Water Research 269 (2025) 122812

Contents lists available at ScienceDirect

Water Research

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

Physicochemical and isotopic similarity between well water and intruding surface water is not synonymous with similarity in prokaryotic diversity and community composition

Kevin J. Lyons [a,*], Vadim Yapiyev [a,1], Kaisa Lehosmaa [b], Anna-Kaisa Ronkanen [c], Pekka M. Rossi [a], Katharina Kujala [a]

[a] *Water, Energy and Environmental Engineering Research Unit, University of Oulu, Oulu, Finland*
[b] *Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland*
[c] *Finnish Environment Institute, Marine and Freshwater Solutions, Oulu, Finland*

ARTICLE INFO

Keywords:
Groundwater
Drinking water
Water supply
Risk assessment
Stable water isotopes
Microbiology

ABSTRACT

Intruding surface water can impact the physicochemical and microbiological quality of groundwater. Understanding these impacts is important because groundwater provides much of the world's potable water, and reduced quality is a potential public health risk. In this study, we monitored six shallow groundwater wells and three surface water bodies in the North Ostrobothnia region of Finland twice monthly for 12 months (October 2021–October 2022) via (i) on-site and off-site measurements of physicochemical water quality parameters, (ii) determination of stable water isotope compositions, and (iii) analysis of microbial communities (via amplicon sequencing of the V3–V4 16S rRNA gene sub-regions). Water from one well showed clear overall physicochemical and isotopic similarity with a nearby pond, as well as temporal fluctuations in water temperature and isotopes that mirrored those of the pond. Isotope mixing analyses suggested that about 80–95 % of the well water comes from the pond. Such large-scale intrusion might be expected to reduce prokaryotic diversity and composition in the aquifer, either by strong influx of surface water taxa or changes to aquifer physicochemistry. Compared to the pond, however, prokaryotic communities from the well showed significantly higher alpha diversity and a composition more similar to a nearby well unaffected by intrusion. The finding that physicochemical and isotopic similarity between well water and intruding surface water is not synonymous with similarity in prokaryotic diversity and community composition makes clear the need for a multi-method approach when studying the impact of surface water intrusion on shallow wells.

1. Introduction

Groundwater provides much of the world's potable water (Sampat, 2000; Zektser and Everett, 2004). Thus, maintaining good groundwater quality is an important public health goal (Figueras and Borrego, 2010; Macler and Merkle, 2000). One potential cause of poor groundwater quality is the intrusion of water from a nearby surface water body – such as a river, lake, or pond – towards a well used for potable water supply (Ascott et al., 2016; Wett et al., 2002; Yang et al., 2023). In many cases, a natural attenuation process takes place in the subsurface – consisting of filtration, biodegradation, adsorption, chemical precipitation, and redox reactions – which can lessen the impact of the intruding surface water on the groundwater (Hiscock and Grischek, 2002; Maliva and Missimer, 2012; Ray et al., 2003). Sometimes, however, the intrusion occurs in such a way, or at such a high rate, that the usual biotic and abiotic subsurface mechanisms cannot function effectively (Ascott et al., 2016; Wett et al., 2002; Yang et al., 2023). In these cases, solid and dissolved materials, as well as microbes may be carried into the groundwater in greater concentrations, thereby influencing both the physicochemical and microbiological quality of the groundwater (Li et al., 2021; Ritter et al., 2002). If the impacts of surface water intrusion are particularly evident and persistent, the affected groundwater might be said to be 'groundwater under the direct influence' of surface water (GUDI) (Chaudhary et al., 2009; Chin and Qi, 2000; Cullimore, 2008; Nnadi and

---

[*] Corresponding author at: PO Box 4300, 90014 University of Oulu, Finland.
E-mail address: kevin.lyons@oulu.fi (K.J. Lyons).
[1] Present address: School of Mining and Geoscience, Nazarbayev University, Astana, Kazakhstan.

https://doi.org/10.1016/j.watres.2024.122812
Received 26 April 2024; Received in revised form 14 November 2024; Accepted 16 November 2024

Fulkerson, 2002; Vasconcelos and Harris, 1992).

Many methods exist to detect and study the effects of surface water intrusion in groundwater wells. For example, measurements of physicochemical water quality parameters (e.g. temperature, pH, electrical conductivity, dissolved oxygen, redox potential, turbidity, and the concentrations of various chemical species and nutrients) can provide useful basic information about the potability of the groundwater and about the kinds of influences the surrounding area has on the well (Edmunds and Shand, 2008; Fitts, 2013; WHO, 2011). Another method involves the determination of stable water isotope compositions, which can be used to relate water samples to different phases of the hydrological cycle (Han et al., 2019; Hunt et al., 2005; Kortelainen, 2011, 2007; Rosenberry et al., 2015; Yang et al., 2023). Due to isotopic fractionation processes associated with evaporation, precipitation waters tend to be depleted with respect to the heavy isotopes ^{2}H and ^{18}O, whereas lakes and oceans tend to be enriched (Gat, 2010, 1996). Thus, different water sources can often (though not always) be distinguished from one another on the basis of their isotope compositions, and simple mixing models can estimate the contributions of different sources to a given water sample. Yet another method involves the analysis of microbial communities from water samples (e.g. via amplicon sequencing of 16S rRNA gene sub-regions) to gain insights into the composition and functional potentials of microbial communities in the studied water bodies and wells (Chik et al., 2020; Clark et al., 2018; Kim et al., 2015). Combining several methods and data types can lead to a diverse understanding of the groundwater–surface water interactions and other factors influencing water quality at various studied sites.

The aim of this study was to investigate the nature of the unintended surface water intrusion suspected to be affecting three groundwater wells in the North Ostrobothnia region of Finland. We applied a variety of methods to this end, including (i) on-site and off-site measurements of physicochemical water quality parameters, (ii) determination of stable water isotope compositions, and (iii) analysis of microbial communities (via amplicon sequencing of the V3–V4 16S rRNA gene sub-regions). We attempted to confirm or refute the presence of surface water intrusion in the chosen wells, as well as to investigate the temporal effects of this intrusion on physicochemical quality and microbiological communities.

2. Materials and methods

2.1. Selection of study areas and descriptions

We investigated three sparsely populated study areas in the North Ostrobothnia region of Finland (labelled Area A, Area B, and Area C here) (Fig. 1). Based on data from the Finnish Meteorological Institute, the annual mean precipitation in this region is about 575 mm, and the annual mean temperature is about +3 °C, although hourly temperature typically ranges between about –25 °C and +25 °C throughout the year. There are approximately five and a half months of snow cover annually, with the main period of snow cover typically lasting from November/ December until April/May.

The three study areas were chosen because they each contain a groundwater well which, based on the findings of previous studies, is suspected to be under the influence of surface water intrusion (Lyons et al., 2023, 2021; Yapiyev et al., 2023); these wells are hereafter known as the 'suspect' wells. In each of the three areas, we also investigated a surface water body located near the suspect well and thought to be the source of the intrusion, as well as a 'comparison' well located farther from the surface water body and thought to be free from its influence. Hence, our investigation involved nine sampling/measurement sites in total: six groundwater wells and three surface water bodies. All wells investigated here are used for potable water supply and provide 65–750 m^{3} of water per day to networks serving a few hundred to several thousand users. For Area A, Area B, and Area C, respectively, the studied surface water body is a river (with a mean flow rate of about 100 m^{3}/s), a water-filled disused gravel extraction pit (i.e. a gravel pit pond), and a

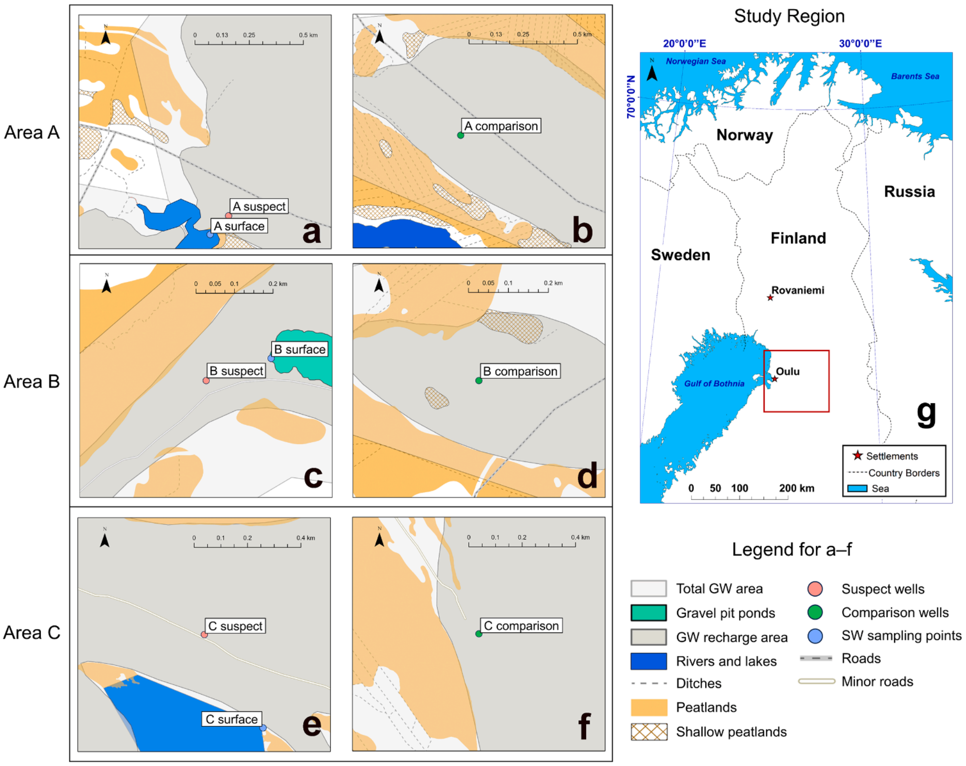

Fig. 1. Maps showing the locations of the nine sampling/measurement sites within the three study areas (a–f) and the approximate study region, shown as a red square (g). The outlines of the surface water bodies have been artificially simplified to preserve the confidentiality of the well locations. Abbreviations: GW = groundwater; SW = surface water

