSI multiparameter water quality sondes were mounted on flow-through columns in this study so that there would be no additional risk to the water going to the users from the implementation of the project. However, this approach proved challenging at well 3 and the column and the analyser in it had to be cleaned regularly. As for microbiological indicators, SSRC were detected at relatively high levels at most timepoints in this well (peaking in August and September; 20 CFU/L each time), and the GenBac3 MST marker (at levels below the quantification limit) and somatic coliphages (1 PFU/100 mL) were detected in April and May respectively, indicating faecal contamination. Well 3 was the only well in which somatic coliphages were detected, and almost the only well in which SSRC and GenBac3 were detected. Gram-negative DNA gene copies were also greatest in well 3 (springtime: 3.23×10^5 GC/100 mL; autumntime: 2.72×10^5 GC/100 mL). Counts of coliform bacteria peaked in June (185 CFU/L) and heterotrophic bacteria in November (1500 CFU/L). Clear signs of evaporation were observed in the isotopic signatures of most of the untreated groundwater samples collected during this study, strongly suggesting that surface water which has been subjected to evaporation in summer had entered the well. All amplicon libraries from well 3, both DNA- and RNA-derived, had high levels of the *Gallionella* genus of iron-oxidizing bacteria (range: 16.5–72.9 %). The springtime DNA-derived amplicon library for well 3 had relatively high levels of the class Clostridia (23.2 %), of which a considerable portion was the *Blautia* genus (10.8 %), a mammalian gut bacterium which may indicate faecal contamination. Overall, the simultaneous presence of high Fe levels, microbiological indicators such as SSRC, GenBac3 and somatic coliphages, a strong probability of surface water intrusion based on stable water isotopes, and high relative abundances of iron-oxidizing bacteria and mammalian gut bacteria suggest considerable problems with water quality at this site, and in fact, this well was finally shut down for good on 20.2.2021 due to these persistent problems.

3.6.4. Well 4

Well 4 had the highest median EC and pH values in this study. The high EC values may be due to a high chloride content in the groundwater (unpublished data from the North Savo Environmental Centre). No prominent peaks of conventional microbiological indicators were detected for this well, and isotopic signatures of untreated groundwater samples from this well did not show any obvious signs of evaporation. The autumntime DNA-derived library from well 4 had a high relative abundance of Parcubacteria (31.7 %), but this is not unusual as the Patescibacteria superphylum has previously been shown to dominate DNA-derived 16S rRNA gene amplicons in groundwater environments (Bruno et al., 2017; Herrmann et al., 2019; Kumar et al., 2017; Schwab et al., 2017). No springtime libraries were sequenced for well 4, and so a seasonal comparison could not be made. Overall, the behaviour of this well was perhaps the least unusual of all wells studied here, and no surface water intrusion or faecal contamination was detected.

4. Discussion

4.1. Evaluation of real-time online monitoring and EWS as implemented in this study

In establishing the real-time online monitoring systems in this study, we were able to assess (i) the quality of untreated groundwater in the wells, (ii) the reliability of the real-time online monitoring measurements, and (iii) potential challenges to sensor performance (e.g. high turbidity). We also demonstrated how real-time online monitoring data of groundwater quality can be collected, transferred, stored, and processed efficiently. The real-time data were largely in fair agreement with periodic manual on-site and laboratory measurements, suggesting that real-time online monitoring is a valid monitoring approach for these wells. However, we also investigated any deviations in the real-time data and tried to explain them. Some deviations were due to maintenance-related reasons, and were not therefore true groundwater quality anomalies. Nevertheless, the ability of the real-time online monitoring systems to capture these events suggests that they are also likely effective in capturing true anomalies.

4.2. Evaluation of stable water isotope analyses as implemented in this study

Measurements of stable water isotopes (i.e. $^2H/^1H$ and $^{18}O/^{16}O$ ratios) were most informative at wells 3 and 4. The isotopic signatures of untreated groundwater samples from well 3 indicated surface water intrusion, and those from wells 1, 2 and 4 did not. Prior to this study, the groundwater quality in well 4 was thought to be influenced by a nearby lake. Although surface water samples were not collected from the lake during this study, isotopic signatures of lake samples taken in August 2021 exhibited strong signs of evaporation (unpublished data from the North Savo Environmental Centre), meaning that stable water isotope analysis is a good method to detect possible surface water intrusion at this site. Nevertheless, signs of evaporation were not detected in the well either during our study or in August 2021. Hence, intrusion of water from this surface water body into the well can probably be ruled out. At wells 1 and 2, even the samples from nearby surface waters at times did not exhibit strong signs of evaporation

(especially during the snowmelt period), and hence, problematic intrusion of surface water into the groundwater at these sites could be neither confirmed or refuted by isotopic analyses.

