OULU 2024
C 963

ACTA UNIVERSITATIS OULUENSIS

C
TECHNICA

Kevin Lyons

POTABLE WATER FROM SHALLOW WELLS

A MULTI-METHOD STUDY OF FACTORS INFLUENCING PHYSICOCHEMICAL AND MICROBIOLOGICAL QUALITY

UNIVERSITY OF OULU GRADUATE SCHOOL;
UNIVERSITY OF OULU,
FACULTY OF TECHNOLOGY

ACTA UNIVERSITATIS OULUENSIS
C Technica 963

KEVIN LYONS

POTABLE WATER FROM SHALLOW WELLS

A multi-method study of factors influencing physicochemical and microbiological quality

Academic dissertation to be presented with the assent of the Doctoral Programme Committee of Technology and Natural Sciences of the University of Oulu for public defence in Martti Ahtisaari auditorium (L2), Linnanmaa, on 3 October 2024, at 12 noon

UNIVERSITY OF OULU, OULU 2024

Acta Univ. Oul. C 963, 2024

Supervised by
Docent Katharina Kujala
Assistant Professor Pekka Rossi
Docent Anna-Kaisa Ronkanen

Reviewed by
Professor Christian Griebler
Professor Paul Hynds

Opponent
Associate Professor Martin Taubert

ISBN 978-952-62-4210-1 (Paperback)
ISBN 978-952-62-4211-8 (PDF)

ISSN 0355-3213 (Printed)
ISSN 1796-2226 (Online)

Cover Design
Raimo Ahonen

PUNAMUSTA
TAMPERE 2024

Lyons, Kevin, Potable water from shallow wells. A multi-method study of factors influencing physicochemical and microbiological quality
University of Oulu Graduate School; University of Oulu, Faculty of Technology
Acta Univ. Oul. C 963, 2024
University of Oulu, P.O. Box 8000, FI-90014 University of Oulu, Finland

Abstract

While shallow wells (e.g. well depth < 15 m) provide easy and affordable access to water, they can also be vulnerable to influences from the surrounding environment due to the thinness of the overlying soil layer and the presence of groundwater–surface water interactions.

Here, water from thirty-three shallow wells used for potable water supply in northern and central Finland was assessed to better understand the nature of these influences and their effects on well water quality.

Well water was assessed by (i) manual and real-time online measurements of physicochemical parameters, (ii) counts of microbiological indicators, (iii) determination of stable water isotope compositions, and (iv) microbial community analyses. Several of these methods were also applied to water from nearby surface water bodies to investigate groundwater–surface water interactions.

Physicochemical and microbiological water quality data from several wells appeared to show evidence of surrounding influences (e.g. agriculture, faecal contamination, road salt, and surface water intrusion) – and real-time online monitoring demonstrated potential for more widespread use. Snapshot sampling of stable water isotope compositions was itself sufficient to reveal indications of surface water intrusion at several wells, whereas more frequent sampling revealed temporal variation in the groundwater–surface water interactions. One well near a gravel pit pond was of particular interest: (i) isotope mixing models estimated that 80–95% of the well water comes from the pond, (ii) well water exhibited strong overall physicochemical similarity to the pond, and (iii) well water temperature and isotope compositions showed temporal fluctuations that strongly mirrored those of the pond. 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, suggesting that resident well communities are not strongly impacted by influx of surface water taxa.

Well water monitoring often relies on periodic on-site measurements, as well as manual sampling and laboratory analyses. This thesis shows how that approach can be combined with alternative methods to provide a more holistic understanding of the influences affecting shallow wells, producing findings with relevance to public health, water supply management, and subsurface microbial ecology.

Keywords: drinking water, groundwater, microbial ecology, microbiology, monitoring, stable water isotopes, water quality, wells

Lyons, Kevin, Vesihuoltolaitosten kaivoveden monitorointi. Monimenetelmällinen lähestymistapa kemiallisen ja mikrobiologisen laatuvaihtelun syiden tarkasteluun

Oulun yliopiston tutkijakoulu; Oulun yliopisto, teknillinen tiedekunta
Acta Univ. Oul. C 963, 2024
Oulun yliopisto, PL 8000, 90014 Oulun yliopisto

Tiivistelmä

Matalien pohjavesiesiintymien kaivot (kaivon syvyys < 15 m) mahdollistavat verrattain helpon vedenoton ja ovat yleisin vedenoton muoto Suomessa. Toisaalta nämä kaivot voivat olla haavoittuvia ohuen maaperäkerroksen ja mahdollisten pohjavesi-pintavesi-interaktioiden takia. Tässä tutkimuksessa tarkasteltiin kolmeakymmentäkolmea vedenottamon kaivoa Pohjois- ja Keski-Suomessa, jotta haavoittuvuusriskejä ja niiden vaikutuksia vedenlaatuun voitaisiin ymmärtää paremmin.

Kaivovesiä tarkasteltiin (i) manuaalisilla ja jatkuvatoimisilla fysikaaliskemiallisten parametrien mittauksilla, (ii) mikrobiologisten indikaattorien laskennalla, (iii) veden stabiilien isotooppien koostumuksilla ja (iv) mikrobiyhteisöjen analyyseilla. Useita näistä menetelmistä käytettiin myös läheisten pintavesien analysointiin pohjavesi-pintavesi-interaktion tarkasteluun.