lake (Fig. S1). All wells are located in sand-and-gravel aquifers, and in most cases the aquifers are 4–7 m above their surroundings (e.g. peatlands or surface water), though the aquifers containing suspect well A and comparison well C are about 24 m and 20 m above their surroundings, respectively. Cross-sections of the suspect wells are shown in Fig. S2.

2.2. Groundwater and surface water sampling

Six groundwater wells and three surface water bodies were monitored twice-monthly for 12 months (October 2021–October 2022) via (i) on-site and off-site measurements of physicochemical water quality parameters, (ii) determination of stable water isotope compositions, and (iii) analysis of microbial communities (via amplicon sequencing of the V3–V4 16S rRNA gene sub-regions). Exact sampling dates are given in Supplementary Methods S1.

During each sampling round, untreated groundwater was collected from a groundwater sampling tap in each of the six wells (after flame-sterilising the tap), and surface water was collected from the three surface water bodies via aseptic methods which varied depending on time of year: when the surface water bodies were free of ice, surface water was collected by repeatedly filling a steam-sterilised 1 L polypropylene bottle attached to the end of a 4 m long metal sampling pole; and during ice-cover periods, an ice drill was used to make a hole in the ice and either (i) an ethanol-sterilised 2 L Limnos water sampler (GWM-Engineering Oy, Kuopio, Finland) or (ii) a steam-sterilised 1 L polypropylene bottle held in an ethanol-sterilised gloved hand was used to retrieve water from below the ice. Surface water samples were collected from approximately the same location during every sampling round regardless of the time of year. At each groundwater well and surface water body, water was collected in (i) a clean plastic bucket (to enable on-site measurements of physicochemical water quality parameters), (ii) a 1 L polypropylene bottle (to enable off-site measurements of physicochemical water quality parameters), (iii) 2 × 15 mL conical polypropylene tubes (filled with no headspace; one to enable off-site analysis of turbidity, and one to enable off-site analysis of stable water isotopes), and (iv) 3 × 500 mL sterile low-density polyethylene Whirl-Pak® bags (to enable off-site analysis of microbial communities). All collected samples were transported from the sampling sites to the laboratory in polystyrene cooler boxes containing cold freezer blocks, and were then stored in the dark at 4 °C, with the exception of the Whirl-Pak® bags, which were stored at −20 °C.

2.3. On-site and off-site measurements of physicochemical water quality parameters and stable water isotopes

Temperature, pH, redox potential (ORP), dissolved oxygen (DO), and electrical conductivity (EC), were measured on site from the water collected in the clean plastic bucket, using a MultiLine® Multi 3630 IDS portable multi-parameter meter and associated sensors (WTW, Weilheim, Germany). All other physicochemical water quality measurements were performed off site, at either the University of Oulu or Oulanka Research Station (a regional unit of the University of Oulu). All redox potential values reported in this work have been converted to standard hydrogen electrode values.

The turbidity of the water samples was measured using a TB 210 IR infrared turbidity meter (Lovibond® Water Testing, Dortmund, Germany). The calcium (Ca), magnesium (Mg), sodium (Na), and potassium (K) concentrations in the water samples were measured via inductively coupled plasma–optical emission spectrometry (ICP-OES), and the chloride (Cl^-) and sulfate (SO_4^{2-}) concentrations were measured via ion chromatography (IC). The total alkalinity of the water samples (as $CaCO_3$ and as HCO_3^-) was measured using a YSI 9500 photometer (YSI, Ohio, USA) via the 'Alkaphot' colorimetric method (an unpublished method developed by Palintest Ltd, UK). The iron (Fe) concentrations in the water samples were measured using a SPECTROstar® Nano microplate reader (BMG LABTECH, Ortenberg, Germany) via a colorimetric method involving 1,10-phenanthroline monohydrate and ascorbic acid (Mortatti et al., 1982). Total phosphorus (TP), total nitrogen (TN), ammonium nitrogen (NH_4^+-N), combined nitrate and nitrite nitrogen (NO_3^-+NO_2^--N), and phosphate (PO_4^{3-}) were measured according to manufacturer's instructions with a SEAL AA500 AutoAnalyzer (SEAL Analytical GmbH, Norderstedt, Germany), and dissolved organic carbon (DOC) was measured with a Shimadzu TOC-L (Shimadzu, Japan). Further details of the physicochemical methods are given in Supplementary Methods S2.

The stable water isotope composition of oxygen ($\delta^{18}O$) and hydrogen (δ^2H) was determined from water samples using cavity ring-down spectroscopy with a Picarro L2140-i analyser (Picarro, Inc., Santa Clara, CA, USA) at the University of Oulu. The results were normalized to VSMOW2 (Vienna Standard Mean Ocean Water) scale in parts per thousand (‰). Mean precision for $\delta^{18}O$ and δ^2H, as assessed by the Picarro ChemCorrect^{TM} software, was 0.027 ‰ and 0.181 ‰, respectively. The mean percent error for $\delta^{18}O$ and δ^2H was estimated to be 0.073 % and 0.417 %, respectively; these estimations were made by repeated measurement ($n = 22$) of each of two standards with known $\delta^{18}O$ and δ^2H values: (i) Tap water ($\delta^{18}O = -11.2$ ‰ and $\delta^2H = -84.5$ ‰), and (ii) Viro ($\delta^{18}O = -21$ ‰ and $\delta^2H = -156$ ‰) (Noor et al., 2023). These standards were measured at the beginning, middle, and end of each batch of samples (to monitor measurement quality).

2.4. Principal component analysis of physicochemical water quality data

Principal component analysis (PCA) of the physicochemical water quality data was conducted in R (R Core Team, 2024). The raw data were first cross-checked against notes made during sample collection and information obtained from the well managers, and dubious data points were removed. Missing values were then imputed using the imputePCA() function of the 'missMDA' R package, and the uncertainty of the imputed data was visualized by multiple imputation using the MIPCA() function of the same package (Josse and Husson, 2016). Plotted PC axes were very stable, indicating that the uncertainty of the imputed data was low. The PCA itself was carried out using the PCA() function of the 'FactoMineR' R package (Lê et al., 2008), resulting in plots of individuals (water samples) and variables. Final figures were made with the 'ggplot2' R package (Wickham, 2016).

2.5. Stable water isotope methods

The stable water isotope composition of precipitation for a region can be plotted in dual isotope ($\delta^{18}O$–δ^2H) space falling around a local meteoric water line (LMWL). The LMWL is a regression line of precipitation samples with a slope and intercept. In this study, we used the recently published precipitation-weighted least square regression (PWLSR) MWL for Northern Finland, based on long-term data from Rovaniemi (location shown in Fig. 1g) (Yapiyev et al., 2023). Local evaporation lines (LELs) from the Rokua and Posio areas were also used to help visualize the local evaporation effects (Isokangas et al., 2015; Nora et al., 2019).

We applied line-conditioned excess (lc-excess*) to identify what changes the isotope composition of water samples had undergone on the land surface (Landwehr and Coplen, 2004).

$$\text{lc-excess}^* = \left(\delta^2\text{H} - \text{a}\,8\delta^{18}\text{O} - \text{b}\right)/\text{S}. \tag{1}$$

where, *a* and *b* are the slope of and intercept of the MWL for Rovaniemi ($\delta^2H = 7.7915\delta^{18}O + 6.4552$), and *S* is the uncertainty parameter (Landwehr and Coplen, 2004). The uncertainty parameter value ($S = 1.27$) was taken from (Yapiyev et al., 2023), which used the same equipment.