4.3. Potential indicator taxa from 16S rRNA amplicon sequencing libraries

4.3.1. Epipyxis and Neotessella volvocina as potential indicators of surface water intrusion at well 2

Chloroplast sequences from the freshwater microalgal genus *Epipyxis* and the freshwater algal species *Neotessella volvocina* were observed in the springtime DNA- (relative abundance: 7.9 % and 5.4.%, respectively) and RNA-derived (26.9 % and 2.5 %, respectively) amplicon libraries from well 2. Algae are known surface water indicators (Gollnitz et al., 2003; Moulton-Hancock et al., 2000; Robertson and Edberg, 1997; USEPA, 1992), and are not often found in groundwater as they typically require light for photosynthesis. However, well 2 is close to a river, and there is a known risk of surface water flowing towards the well in the springtime each year. Hence, the detection of algae or chloroplast sequences in the groundwater could be a useful indicator of surface water intrusion at this site in springtime. More support for this idea comes from a recent study which showed that relative abundance values for chloroplast sequences increase during ice cover in the Keweenaw Waterway in Michigan, USA, reaching a peak just after ice melt (Butler et al., 2019). The regions of northern and central Finland where our study sites are located have very similar climates to Michigan, so it is not entirely unreasonable to assume that levels of chloroplast sequences in Finnish lakes and rivers might also reach a peak in springtime, around the time of ice melt. These high levels could make detection of chloroplast sequences in nearby groundwaters more likely in springtime. High proportions of *Epipyxis* have also previously been reported in surface water samples from an acidic opencast pit lake in Sherlovaya Gora, Russia (Gavrilov et al., 2019). Both of these studies used 16S rRNA gene amplicon sequencing, and made use of the SILVA rRNA database for assigning taxonomic classifications, as here. However, given that 16S rRNA amplicon sequencing is not typically the method of choice for studying algae, the potential use of algae as surface water indicators at the sites studied here could be further investigated in future through the use of methods such as quantitative real-time PCR (qPCR) or biological activity reaction tests (ALGE-BART™, Droycon Bioconcepts Inc., Regina, Canada).

4.3.2. Gallionella and Blautia at well 3

All amplicon libraries from well 3, both DNA- and RNA-derived, had high levels of the *Gallionella* genus of iron-oxidizing bacteria (range: 16.5–72.9 %). This was not entirely surprising as this well also had low DO, high turbidity, and high Fe and Mn levels, and observable brown, iron-related staining on piping in the groundwater well works. Problems associated with iron-oxidizing bacteria are a common nuisance in groundwater works (Emerson and De Vet, 2015). However, a recent study found that *Gallionella* correlated with non-operational groundwater wells in eastern Russia (Braun et al., 2016). This correlation also agrees quite nicely with the history of well 3 in our study, which was previously bypassed for extended periods (e.g. from 29.11.2017–16.5.2018, and from 31.7.2018–7.2.2019). The springtime DNA-derived amplicon library for well 3 had relatively high levels of the class Clostridia (23.2 %), of which a considerable portion was the *Blautia* genus (10.8 %), a mammalian gut bacterium which may indicate faecal contamination. Members of the class Clostridia are also anaerobic, so their presence in largely aerobic groundwater (the median DO for well 3 was 3 mg/L) may indicate recent intrusion. However, the DO for well 3 also sometimes dropped below 1, which may indicate that anaerobic conditions sometimes prevail in this well. Well 3 was also the only well where spores of sulphite-reducing clostridia (SSRC) were detected at every timepoint (Fig. 1C).

4.4. Potential for more widespread use of these methods in groundwater quality monitoring

Many previous studies have described the development and use of continuous or real-time methods for monitoring groundwater bodies (Calderwood et al., 2020; Drage and Kennedy, 2020; Oppus et al., 2020; Velasquez-Orta et al., 2017). However, larger-scale implementations tend to focus almost exclusively on the monitoring of groundwater level (Calderwood et al., 2020; Drage and Kennedy, 2020; Lee et al., 2007), suggesting that the potential public health benefits of monitoring physicochemical parameters have not yet been fully realised. Long-term continuous or real-time monitoring of physicochemical parameters could be implemented at shallow groundwater wells in at least two ways: (i) the well managers could establish a monitoring system themselves, or (ii) the work could be outsourced to an environmental monitoring company. The start-up costs (e.g. sensors, data transmission equipment, installation, and training) and running costs (e.g. equipment maintenance, calibration, and data transmission) of the first option could be within the budgets of larger water suppliers, but unfeasible for small-scale community water suppliers such as those featured in this study. For smaller suppliers, outsourcing may be a more realistic option. Alternatively, water suppliers could partner with local authorities, universities, or other research institutes to conduct continuous or real-time monitoring on a temporary basis to investigate specific problems – this was the kind of approach adopted in our study.

The analysis of stable water isotopes is often an effective means of identifying signs of surface water intrusion in groundwater wells (Hunt et al., 2005). Many commercial laboratories can perform the analysis, which requires only a small volume of water, and the cost has fallen with the transition from isotope ratio mass spectrometry to the now cheaper laser-based technologies (Stumpp et al., 2018), making it feasible for small-scale community water suppliers to incorporate this method into their regular monitoring programmes. Again, as with the continuous or real-time monitoring, water suppliers could alternatively partner with local authorities, universities, or other research institutes that have the necessary equipment and training to conduct the analysis – either on a short-term or long-term basis.

Modern molecular biology methods such as 16S rRNA amplicon sequencing and shotgun metagenomics can provide valuable insights into the composition and spatiotemporal variation of microbial communities in aquatic environments (Clark et al., 2018). These methods provide a big-picture overview of the communities and their diversity, but can also be used to identify potential indicator taxa that may be of special relevance to suspected site-specific water quality issues (e.g. potential indicators of faecal contamination or surface water intrusion). The indicator approach has a long history in water quality monitoring (Saxena et al., 2015), and these modern molecular methods could be more widely implemented as a way of corroborating and extending the findings of conventional indicator-based analyses. Once potential indicator taxa have been identified, qPCR could also be applied to measure absolute abundances (Converse et al., 2009; Haugland et al., 2010; Kärkkäinen et al., 2010; Pitkänen et al., 2013). Water suppliers could avail of these methods via commercial laboratories, local authorities, or via universities and other research institutes.

Overall, the potential usefulness of these supplementary methods should be assessed by water suppliers on a case-by-case basis, in conjunction with the available resources and with an understanding of suspected site-specific problems.

4.5. Study limitations and recommendations for future work

This study had several limitations, which could be remedied in future work. Firstly, although the real-time online monitoring programme implemented in this study was largely successful, the water quality sondes should ideally be calibrated and maintained more often during long-term monitoring to ensure measurement accuracy. Good quality measurement data is of primary importance for the functioning of a real-time online monitoring system or EWS. Hence, we recommend that researchers consider the following matters during the planning phase of their projects, in conjunction with the available financial and time resources: (i) identification of a representative measurement point, (ii) correct installation of measuring instruments,

(iii) measurement accuracy of the selected sensors, (iv) reliability of data transfer between sensor and data logger, (v) reliability of data transfer between data logger and cloud, (vi) pre-processing of raw data to remove potential noise in the measurements. Secondly, the analysis of stable water isotopes to investigate groundwater–surface water interactions in this study was limited in at least two ways: (i) water samples from nearby surface water bodies were taken only for wells 1 and 2 (not 3 and 4; although a lake near well 4 was sampled for isotopes in August 2021), and (ii) the surface water sources near wells 1 and 2 did not exhibit very strong signs of evaporation, limiting the ability of isotopic analysis to reveal surface water intrusion in the groundwater. In future, samples should ideally be taken from all wells and nearby surface water sources, and, in cases where the nearby surface waters do not exhibit very strong signs of evaporation, alternative methods for detecting surface water intrusion should also be explored. Thirdly, no springtime 16S libraries were sequenced for well 4, and so a seasonal comparison of 16S taxonomic profiles could not be made for that site; this was an unavoidable consequence of samples being lost in transit to the sequencing company. Finally, given that threats to groundwater quality can vary in a site-specific manner (as shown in this and many other studies) we recommend that future groundwater quality monitoring programmes identify and focus on parameters most likely to prove useful in detecting potential site-specific water quality problems.

5. Conclusion

The aim of this study was to examine the potential benefits of enhancing the conventional monitoring of shallow groundwater wells with (i) real-time online monitoring, (ii) stable water isotope analysis, and (iii) 16S rRNA amplicon sequencing. The idea was that expanding the diversity of monitoring methods applied could provide a more diverse array of data, helping to build a more comprehensive understanding of the studied systems. Our work revealed some of the potential challenges of using these methods in shallow groundwater wells, but also how these methods can be used to enhance water quality monitoring by adding new insights and corroborating the findings of conventional methods.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.161199.