In the case of surface water intrusion in a suspect well we used two-component end-member mixing analysis with lc-excess* as a tracer:

$$R_{surface} = \left(\text{lc-excess}^{*}{}_{(gw)} - \text{lc-excess}^{*}{}_{(source)}\right) / \left(\text{lc-excess}^{*}{}_{(surface)} - \text{lc-excess}^{*}{}_{(source)}\right) \tag{2}$$

where $R_{surface}$ is the fraction of surface water in the suspect well; and lc-excess* $_{(gw)}$, lc-excess* $_{(source)}$, and lc-excess* $_{(surface)}$ are isotopic compositions of water in the suspect well, a nearby comparison well (without intrusion), and a surface water body (from which the intrusion is suspected to arise), respectively.

2.6. Membrane filtration, DNA extractions, and 16S rRNA gene amplicon sequencing

To enable extraction of DNA for 16S rRNA gene amplicon sequencing, the water samples collected in the Whirl-Pak® bags were filtered in a Kojair BioWizard laminar flow hood. Samples were filtered via vacuum filtration through sterile porcelain Buchner filtering funnels equipped with 0.1 μm sterile polyethersulfone membrane filters (Sartorius, Göttingen, Germany). The median volume filtered for groundwater and surface water samples was 1100 mL and 550 mL, respectively. After filtering, the membrane filters were cut in half with a sterile scalpel and each half stored at −20 °C in a sterile 1.5 mL Eppendorf tube until extraction. DNA was extracted from half of each filter using the ZymoBIOMICS™ DNA Miniprep Kit (Zymo Research, CA, USA), and sent (with a negative extraction control) to Novogene (Cambridge, UK) for amplicon sequencing of the V3–V4 16S rRNA gene sub-regions. The primers used by Novogene for amplifying the V3–V4 sub-regions were 5’-CCTAYGGGRBGCASCAG-3’ and 5’-GGACTACNNGGGTATCTAAT-3’, which are similar but not identical to the Bakt_341F and Bakt_805R primers reported by (Herlemann et al., 2011). At Novogene, 250 bp paired-end reads were generated via sequencing on the Illumina NovaSeq™ 6000 platform and raw reads were trimmed to remove adapters and sequencing primers.

2.7. Processing of the 16S rRNA gene amplicon sequencing data

A total of 198 DNA extractions were sent to Novogene for 16S rRNA gene amplicon sequencing. Of these, the extractions from 129 water samples and 1 negative extraction control were successfully sequenced, and a total of >13,000,000 reads were generated (mean number of reads per sample was >100,000). Reads, as received from Novogene, were imported into QIIME 2™ (Bolyen et al., 2019), and all adapters, amplicon primers, and unknown nucleotides (Ns) remaining in the reads were removed using ‘qiime cutadapt trim-paired’ (Martin, 2011). The adapter trimming sequences 5’-AGATCGGAAGAGCACACGTCTGAACTCCAGTCA-3’ and 5’-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-3’ were used for adapter removal. Sequences were denoised by using ‘qiime dada2 denoise-paired’ with no additional trimming, thereby generating a table of amplicon sequence variants (ASVs) (Callahan et al., 2016).

2.8. Taxonomic classification and diversity analyses

Taxonomic classification of ASVs was performed 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., 2021). Alpha and beta diversity metrics were calculated via the QIIME 2™ ‘diversity core-metrics-phylogenetic’ command, using a sampling depth of 18,187 (Halko et al., 2011; Sørensen, 1948). For each alpha diversity metric, tests for significant differences between sampling/measurement sites were performed in R. Shapiro-Wilk tests were first used to assess whether the data for each sampling/measurement site were normally distributed. As not all data were normally distributed, Kruskal-Wallis tests were used to assess whether there were significant differences among the sites, and pairwise comparisons were performed using Dunn’s tests (via the dunn.test() function of the ‘dunn.test’ R package) producing adjusted p-values with Bonferroni correction for multiple comparisons (Dinno, 2022). Adjusted p-values < 0.05 in Dunn’s tests were considered indicative of significant pairwise differences between sites. Non-metric multidimensional scaling (nMDS) of the QIIME 2™ ASV table was performed via the metaMDS() function of the ‘vegan’ R package using the Bray–Curtis dissimilarity metric (Oksanen, 2020). Water quality variables were fit to the nMDS plots using the envfit() function from the same package. Only variables that were significantly correlated with the ordination axes (i.e. those identified in ‘envfit’ with a P value of < 0.05 following permutation test) were included in the figures. Final taxonomic barplot figures and diversity figures were made with the ‘ggplot2’ R package (Wickham, 2016). Significant differences between the sampling/measurements sites in each study area were sought via permutational multivariate analysis of variance (PERMANOVA) using the adonis2() function from the ‘vegan’ R package (Oksanen, 2020), and pairwise comparisons were performed by pairwise PERMANOVA using the pairwise.adonis2() function of the ‘pairwiseAdonis’ R package (Martinez Arbizu, 2020), with subsequent Bonferroni correction for multiple comparisons. The number of permutations was 999 for both PERMANOVA and pairwise PERMANOVA. Adjusted p-values < 0.05 were considered indicative of significant pairwise differences.

2.9. Sequencing data availability

The 16S amplicon sequencing data for this study – as received from Novogene – have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under primary accession number PRJEB66003 (https://www.ebi.ac.uk/ena/browser/view/PRJEB66003).

3. Results and discussion

3.1. Suspect wells A, B, and C showed weak, strong, and moderate physicochemical similarity with nearby surface water

Principal component analysis (PCA) of physicochemical water quality data revealed that the three suspect wells differed in their degree of physicochemical similarity with nearby surface water. In Area A, samples from the suspect well were distant to samples from the nearby river, suggesting weak physicochemical similarity (Fig. 2a). In Area B, samples from the suspect well overlapped completely with samples from the nearby gravel pit pond, suggesting strong physicochemical similarity (Fig. 2b). And in Area C, samples from the suspect well grouped somewhat closely to samples from the nearby lake, suggesting moderate physicochemical similarity (Fig. 2c).

The lack of physicochemical similarity between suspect well A and the nearby river was surprising, as this well had the shortest distance to surface water of the three suspect wells (~60 m vs. ~180 m and ~220 m), and water level data collected from this well and the nearby river showed that river level rose above well water level during the springtime snowmelt period in May 2022 (Fig. S3), suggesting increased subsurface flow of river water towards the well (Hunt et al., 2005; Winter et al., 1998). Nevertheless, no obvious physicochemical fluctuations were detected in suspect well A around that time. By contrast, a recent German study of surface water intrusion during the onset of riverbank filtration (RBF) in February 2017 found logical reductions in well water temperature, EC, and major ion concentrations; and increases