CRediT authorship contribution statement

Kevin J. Lyons: Conceptualization, Formal analysis, Investigation, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Jenni Ikonen:** Conceptualization, Investigation, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Anna-Maria Hokajärvi:** Conceptualization, Investigation, Data curation, Visualization, Writing – review & editing. **Teemu Räsänen:** Conceptualization, Investigation, Data curation, Visualization, Writing – review & editing. **Tarja Pitkänen:** Conceptualization, Investigation, Writing – review & editing, Supervision, Project administration. **Ari Kauppinen:** Conceptualization, Investigation, Writing – review & editing. **Katharina Kujala:** Conceptualization, Investigation, Writing – review & editing, Supervision. **Pekka M. Rossi:** Conceptualization, Investigation, Data curation, Visualization, Writing – review & editing, Supervision, Project administration. **Ilkka T. Miettinen:** Conceptualization, Investigation, Data curation, Visualization, Writing – review & editing, Supervision, Project administration.

Data availability

Data will be made available on request.

Declaration of competing interest

Ilkka T. Miettinen reports financial support was provided by Finnish Ministry of Agriculture and Forestry (Maa- ja metsätalousministeriö). Kevin J. Lyons reports financial support was provided by Land and Water Technology Support Association (Maa- ja vesitekniikan tuki ry). Kevin J. Lyons reports financial support was provided by KAUTE Foundation (Kaupallisten ja teknillisten tieteiden tukisäätiö KAUTE).

Acknowledgements

The authors acknowledge the well operators for their cooperation and participation in the study, and the CSC: IT Centre for Science (CSC: tietotekniikan keskus) for providing computational resources. This work received significant financial support from the Finnish Ministry of Agriculture and Forestry (Maa- ja metsätalousministeriö), as a Blue Bioeconomy Government Key Project entitled “Ensuring the safety of groundwater through real-time monitoring”. K.J.L. was additionally supported by personal grants from the Land and Water Technology Support Association (Maa- ja vesitekniikan tuki ry) (project ID: 4261 Vesihuolto) and the KAUTE Foundation (Kaupallisten ja teknillisten tieteiden tukisäätiö KAUTE) (project ID: 20201135).

References

Allen, M.J., Edberg, S.C., Reasoner, D.J., 2004. Heterotrophic plate count bacteria—what is their significance in drinking water? Int. J. Food Microbiol. 92, 265–274. https://doi.org/10.1016/j.ijfoodmicro.2003.08.017.

Anderson, C.W., 2005. National Field Manual for the Collection of Water-Quality Data. U.S. Geological Survey Techniques of Water-Resources Investigations, Book 9. Chapter A6. Section 6.7. Turbidity. U.S. Geological Survey, Reston, VA https://doi.org/10.3133/twri09A6.7.

Banna, M.H., Imran, S., Francisque, A., Najjaran, H., Sadiq, R., Rodriguez, M., Hoorfar, M., 2014. Online drinking water quality monitoring: review on available and emerging technologies. Crit. Rev. Environ. Sci. Technol. 44, 1370–1421. https://doi.org/10.1080/10643389.2013.781936.

Bernhard, A.E., Field, K.G., 2000a. A PCR assay to discriminate human and ruminant feces on the basis of host differences in bacteroides-prevotella genes encoding 16S rRNA. Appl. Environ. Microbiol. 66, 4571–4574. https://doi.org/10.1128/AEM.66.10.4571-4574.2000.

Bernhard, A.E., Field, K.G., 2000b. Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Appl. Environ. Microbiol. 66, 1587–1594. https://doi.org/10.1128/AEM.66.4.1587-1594.2000.

Bokulich, N.A., Kaehler, B.D., Rideout, J.R., Dillon, M., Bolyen, E., Knight, R., Huttley, G.A., Gregory Caporaso, J., 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome 6, 90. https://doi.org/10.1186/s40168-018-0470-z.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K.B., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.-X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A.V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37, 852–857. https://doi.org/10.1038/s41587-019-0209-9.

Braun, B., Schröder, J., Knecht, H., Szewzyk, U., 2016. Unraveling the microbial community of a cold groundwater catchment system. Water Res. 107, 113–126. https://doi.org/10.1016/j.watres.2016.10.040.

Brester, C., Ryzhikov, I., Siponen, S., Jayaprakash, B., Ikonen, J., Pitkänen, T., Miettinen, I.T., Torvinen, E., Kolehmainen, M., 2020. Potential and limitations of a pilot-scale drinking water distribution system for bacterial community predictive modelling. Sci. Total Environ. 717, 137249. https://doi.org/10.1016/j.scitotenv.2020.137249.

Bruno, A., Sandionigi, A., Rizzi, E., Bernasconi, M., Vicario, S., Galimberti, A., Cocuzza, C., Labra, M., Casiraghi, M., 2017. Exploring the under-investigated “microbial dark matter” of drinking water treatment plants. Sci. Rep. 7, 44350. https://doi.org/10.1038/srep44350.

Burri, N.M., Weatherl, R., Moeck, C., Schirmer, M., 2019. A review of threats to groundwater quality in the anthropocene. Sci. Total Environ. 684, 136–154. https://doi.org/10.1016/j.scitotenv.2019.05.236.

Butler, T.M., Wilhelm, A.-C., Dwyer, A.C., Webb, P.N., Baldwin, A.L., Techtmann, S.M., 2019. Microbial community dynamics during lake ice freezing. Sci. Rep. 9, 6231. https://doi.org/10.1038/s41598-019-42609-9.

Calderwood, A.J., Pauloo, R.A., Yoder, A.M., Fogg, G.E., 2020. Low-cost, open source wireless sensor network for real-time, scalable groundwater monitoring. Water 12, 1066. https://doi.org/10.3390/w12041066.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: high-resolution sample inference from illumina amplicon data. Nat. Methods 13, 581–583. https://doi.org/10.1038/nmeth.3869.

Capodaglio, A.G., Callegari, A., 2009. Online monitoring technologies for drinking water systems security. In: Hlavinek, P., Popovska, C., Marsalek, J., Mahrikova, I., Kukharchyk, T. (Eds.), Risk Management of Water Supply and Sanitation Systems, NATO Science for Peace and Security Series C: Environmental Security. Springer Netherlands, Dordrecht, pp. 153–179 https://doi.org/10.1007/978-90-481-2365-0_15.

Chik, A.H.S., Emelko, M.B., Anderson, W.B., O'Sullivan, K.E., Savio, D., Farnleitner, A.H., Blaschke, A.P., Schijven, J.F., 2020. Evaluation of groundwater bacterial community composition to inform waterborne pathogen vulnerability assessments. Sci. Total Environ. 743, 140472. https://doi.org/10.1016/j.scitotenv.2020.140472.

Clark, D.R., Ferguson, R.M.W., Harris, D.N., Matthews Nicholass, K.J., Prentice, H.J., Randall, K.C., Randell, L., Warren, S.L., Dumbrell, A.J., 2018. Streams of data from drops of water: 21st century molecular microbial ecology. WIREs Water 5, e1280. https://doi.org/10.1002/wat2.1280.

Converse, R.R., Blackwood, A.D., Kirs, M., Griffith, J.F., Noble, R.T., 2009. Rapid QPCR-based assay for fecal bacteroides spp. As a tool for assessing fecal contamination in recreational waters. Water Res. 43, 4828–4837. https://doi.org/10.1016/j.watres.2009.06.036.

Dick, L.K., Field, K.G., 2004. Rapid estimation of numbers of fecal bacteroidetes by use of a quantitative PCR assay for 16S rRNA genes. Appl. Environ. Microbiol. 70, 5695–5697. https://doi.org/10.1128/AEM.70.9.5695-5697.2004.

Drage, J., Kennedy, G., 2020. Building a low-cost, internet-of-things, real-time groundwater level monitoring network. Groundwater Monit. R 40, 67–73. https://doi.org/10.1111/gwmr.12408.

Emerson, D., De Vet, W., 2015. The role of FeOB in engineered water ecosystems: a review. J. Am. Water Works Assoc. 107, E47–E57. https://doi.org/10.5942/jawwa.2015.107.0004.

Fältmarsch, R.M., Åström, M.E., Vuori, K.-M., 2008. Environmental risks of metals mobilised from acid sulphate soils in Finland: a literature review. Boreal Environ. Res. 13, 444–456.

Famiglietti, J.S., 2014. The global groundwater crisis. Nature Clim Change 4, 945–948. https://doi.org/10.1038/nclimate2425.

Favere, J., Buysschaert, B., Boon, N., De Gusseme, B., 2020. Online microbial fingerprinting for quality management of drinking water: full-scale event detection. Water Res. 170, 115353. https://doi.org/10.1016/j.watres.2019.115353.

Figueras, M.J., Borrego, J.J., 2010. New perspectives in monitoring drinking water microbial quality. IJERPH 7, 4179–4202. https://doi.org/10.3390/ijerph7124179.

Fricker, E.J., Illingworth, K.S., Fricker, C.R., 1997. Use of two formulations of colilert and QuantiTrayTM for assessment of the bacteriological quality of water. Water Res. 31, 2495–2499. https://doi.org/10.1016/S0043-1354(96)00342-9.