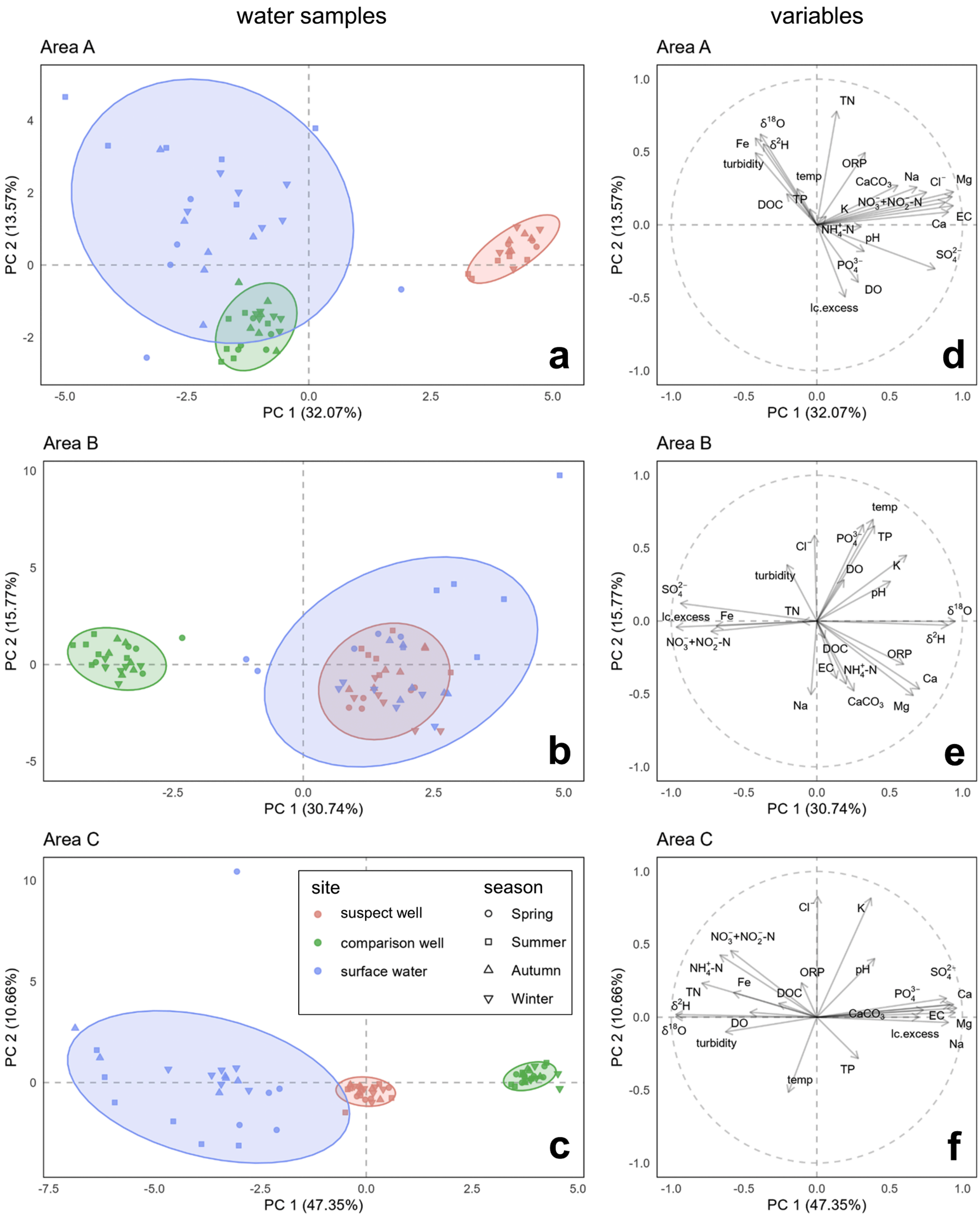

Fig. 2. Principal component analysis (PCA) plots for individuals (i.e. water samples) (a–c), and variables (d–f) from study areas A, B, and C, respectively. Seasons are winter (January–March), spring (April–June), summer (July–September) and autumn (October–December). The ellipses in plots a–c enclose approximately 95 % of the data points in each group.

in pH, DO, DOC, and nitrate (Fillinger et al., 2021). One possible explanation for the lack of fluctuations observed in suspect well A is that the pumping rate of this well during late April–early June 2022 was about 10 % of its typical rate. This may have prevented major effects of river water intrusion on well water physicochemistry during the springtime snowmelt period. The well is not usually pumped during late April–early June due to potential well water quality risks. During this study, however, some pumping was continued (without supplying the water to users) so that the interaction between surface water and groundwater could be examined.

PCA also revealed different degrees of physicochemical similarity between the comparison wells and nearby surface water. In Areas B and C, samples from the comparison well were distant from both suspect well and surface water, as expected under the assumption of little to no surface water influence (Fig. 2b and 2c). In Area A, however, samples from the comparison well grouped closely to samples from the nearby river suggesting unexpected physicochemical similarity, and thereby unexpectedly raising the possibility of surface water intrusion at comparison well A (Fig. 2a). However, isotope data (described in the next section) do not support this idea.

In all study areas, physicochemical water quality exhibited more seasonal variation in surface water bodies than in wells, as shown by the surface water ellipses being relatively larger than the well ellipses in PCA plots (Fig. 2a–c). Similarly, suspect well B also exhibited more variation than any of the other wells, presumably due to strong surface water influence (Fig. 2b).

Suspect well A had the highest EC, Mg, Cl$^-$, Na, K and NO$_3^-$+NO$_2^-$-N of all sites; comparison well B had the highest Fe; and comparison well C had the highest Ca and SO$_4^{2-}$ (Fig. S4). These variables were also among those that most distinguished these sites from others in the PCA (Fig. 2d–f).

3.2. *Suspect wells B and C showed strong and moderate indications of surface water intrusion in stable water isotope data*

Isotope compositions from suspect wells B and C were more enriched in ^{2}H and ^{18}O relative to all other wells and deviated to the right of the local meteoric water line (LMWL) in dual isotope plots, in closer alignment with the local evaporation lines (LELs) (Fig. 3b and 3c). Isotope data from suspect well B overlapped strongly with data from the nearby gravel pit pond (Fig. 3b), whereas data from suspect well C plotted approximately midway between data from the comparison well and data

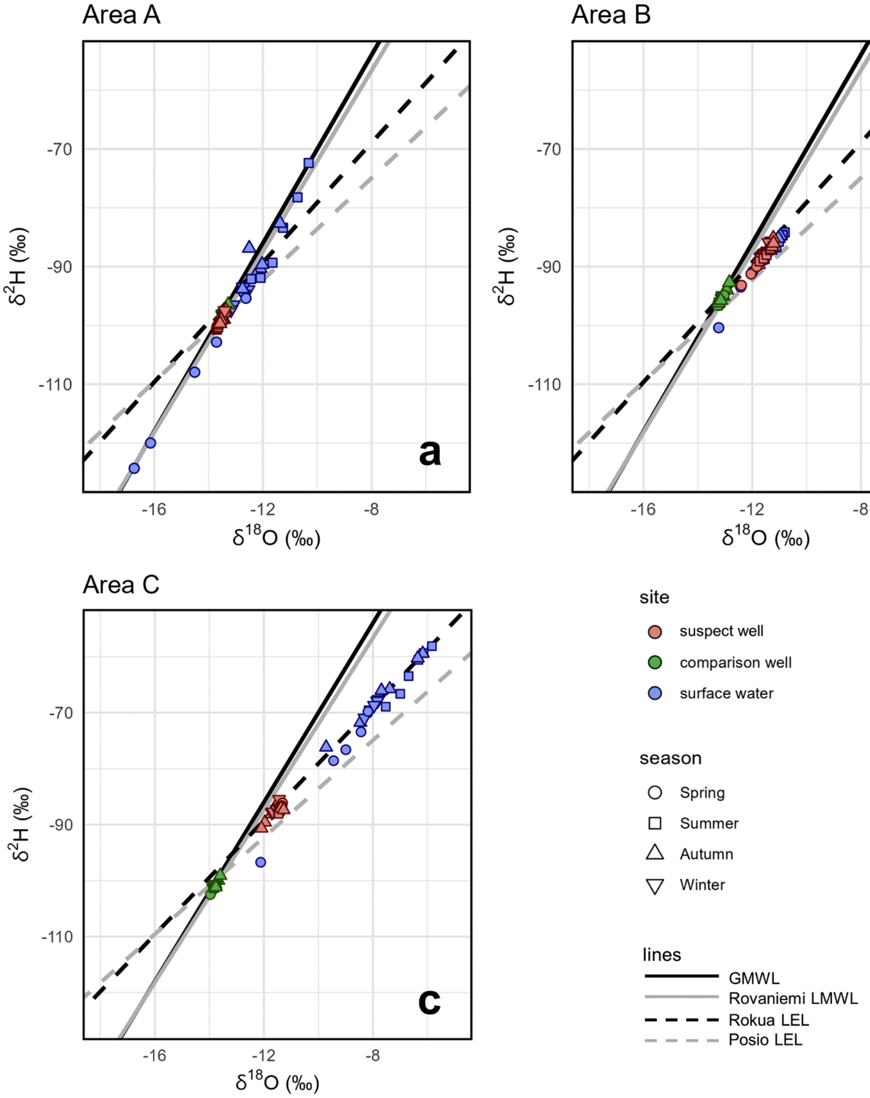

Fig. 3. The stable water isotope compositions, in dual isotope (δ^{18}O–δ^2H) space, of collected samples from wells and surface water bodies in Area A (a), Area B (b), and Area C (c). Abbreviations: GMWL = global meteoric water line; LMWL = local meteoric water line; LEL = local evaporation line. Equations for the GMWL, Rovaniemi LMWL, and Rokua and Posio LELs were taken from previous studies (Craig, 1961; Isokangas et al., 2015; Nora et al., 2019; Rossi et al., 2015; Yapiyev et al., 2023).

from the nearby lake (Fig. 3c). In other words, suspect wells B and C showed strong and moderate indications of surface water intrusion respectively.

Data from all comparison wells and suspect well A plotted along the LMWL. These wells had similar values and small standard deviations (Table S1). Of all studied sites, compositions from the lake were the most enriched in ^{2}H and ^{18}O (Fig. 3c).

Previous work has shown that values of lc-excess for precipitation and groundwater typically fluctuate around 0, whereas the lc-excess values of water bodies undergoing evaporation are negative (Liu et al., 2020). Here, mean lc-excess* values for suspect wells B and C were both below −3.3, indicating surface water intrusion at these wells (Table S1). Mean lc-excess* values for suspect well B and the nearby gravel pit pond were strikingly similar (−3.48 and −3.99), indicating a strong surface water influence on this well. Mean lc-excess* values for suspect well C and the nearby lake were less similar (−3.35 and −11.23), suggesting a more moderate surface water influence. Mean lc-excess* values for suspect well A and the three comparison wells ranged between −0.2 and +0.25, suggesting a relative lack of surface water influence, despite noticeable physicochemical similarity between the river and comparison well A in the PCA plot (Fig. 2a). Of all studied sites, compositions from the lake had the lowest mean values for δ^2H, δ^{18}O, and lc-excess*, indicating strong enrichment of ^{2}H and ^{18}O due to open-water evaporation.