Gat, J.R., 1996. Oxygen and hydrogen isotopes in the hydrologic cycle. Annu. Rev. Earth Planet. Sci. 24, 225–262. https://doi.org/10.1146/annurev.earth.24.1.225.

Gat, J., 2010. Isotope hydrology: a study of the water cycle. Series on Environmental Science and Management. Imperial College Press, London https://doi.org/10.1142/p027.

Gavrilov, S.N., Korzhenkov, A.A., Kublanov, I.V., Bargiela, R., Zamana, L.V., Popova, A.A., Toshchakov, S.V., Golyshin, P.N., Golyshina, O.V., 2019. Microbial communities of polymetallic deposits' acidic ecosystems of continental climatic zone with high temperature contrasts. Front. Microbiol. 10, 1573. https://doi.org/10.3389/fmicb.2019.01573.

Gollnitz, W.D., Clancy, J.L., Whitteberry, B.L., Vogt, J.A., 2003. RBF as a microbial treatment process. J. Am. Water Works Assoc. 95, 56–66. https://doi.org/10.1002/j.1551-8833.2003.tb10511.x.

Greenberg, A.E., Clesceri, L.S., Eaton, E.D., 1992. Standard Methods for the Examination of Water and Wastewater. 18thedn. American Public Health Association, Washington DC, USA.

Halko, N., Martinsson, P.-G., Shkolnisky, Y., Tygert, M., 2011. An algorithm for the principal component analysis of large data sets. SIAM J. Sci. Comput. 33, 2580–2594. https://doi.org/10.1137/100804139.

Haugland, R.A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L., Shanks, O.C., 2010. Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected bacteroidales species and human fecal waste by qPCR. Syst. Appl. Microbiol. 33, 348–357. https://doi.org/10.1016/j.syapm.2010.06.001.

Herlemann, D.P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., Andersson, A.F., 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME J 5, 1571–1579. https://doi.org/10.1038/ismej.2011.41.

Herrmann, M., Wegner, C.-E., Taubert, M., Geesink, P., Lehmann, K., Yan, L., Lehmann, R., Totsche, K.U., Küsel, K., 2019. Predominance of cand. Patescibacteria in groundwater is caused by their preferential mobilization from soils and flourishing under oligotrophic conditions. Front. Microbiol. 10, 1407. https://doi.org/10.3389/fmicb.2019.01407.

Højris, B., Kornholt, S.N., Christensen, S.C.B., Albrechtsen, H.-J., Olesen, L.S., 2018. Detection of drinking water contamination by an optical real-time bacteria sensor. H2Open J. 1, 160–168. https://doi.org/10.2166/h2oj.2018.014.

Hunt, R.J., Coplen, T.B., Haas, N.L., Saad, D.A., Borchardt, M.A., 2005. Investigating surface water–well interaction using stable isotope ratios of water. J. Hydrol. 302, 154–172. https://doi.org/10.1016/j.jhydrol.2004.07.010.

Ikonen, J., Pitkänen, T., Kosse, P., Ciszek, R., Kolehmainen, M., Miettinen, I.T., 2017. On-line detection of Escherichia coli intrusion in a pilot-scale drinking water distribution system. J. Environ. Manag. 198, 384–392. https://doi.org/10.1016/j.jenvman.2017.04.090.

Isokangas, E., Rozanski, K., Rossi, P.M., Ronkanen, A.-K., Kløve, B., 2015. Quantifying groundwater dependence of a sub-polar lake cluster in Finland using an isotope mass balance approach. Hydrol. Earth Syst. Sci. 19, 1247–1262. https://doi.org/10.5194/hess-19-1247-2015.

Isomäki, E., Valve, M., Kivimäki, A.-L., 2006. Small waterworks in Finland. Presented at the 5th Nordic Drinking Water Conference, 5th Nordic Drinking Water Conference. Reykjavík, Iceland, pp. 91–95.

Isomäki, E., Valve, M., Kivimäki, A.-L., Lahti, K., Suomen ympäristökeskus, 2008. Operation and Maintenance of Small Waterworks. Finnish Environment Institute: Edita Publishing, Helsinki.

Kärkkäinen, P.M., Valkonen, M., Hyvärinen, A., Nevalainen, A., Rintala, H., 2010. Determination of bacterial load in house dust using qPCR, chemical markers and culture. J. Environ. Monit. 12, 759–768. https://doi.org/10.1039/B917937B.