Isotope mixing analyses suggested that about 80–95 % of the water in suspect well B comes from the nearby gravel pit pond (Fig. 4a), whereas about 25–40 % of the water in suspect well C comes from the nearby lake (Fig. 4b). Thus, the relationship between suspect well B and the nearby gravel pit pond could perhaps be described as a bank filtration system. This is something which had not previously been considered by the well managers at this site and is likely to have consequences for how the well is managed into the future. Notably, however, the well managers reported that routine physicochemical and microbiological monitoring has not revealed any significant health threats at this well, despite the strong intrusion – and that, in fact, this well has the best water quality of all the wells they manage in the region.

3.3. Fluctuations in temperature and isotopes indicated surface water intrusion at suspect well B

In Area B, a wave of increased water temperature was observed in the suspect well beginning in June and coming to a peak of 12.7 °C in September (the highest temperature recorded in any of the wells) (Fig. 5 and S3). This wave was likely caused by the intrusion of warmer water from the nearby gravel pit pond (Anderson, 2005; Foulquier et al., 2011; Hunt et al., 2005; Kalbus et al., 2006), and appears to be an annual, recurring phenomenon as indicated by the presence of a downward temperature trend at the start of the study period (late October 2021 to February 2022), and at the end (October 2022). Thus, water from this well appears to fit the USEPA definition of 'groundwater under the direct influence of surface water', which lists "significant and rapid shifts in turbidity, temperature, conductivity, or pH which closely correlate to climatological or surface water conditions" as one of the defining characteristics (Chaudhary et al., 2009; Chin and Qi, 2000). None of the other wells exhibited similar temperature fluctuations, having temperatures mostly between 4 and 9 °C throughout the year. Temperatures in all surface waters were near 0 °C from mid-November to early April, after which they began to rise, coming to an annual peak of between 20 and 22.5 °C in July before beginning to fall again.

Water temperature peaked (at 21.3 °C) in the gravel pit pond in late July, and peaked (at 12.7 °C) in suspect well B in early September – remaining high (>12.5 °C) until early-to-mid October (Fig. 5). If the intrusion of warmer water from the pond is the cause of increased well water temperature, then the travel time between pond and well during the summer–early autumn period may be approximately 6–10 weeks (Hunt et al., 2005).

A noticeable drop in δ^{18}O and δ^2H occurred in all surface waters around the time of the springtime snowmelt (during April–May) (Fig. 5). A similar drop occurred in suspect well B after a 3-week delay, possibly suggesting a shorter travel time between pond and well during this period. The drop in δ^{18}O and δ^2H in suspect well B was likely due to the connection with the pond (rather than the direct input of snowmelt water through the subsurface) because a similar drop was not seen at any of the other wells. δ^{18}O, δ^2H, and lc-excess* values from suspect well B were overall very similar to values from the nearby pond, suggesting a close groundwater–surface water connection, with values being near identical during November–March and slightly less similar during May–October (probably due to open-water evaporation in the pond). δ^{18}O, δ^2H, and lc-excess* values were relatively stable in the other wells, except for a gradual decline in lc-excess* in suspect well C (from about −2.3 to −4.7), which may indicate moderate long-term intrusion from the nearby lake.

3.4. Significantly higher prokaryotic alpha diversity in suspect well B relative to the nearby pond

Despite strong indications of surface water intrusion at suspect well B, prokaryotic communities in well water samples had significantly higher values of all alpha diversity metrics than communities from the nearby pond ($p < 0.05$) (Fig. 6). This might be considered surprising, given that much of the water from this well likely comes from the pond, and the prokaryotic communities of surface waters tend to be less diverse than those of groundwater (Fillinger et al., 2021; Ji et al., 2022a, 2022b). However, a recent study of a RBF system in a sand and gravel aquifer in Germany, which involved collection of water samples both before and after restarting the system after years of disuse, found that

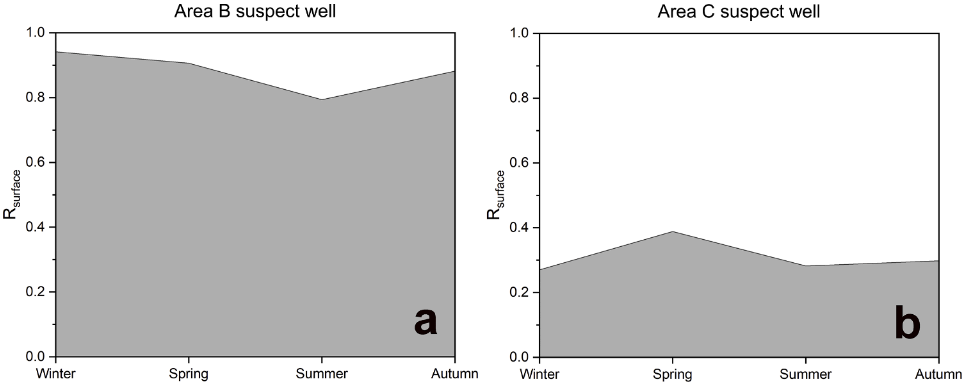

Fig. 4. The fraction of surface water in suspect well B (a) and suspect well C (b). Seasons are winter (January–March), spring (April–June), summer (July–September) and autumn (October–December). $R_{surface}$ was not calculated for suspect well A because no indications of intrusion were detected in dual-isotope plots.

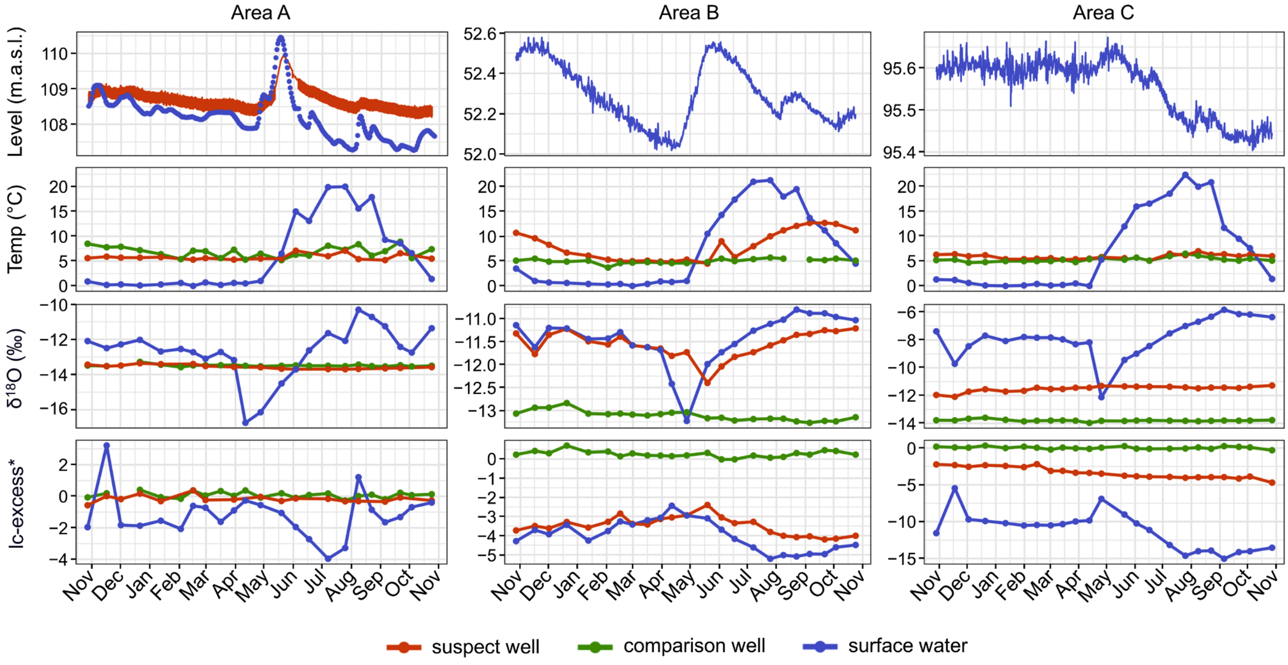

Fig. 5. Temporal variation in water level, water temperature, δ18O, and lc-excess* values for all sites throughout the study period. Note that, except in the case of temperature, the y-axis ranges are different for each study area (to better show the variation). No water level data were collected for wells in study areas B and C.