Kauppinen, A., Pitkänen, T., Al-Hello, H., Maunula, L., Hokajärvi, A.-M., Rimhanen-Finne, R., Miettinen, I.T., 2019. Two drinking water outbreaks caused by wastewater intrusion including sapovirus in Finland. IJERPH 16, 4376. https://doi.org/10.3390/ijerph16224376.

Kim, H., Kaown, D., Mayer, B., Lee, J.-Y., Hyun, Y., Lee, K.-K., 2015. Identifying the sources of nitrate contamination of groundwater in an agricultural area (Haean basin, Korea) using isotope and microbial community analyses. Sci. Total Environ. 533, 566–575. https://doi.org/10.1016/j.scitotenv.2015.06.080.

Konikow, L.F., Kendy, E., 2005. Groundwater depletion: a global problem. Hydrogeol. J. 13, 317–320. https://doi.org/10.1007/s10040-004-0411-8.

Kumar, S., Herrmann, M., Thamdrup, B., Schwab, V.F., Geesink, P., Trumbore, S.E., Totsche, K.-U., Küsel, K., 2017. Nitrogen loss from pristine carbonate-rock aquifers of the hainich critical zone exploratory (Germany) is primarily driven by chemolithoautotrophic anammox processes. Front. Microbiol. 8, 1951. https://doi.org/10.3389/fmicb.2017.01951.

Lee, J.-Y., Kwon, K., 2016. Current status of groundwater monitoring networks in Korea. Water 8, 168. https://doi.org/10.3390/w8040168.

Lee, J.-Y., Yi, M.-J., Yoo, Y.-K., Ahn, K.-H., Kim, G.-B., Won, J.-H., 2007. A review of the national groundwater monitoring network in Korea. Hydrol. Process. 21, 907–919. https://doi.org/10.1002/hyp.6282.

Lyons, K.J., Hokajärvi, A.-M., Ikonen, J., Kauppinen, A., Miettinen, I.T., Pitkänen, T., Rossi, P.M., Kujala, K., 2021. Surface water intrusion, land use impacts, and bacterial community composition in shallow groundwater wells supplying potable water in sparsely populated areas of a boreal region. Microbiol Spectr, e00179-21 https://doi.org/10.1128/Spectrum.00179-21.

Macler, B.A., Merkle, J.C., 2000. Current knowledge on groundwater microbial pathogens and their control. Hydrogeol. J. 8, 29–40. https://doi.org/10.1007/PL00010972.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 17, 10–12. https://doi.org/10.14806/ej.17.1.200.

Moulton-Hancock, C., Rose, J.B., Vasconcelos, G.J., Harris, S.I., Klonicki, P.T., Sturbaum, G.D., 2000. Giardia and cryptosporidium occurrence in groundwater. J. Am. Water Works Assoc. 92, 117–123. https://doi.org/10.1002/j.1551-8833.2000.tb09010.x.

Nora, J., Rossi, P., Sanaksenaho, R., Lindholm, A., 2019. Lapin POSKI2 – hankkeen erillisselvitys: Isotooppitutkimukset (POSKI2 project, Lapland – Separate Study on Water Isotopes. Geological Survey of Finland, Groundwater Unit, Rovaniemi, Finland.

Oksanen, J., 2020. vegan: Community Ecology Package. R Package.

Oppus, C., Guico, M.L., Claro Monje, J., Leah Guzman Annael Domingo, M.A., Ngo, G., Retirado, M.G., Chris Kwong, J., 2020. Remote and real-time sensor system for groundwater level and quality. 2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE). Presented at the 2020 IEEE Eurasia Conference on IOT. Communication and Engineering (ECICE), IEEE, Yunlin, Taiwan, pp. 152–155 https://doi.org/10.1109/ECICE50847.2020.9301948.

Parlov, J., Kovač, Z., Nakić, Z., Barešić, J., 2019. Using water stable isotopes for identifying groundwater recharge sources of the unconfined alluvial Zagreb aquifer (Croatia). Water 11, 2177. https://doi.org/10.3390/w11102177.

Pitkänen, T., Ryu, H., Elk, M., Hokajärvi, A.-M., Siponen, S., Vepsäläinen, A., Räsänen, P., Santo Domingo, J.W., 2013. Detection of fecal bacteria and source tracking identifiers in environmental waters using rRNA-based RT-qPCR and rDNA-based qPCR assays. Environ. Sci. Technol. 47, 13611–13620. https://doi.org/10.1021/es403489b.

Pitkänen, T., Juselius, T., Isomäki, E., Miettinen, I., Valve, M., Kivimäki, A.-L., Lahti, K., Hänninen, M.-L., 2015. Drinking water quality and occurrence of Giardia in Finnish small groundwater supplies. Resources 4, 637–654. https://doi.org/10.3390/resources4030637.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596. https://doi.org/10.1093/nar/gks1219.

Reasoner, D.J., Geldreich, E.E., 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl. Environ. Microbiol. 49, 1–7. https://doi.org/10.1128/AEM.49.1.1-7.1985.

Robertson, J.B., Edberg, S.C., 1997. Natural protection of spring and well drinking water against surface microbial contamination. I. Hydrogeological parameters. Crit. Rev. Microbiol. 23, 143–178. https://doi.org/10.3109/10408419709115134.

Robeson, M.S., O'Rourke, D.R., Kaehler, B.D., Ziemski, M., Dillon, M.R., Foster, J.T., Bokulich, N.A., 2020. RESCRIPt: reproducible sequence taxonomy reference database management. PLoS Comput. Biol. 17, e1009581. https://doi.org/10.1371/journal.pcbi.1009581.

Rossi, P.M., Marttila, H., Jyväsjärvi, J., Ala-aho, P., Isokangas, E., Muotka, T., Kløve, B., 2015. Environmental conditions of boreal springs explained by capture zone characteristics. J. Hydrol. 531, 992–1002. https://doi.org/10.1016/j.jhydrol.2015.11.009.

Sampat, P., 2000. Deep Trouble: The Hidden Threat of Groundwater Pollution (No. Worldwatch Paper 154). Worldwatch Institute, Washington, DC.

Saxena, G., Bharagava, R.N., Kaithwas, G., Raj, A., 2015. Microbial indicators, pathogens and methods for their monitoring in water environment. J. Water Health 13, 319–339. https://doi.org/10.2166/wh.2014.275.

Schwab, V.F., Herrmann, M., Roth, V.-N., Gleixner, G., Lehmann, R., Pohnert, G., Trumbore, S., Küsel, K., Totsche, K.U., 2017. Functional diversity of microbial communities in pristine aquifers inferred by PLFA- and sequencing-based approaches. Biogeosciences 14, 2697–2714. https://doi.org/10.5194/bg-14-2697-2017.

Siefring, S., Varma, M., Atikovic, E., Wymer, L., Haugland, R.A., 2008. Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different

instrument and reagent systems. J. Water Health 6, 225–237. https://doi.org/10.2166/wh.2008.022.
Smith, S.W., 2003. Digital signal processing: a practical guide for engineers and scientists. Demystifying Technology Series. Newnes, Amsterdam; Boston.
Sørensen, T.A., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on danish commons. Biol. Skar. 5, 1–34.
Storey, M.V., van der Gaag, B., Burns, B.P., 2011. Advances in on-line drinking water quality monitoring and early warning systems. Water Res. 45, 741–747.
Stumpp, C., Brüggemann, N., Wingate, L., 2018. Stable isotope approaches in vadose zone research. Vadose Zone J. 17, 180096. https://doi.org/10.2136/vzj2018.05.0096.
Turunen, K., Räsänen, T., Hämäläinen, E., Hämäläinen, M., Pajula, P., Nieminen, S.P., 2020. Analysing contaminant mixing and dilution in river waters influenced by mine water discharges. Water Air Soil Pollut. 231, 317. https://doi.org/10.1007/s11270-020-04683-y.
USEPA, 1992. Consensus Method for Determining Groundwaters Under the Direct Influence of Surface Water Using Microscopic Particulate Analysis (MPA) (No. EPA 910/9-92-029). USEPA, Port Orchard, WA, USA.
Velasquez-Orta, S.B., Werner, D., Varia, J.C., Mgana, S., 2017. Microbial fuel cells for inexpensive continuous in-situ monitoring of groundwater quality. Water Res. 117, 9–17. https://doi.org/10.1016/j.watres.2017.03.040.
Wada, Y., van Beek, L.P.H., van Kempen, C.M., Reckman, J.W.T.M., Vasak, S., Bierkens, M.F.P., 2010. Global depletion of groundwater resources. Geophys. Res. Lett. 37, L20402. https://doi.org/10.1029/2010GL044571.
WHO, 2017. Water Quality and Health - Review of Turbidity: Information for Regulators and Water Suppliers.
Winter, T.C., Harvey, J.W., Franke, O.L., Alley, W.M., 1998. Ground water and surface water: a single resource (Circular). Circular. 1139. USGS Circular. https://doi.org/10.3133/cir1139.
Zektser, I.S., Everett, L.G., 2004. Groundwater Resources of the World and Their Use. UNESCO, Paris. http://unesdoc.unesco.org/images/0013/001344/134433e.pdf.