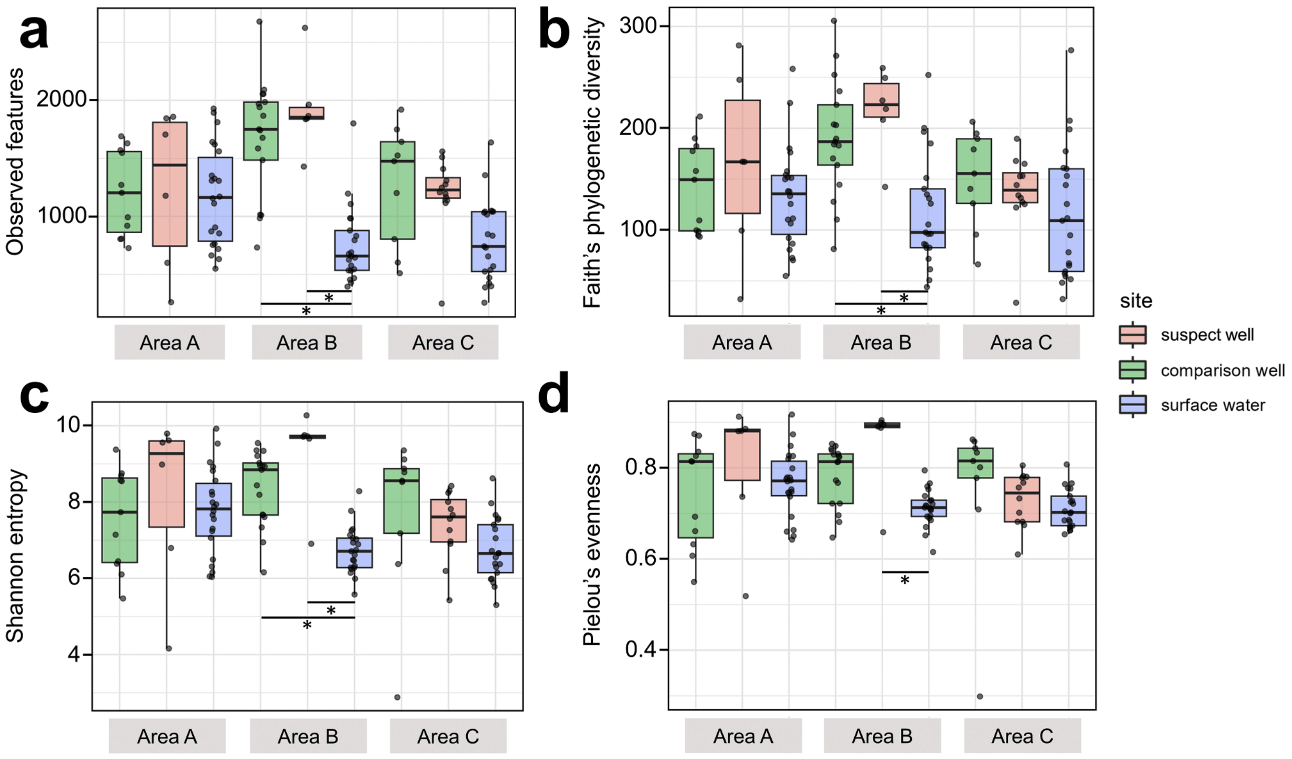

Fig. 6. Box and whisker plots for all wells and surface water bodies, showing values of alpha diversity metrics – (a) Observed features, (b) Faith's phylogenetic diversity, (c) Shannon entropy, (d) Pielou's evenness – as calculated from 16S rRNA gene amplicon sequencing data in QIIME 2™. Asterisks indicate significant differences as determined by adjusted p-values < 0.05 in Dunn's tests.

restarting RBF led to reduced well water alpha diversity in the first seven days, but a subsequent return of diversity to near-original levels after forty-nine days (Fillinger et al., 2021). Thus, assuming many years of near-constant surface water intrusion at this well, the subsurface conditions in the vicinity of the well may have long ago reached the same kind of equilibrium state observed forty-nine days after restarting RBF in the study mentioned above. This would explain why, here, prokaryotic diversity in the suspect well was more similar to that of the comparison well than to that of the gravel pit pond. Suspect well B is also situated ~180 m from the gravel pit pond, whereas the wells that saw the most obvious drops in alpha diversity in the RBF study were located ~20 m and ~70 m from the river. Greater distance between surface water and well may help to somewhat dampen the impact of the intrusion on the prokaryotic communities near the well, through the natural attenuation process associated with subsurface flow (Hiscock and Grischek, 2002). No significant differences in prokaryotic alpha diversity were observed between sampling/measurement sites in the other two study areas.

3.5. Clear separation between prokaryotic communities from suspect well B and the nearby pond

Despite strong indications of surface water intrusion at suspect well B, prokaryotic communities from the suspect well and gravel pit pond in Area B were clearly separated in nMDS plots (Fig. 7). Communities from the suspect well grouped more closely to communities from the comparison well, suggesting similarity between the prokaryotic communities from both wells, despite suspect well communities being strongly exposed to intrusion, and comparison well communities being free of this influence. The similarity between suspect well and comparison well was in agreement with alpha diversity findings.

In all three study areas, ellipses for the wells overlapped more with each other than with surface water ellipses, indicating varying degrees of similarity between the suspect well and comparison well communities. Communities from surface water were most dissimilar to well water communities in Areas B and C, whereas the separation between well water communities and surface water communities was less clear in Area A. Pairwise PERMANOVA of sampling/measurement sites within each study area revealed significant differences for all pairs in all sites ($p < 0.05$), except for the suspect well and comparison well in Area A.

The environmental variables δ^{18}O, d-excess, lc-excess*, SO_4^{2-}, PO_4^{3-}, and Fe were significantly correlated with the ordination axes in all three study areas ($p < 0.05$). The variables turbidity, TN, Ca, Mg, NO_3^-+NO_2^--N, and NH_4^+-N were significantly correlated ($p < 0.05$) with the ordination axes in two study areas each.

3.6. Site-specific differences and temporal variation in dominant prokaryotic groups

There was a marked increase in the relative abundance of the *Crenothrix* genus of methanotrophic bacteria in suspect well C around the time of the springtime snowmelt (Fig. S5), which may be related to the transport of methane or *Crenothrix* bacteria into the well from the nearby lake, though this was not investigated further (Bastviken et al., 2004; Stoecker et al., 2006). Candidatus Moranbacteria, a group belonging to the Candidate Phylum Radiation (CPR) (Brown et al., 2015; Wrighton et al., 2012), also had high relative abundances in this well at several timepoints.

High relative abundances of the *Gallionella* genus of iron-oxidizing bacteria were detected in samples from comparison well B throughout the entire study period; this well also had the highest Fe of all wells. Dominant groups in the surface waters included *Polynucleobacter* (most abundant in the river), the hgcl clade (most abundant in the gravel pit pond and lake), and several others (Fig. S5).

The overall shortage of successfully sequenced samples from suspect wells A and B hindered detection of high-level changes to the prokaryotic communities arising from the impacts of surface water intrusion at

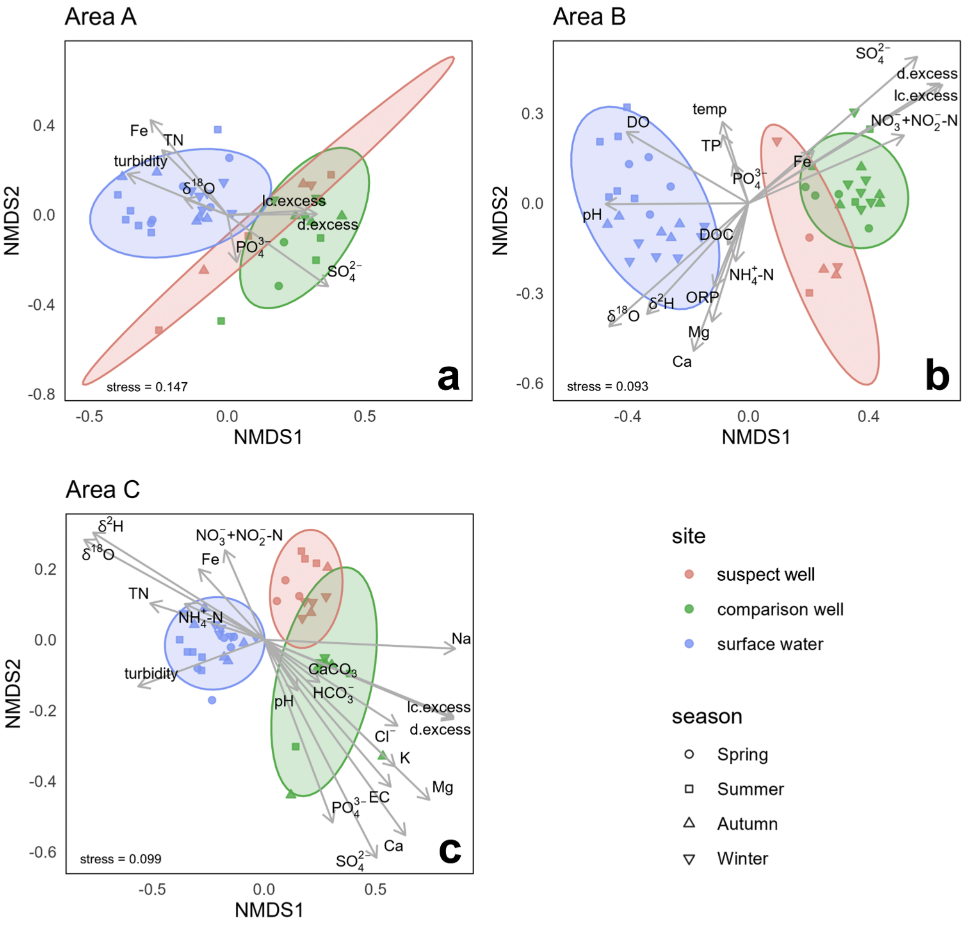

Fig. 7. Non-metric multidimensional scaling (nMDS) plots of Bray–Curtis dissimilarity, for study areas A, B, and C (a–c), as calculated from 16S rRNA gene amplicon sequencing data via the metaMDS() function of the 'vegan' R package. Water quality variables were fit to plots using the envfit() function of the same package. Only variables significantly correlated with the ordination axes ($p < 0.05$) are shown. Seasons are winter (January–March), spring (April–June), summer (July–September) and autumn (October–December). The ellipses in the three plots enclose approximately 95 % of the data points in each group.

these sites, although the communities of suspect well B might be said to appear rather stable at five of the six successfully sequenced timepoints (Fig. S6). Sequencing was more successful for suspect well C. Many of the classes detected at high relative abundances in the nearby lake (e.g. Acidimicrobiia, Actinobacteria, Alphaproteobacteria, Bacteroidia, and Verrucomicrobiae) were not detected at high relative abundances in suspect well C, suggesting that direct influx of surface water microbes is not a strong influence on the groundwater communities at this well, a finding supported by an earlier RBF study (Fillinger et al., 2021).

4. Conclusions

Our work revealed indications of surface water intrusion in shallow wells on the basis of physicochemical and isotopic data. Water from the well with the strongest indications of intrusion exhibited strong physicochemical and isotopic similarity with water from the nearby gravel pit pond, yet higher prokaryotic diversity and a distinct community composition. This finding makes clear the need for a multi-method approach when studying the impact of surface water intrusion on shallow wells.

Nearby surface water is often considered a threat to well water quality. In some cases, however, it may provide unexpected benefits – acting as a reservoir of stored water and supplying this water to wells via bank filtration without noticeably deteriorating well water quality (as in Area B here). Regardless of whether intrusion has a positive or negative impact on well water quality, identifying the source of the intrusion is an important part of water safety planning, enabling well managers to better understand the nature of the impact and to take appropriate action.

Although snapshot sampling of well water isotopes is sometimes sufficient to detect surface water intrusion, our work suggests that a proper understanding of the groundwater–surface water interaction requires that investigators apply a variety of methods to both wells and nearby surface water bodies over several sampling/measurement timepoints. Special attention should perhaps also be given to time periods where the rate of intrusion is expected to be elevated, with sampling being conducted more frequently during those periods to observe in greater detail the changes arising from the interaction. Combining the methods used here with a more detailed account of aquifer characteristics and subsurface flow patterns would likely also enable a better overall understanding of the spatial and temporal variation of the interaction and its consequences for water supply management.

CRediT authorship contribution statement

Kevin J. Lyons: Writing – review & editing, Writing – original draft, Visualization, Investigation, Funding acquisition, Formal analysis, Data curation. **Vadim Yapiyev:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Data curation. **Kaisa Lehosmaa:** Writing – review & editing. **Anna-Kaisa Ronkanen:** Writing – review & editing. **Pekka M. Rossi:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Katharina Kujala:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Kevin J. Lyons reports financial support was provided by The Land and Water Technology Support Association (Maa- ja vesitekniikan tuki ry). Kevin J. Lyons reports financial support was provided by The University of Oulu Scholarship Foundation (Oulun yliopiston tukisäätiö). Kevin J. Lyons reports financial support was provided by The Sven Hallin Research Foundation sr (Sven Hallinin tutkimussäätiö sr). Pekka M. Rossi reports financial support was provided by The Finnish Water Utilities Association (Vesilaitosyhdistys). Vadim Yapiyev reports financial support was provided by Nazarbayev University. Katharina Kujala reports financial support was provided by The Research Council of Finland (Suomen Akatemia). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to acknowledge the contributions of Joni Koivula, Laura Tarvainen, Pietari Pöykkö, Lisa Divet, Tuomo Pitkänen, and Grethel Ruiz (general assistance with field- and laboratory work); Markus Väyrynen, Ilkka Vesavaara, and Juhani Väisänen (IC and ICP analysis); Päivi Salmijärvi (nutrient measurements); and Kashif Noor (stable water isotope measurements). The authors also wish to acknowledge the water utilities and well operators for their co-operation and participation in the study, and the CSC: IT Centre for Science (CSC: tietotekniikan keskus) for providing computational resources. K.J.L. was supported by personal grants from the Land and Water Technology Support Association (Maa- ja vesitekniikan tuki ry) (project ID: 4261 Vesihuolto), the University of Oulu Scholarship Foundation (Oulun yliopiston tukisäätiö) and the Sven Hallin Research Foundation sr (Sven Hallinin tutkimussäätiö sr). P.M.R. received funding from the development fund of the Finnish Water Utilities Association (Vesilaitosyhdistys). V.Y. was supported by Nazarbayev University (CRP research grant 021220CRP2122) during the preparation of the manuscript. K.K. was supported by the Research Council of Finland (Suomen Akatemia) (grant ID: 322753).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2024.122812.

Data availability

Data will be made available on request.

References

Anderson, M.P., 2005. Heat as a ground water tracer. Ground Water 43, 951–968. https://doi.org/10.1111/j.1745-6584.2005.00052.x.

Ascott, M.J., Lapworth, D.J., Gooddy, D.C., Sage, R.C., Karapanos, I., 2016. Impacts of extreme flooding on riverbank filtration water quality. Sci. Total Environ. 554–555, 89–101. https://doi.org/10.1016/j.scitotenv.2016.02.169.

Bastviken, D., Cole, J., Pace, M., Tranvik, L., 2004. Methane emissions from lakes: dependence of lake characteristics, two regional assessments, and a global estimate. Glob. Biogeochem. Cycles 18, GB4009. https://doi.org/10.1029/2004GB002238.

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.

Brown, C.T., Hug, L.A., Thomas, B.C., Sharon, I., Castelle, C.J., Singh, A., Wilkins, M.J., Wrighton, K.C., Williams, K.H., Banfield, J.F., 2015. Unusual biology across a group comprising more than 15% of domain Bacteria. Nature 523, 208–211. https://doi.org/10.1038/nature14486.

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.

Chaudhary, K., Scanlon, B., Scheffe, N., Walden, S., 2009. Review of the state of art: ground water under the direct influence of surface water programs. https://www.beg.utexas.edu/files/content/beg/research/swr/GWUDI-Final-Report.pdf..

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.

Chin, D.A., Qi, X., 2000. Ground water under direct influence of surface water. J. Environ. Eng. 126, 501–508. https://doi.org/10.1061/(ASCE)0733-9372(2000)126:6(501).

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.

Craig, H., 1961. Isotopic variations in meteoric waters. ScienceScience 133, 1702–1703. https://doi.org/10.1126/science.133.3465.1702.

Cullimore, D.R., 2008. Practical Manual of Groundwater Microbiology. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA. https://doi.org/10.1201/9781420008166.

Dinno, A., 2022. Package 'dunn.test'. R package. https://CRAN.R-project.org/package=dunn.test.

Edmunds, W.M., Shand, P., 2008. Natural groundwater quality. Blackwell Publishing Ltd., Malden, MA, USA; Oxford, UK.

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

Fillinger, L., Hug, K., Griebler, C., 2021. Aquifer recharge viewed through the lens of microbial community ecology: Initial disturbance response, and impacts of species sorting versus mass effects on microbial community assembly in groundwater during riverbank filtration. Water Res. 189, 116631. https://doi.org/10.1016/j.watres.2020.116631.

Fitts, C.R., 2013. Groundwater Science, Second Edition. Elsevier Inc. https://doi.org/10.1016/B978-0-12-384705-8.00001-7.

Foulquier, A., Malard, F., Mermillod, F., Montuelle, B., Doledec, S., Volat, B., Gibert, J., 2011. Surface water linkages regulate trophic interactions in a groundwater food web. Ecosystems 14, 1339–1353.

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.

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.

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.

Han, Z., Shi, X., Jia, K., Sun, B., Zhao, S., Fu, C., 2019. Determining the discharge and recharge relationships between lake and groundwater in Lake Hulun using hydrogen and oxygen isotopes and chloride ions. Water 11, 264. https://doi.org/10.3390/w11020264.

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.

Hiscock, K.M., Grischek, T., 2002. Attenuation of groundwater pollution by bank filtration. J. Hydrol. 266, 139–144. https://doi.org/10.1016/S0022-1694(02)00158-0.

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.

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.

Ji, L., Wang, Z., Zhang, L., Zhu, X., Ning, K., 2022a. Determining the primary sources of groundwater bacterial communities in a large-scale plain area: Microbial source tracking and interpretation for different land use patterns. Agric. Ecosyst. Environ. 338, 108092. https://doi.org/10.1016/j.agee.2022.108092.

Ji, L., Zhang, L., Wang, Z., Zhu, X., Ning, K., 2022b. High biodiversity and distinct assembly patterns of microbial communities in groundwater compared with surface water. Sci. Total Environ. 834, 155345. https://doi.org/10.1016/j.scitotenv.2022.155345.

Josse, J., Husson, F., 2016. missMDA: a package for handling missing values in multivariate data analysis. J. Stat. Softw. 70. https://doi.org/10.18637/jss.v070.i01.

Kalbus, E., Reinstorf, F., Schirmer, M., 2006. Measuring methods for groundwater – surface water interactions: a review. Hydrol. Earth Syst. Sci. 10, 873–887. https://doi.org/10.5194/hess-10-873-2006.

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.

Kortelainen, N., 2011. Isotope tracing in groundwater applications. Geol. Sur. Finland Special Paper 49, 279–284. https://tupa.gtk.fi/julkaisu/specialpaper/sp_049_pages_279_284.pdf.

Kortelainen, N., 2007. Isotopic fingerprints in surficial waters stable isotope methods applied in hydrogeological studies. Geolo. Sur. Finland Espoo. http://hdl.handle.net/10138/21197.

Landwehr, J.M., Coplen, T.B., 2004. Line-conditioned excess: a new method for characterizing stable hydrogen and oxygen isotope ratios in hydrologic systems. Presented at the International Conference on Isotopes in Environmental Studies, Aquatic Forum 2004, 25–29 October 2004, Monte-Carlo, Monaco, pp. 98–99. https://inis.iaea.org/collection/NCLCollectionStore/_Public/36/003/36003223.pdf?r=1.

Lê, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. J. Stat. Softw. 25. https://doi.org/10.18637/jss.v025.i01.

Li, P., Karunanidhi, D., Subramani, T., Srinivasamoorthy, K., 2021. Sources and consequences of groundwater contamination. Arch. Environ. Contam. Toxicol. 80, 1–10. https://doi.org/10.1007/s00244-020-00805-z.

Liu, Y., Fang, Y., Hu, H., Tian, F., Dong, Z., Khan, M.Y.A., 2020. Ecohydrological separation hypothesis: review and prospect. Water 12, 2077. https://doi.org/10.3390/w12082077.

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. 9. https://doi.org/10.1128/Spectrum.00179-21 e00179-21.

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. Sci. Total Environ. 864, 161199. https://doi.org/10.1016/j.scitotenv.2022.161199.

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.

Maliva, R., Missimer, T., 2012. Riverbank Filtration, in: Arid Lands Water Evaluation and Management, Environmental Science and Engineering. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 631–645. https://doi.org/10.1007/978-3-642-29104-3_24.

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.

Martinez Arbizu, P., 2020. pairwiseAdonis: Pairwise multilevel comparison using adonis. R package. https://github.com/pmartinezarbizu/pairwiseAdonis.

Mortatti, J., Krug, F.J., Pessenda, L.C.R., Zagatto, E.A.G., Jørgensen, S.S., 1982. Determination of iron in natural waters and plant material with 1,10-phenanthroline by flow injection analysis. Analyst 107, 659–663. https://doi.org/10.1039/AN9820700659.

Nnadi, F.N., Fulkerson, M., 2002. Assessment of groundwater under direct influence of surface water. J. Environ. Sci. Health, Part A A37, 1209–1222. https://doi.org/10.1081/ESE-120005981.

Noor, K., Marttila, H., Welker, J.M., Mustonen, K.-R., Kløve, B., Ala-aho, P., 2023. Snow sampling strategy can bias estimation of meltwater fractions in isotope hydrograph separation. J. Hydrol. 627, 130429. https://doi.org/10.1016/j.jhydrol.2023.130429.

Nora, J., Rossi, P., Sanaksenaho, R., Lindholm, A., 2019. POSKI2 project, Lapland – separate study on water isotopes (Lapin POSKI2 – hankkeen erillisselvitys: Isotooppitutkimukset). [in Finnish] Geological Survey of Finland. Groundwater Unit, Rovaniemi. Finland. http://projects.gtk.fi/export/sites/projects/poski_lappi/aineistot/LapinPOSKI2_selvitys_isotooppitutkimukset.pdf.

Oksanen, J., 2020. vegan: Community ecology package. R package. https://cran.r-project.org/package=vegan.

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.

R Core Team, 2024. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Ray, C., Melin, G., Linsky, R.B, 2003. Riverbank filtration: improving source-water quality. Water Science and Technology Library. Kluwer Academic Publishers, Dordrecht, The Netherlands. https://doi.org/10.1007/0-306-48154-5.

Ritter, L., Solomon, K., Sibley, P., Hall, K., Keen, P., Mattu, G., Linton, B., 2002. Sources, pathways, and relative risks of contaminants in surface water and groundwater: a perspective prepared for the Walkerton Inquiry. J. Toxicol. Environ. Health, Part A, 65, 1–142. https://doi.org/10.1080/152873902753338572.

Robeson, M.S., O'Rourke, D.R., Kaehler, B.D., Ziemski, M., Dillon, M.R., Foster, J.T., Bokulich, N.A., 2021. RESCRIPt: Reproducible sequence taxonomy reference database management. PLOS Comput. Biol. 17, e1009581. https://doi.org/10.1371/journal.pcbi.1009581.

Rosenberry, D.O., Lewandowski, J., Meinikmann, K., Nützmann, G., 2015. Groundwater - the disregarded component in lake water and nutrient budgets. Part 1: effects of groundwater on hydrology. Hydrol. Process. 29, 2895–2921. https://doi.org/10.1002/hyp.10403.

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. Worldwatch Institute, Washington, DC (No. Worldwatch paper 154). https://archive.org/details/deeptrouble00paya.

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.

Stoecker, K., Bendinger, B., Schöning, B., Nielsen, P.H., Nielsen, J.L., Baranyi, C., Toenshoff, E.R., Daims, H., Wagner, M., 2006. Cohn's *Crenothrix* is a filamentous methane oxidizer with an unusual methane monooxygenase. Proc. Natl. Acad. Sci. 103, 2363–2367. https://doi.org/10.1073/pnas.0506361103.

Vasconcelos, J., Harris, S., 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. https://nepis.epa.gov/Exe/ZyPDF.cgi/P100C58D.PDF?Dockey=P100C58D.PDF.

Wett, B., Jarosch, H., Ingerle, K., 2002. Flood induced infiltration affecting a bank filtrate well at the River Enns. Austria. J. Hydrol. 266, 222–234. https://doi.org/10.1016/S0022-1694(02)00167-1.

WHO, 2011. Guidelines for Drinking-Water Quality, 4th edn. World Health Organization, Geneva.

Wickham, H., 2016. ggplot2: elegant graphics for data analysis. Springer, Switzerland. https://doi.org/10.1007/978-3-319-24277-4. Second edition. ed, Use R!

Winter, T.C., Harvey, J.W., Franke, O.L., Alley, W.M., 1998. Ground water and surface water: a single resource, U.S. Geological Survey Circular 1139. https://doi.org/10.3133/cir1139.

Wrighton, K.C., Thomas, B.C., Sharon, I., Miller, C.S., Castelle, C.J., VerBerkmoes, N.C., Wilkins, M.J., Hettich, R.L., Lipton, M.S., Williams, K.H., Long, P.E., Banfield, J.F., 2012. Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. Science 337, 1661–1665. https://doi.org/10.1126/science.1224041.

Yang, F., Yue, S., Wu, X., Zhang, C., Li, D., Zhu, R., 2023. Effects of flood inundation on biogeochemical processes in groundwater during riverbank filtration. J. Hydrol. 617, 129101. https://doi.org/10.1016/j.jhydrol.2023.129101.

Yapiyev, V., Rossi, P.M., Ala-Aho, P., Marttila, H., 2023. Stable water isotopes as an indicator of surface water intrusion in shallow aquifer wells: a cold climate perspective. Water Resour. Res. 59. https://doi.org/10.1029/2022WR033056 e2022WR033056.

Zektser, I.S., Everett, L.G., 2004. Groundwater Resources of the World and Their Use. UNESCO, Paris. https://unesdoc.unesco.org/ark:/48223/pf0000134433.