Fysikaaliskemialliset ja mikrobiologiset vedenlaatumittaukset osoittivat usean kaivon osalta alueen ympäröivien tekijöiden vaikutusta (kuten maanviljely, fekaalinen kontaminaatio, tien suolaus ja pintaveden intruusio). Jatkuvatoimiset mittaukset osoittivat mittauksissa käytettävyyttään: menetelmille olisi laajamittaista käyttöpotentiaalia. Veden stabiilien isotooppien koostumuksen analysointi kertanäytteellä paljasti usealla kohteella pintaveden intruusiovaikutusta kaivovedessä. Pidempiaikainen isotooppien seuranta puolestaan osoitti interaktion ajallista vaihtelevuutta. Yksi tutkituista kaivoista, joka sijaitsi lähellä vedellä täyttynyttä sorakuoppaa, osoittautui mielenkiintoiseksi: (i) isotooppeihin perustuva komponenttianalyysi arvioi pintaveden osuuden kaivovedestä olevan 80–95 %, (ii) fysikaaliskemialliset parametrit osoittivat vahvaa samankaltaisuutta veden ja sorakuopan veden välillä ja (iii) lämpötilan ja isotooppien koostumuksen vaihtelu heijasteli samoja ajallisia vaihteluja kuin mitä mitattiin sorakuopan vedestä. Sorakuopan veteen verrattuna prokaryoottiyhteisöjen alfadiversiteetti kaivossa oli kuitenkin huomattavasti korkeampi ja yhteisöjen koostumus lähempänä läheistä vertailukaivoa, jossa intruusiota ei ollut havaittu. Tämä viittaa siihen, että kaivoon virtaavan pintaveden yhteisöt eivät vaikuta vahvasti kaivon mikrobiyhteisöön.

Kaivojen monitorointi nojaa monessa tapauksessa ajoitettuihin näytteenottoihin manuaali- sekä laboratorioanalyyseillä. Tämä tutkimus osoitti, että tätä perinteistä monitorointia voidaan yhdistää vaihtoehtoisiin menetelmiin, jotka tuottavat kokonaisvaltaisemman ymmärryksen kaivoihin vaikuttavista tekijöistä. Tällä ymmärryksellä on merkityksensä niin kansanterveyden, vesien hallinnoinnin kuin maaperän mikrobiologian osalta.

Asiasanat: isotoopit, juomavesi, kaivot, mikrobiologia, pohjavesi, valvonta, vedenlaatu, ympäristömikrobiologia

To my supervisors (for their patience)

Acknowledgements

The work described in this thesis would not have been possible without the help of my doctoral supervisors (Katharina Kujala, Pekka Rossi, and Anna-Kaisa Ronkanen), our collaborators (Anna-Maria Hokajärvi, Jenni Ikonen, Jussi Jyväsjärvi, Ari Kauppinen, Kaisa Lehosmaa, Ilkka Miettinen, Tarja Pitkänen, Timo Räsänen, Vadim Yapiyev), and all others who assisted with field and laboratory work throughout the project (Ilkka Dahlström, Lisa Divet, Elisangela Heiderscheidt, Tiina Heiskanen, Jussi Keränen, Joni Koivula, Ville Komulainen, Paavo Kuusela, Kari Moisio, Kashif Noor, Pekka Pesälä, Tuomo Pitkänen, Pietari Pöykkö, Tarja Rahkonen, Tuomo Reinikka, Grethel Ruiz, Päivi Salmijärvi, Laura Tarvainen, Taina Uusitalo, Ilkka Vesavaara, Juhani Väisänen, Markus Väyrynen). Thanks also to my follow-up group (Pertti Ala-Aho, Jenni Hultman, and Kaisa Lehosmaa) for following my progress and giving useful advice along the way.

I would like to thank MVTT (Maa ja vesitekniikan tuki ry) and the KAUTE Foundation (KAUTE-säätiö) for supporting the work through personal grants awarded specifically to me; as well as all the funders that supported my co-authors and the various sub-projects upon which this thesis is based: MVTT, the Finnish Water Utilities Association (Vesilaitosyhdistys), the Finnish Ministry of Agriculture and Forestry (Maa- ja metsätalousministeriö), and the Research Council of Finland (Suomen Akatemia). Thanks also to the CSC: IT Centre for Science (CSC: tietotekniikan keskus) for providing computational resources. I would like to thank the University of Oulu Graduate School (UniOGS) for awarding me a working grant to write the first draft of this thesis, and awarding travel grants supporting my attendance at two major international conferences during my time as a doctoral researcher: namely, the IWA World Water Congress & Exhibition in September 2022 (held in Copenhagen, Denmark) and the FEMS2023 Congress of European Microbiologists in July 2023 (held in Hamburg, Germany). Thanks also to the University of Oulu Scholarship Foundation (Oulun yliopiston tukisäätiö) and the Sven Hallin Research Foundation sr (Sven Hallinin tutkimussäätiö sr) for awarding personal grants that supported me through the final stages of writing and revising.

Last but not least, I would like to thank my wife, Anniina, for putting up with me as I wrestled and waded through this lengthy process. *Grá mo chroí thú!*

Oulu, July 12th 2024 Kevin Lyons

List of abbreviations and symbols

ANOVA	analysis of variance
ASV	amplicon sequence variant
Ca	calcium
Cl^-	chloride
d-excess	deuterium excess
DNA	deoxyribonucleic acid
DO	dissolved oxygen
DOC	dissolved organic carbon
EC	electrical conductivity
Fe	iron
FMI	Finnish Meteorological Institute
K	potassium
lc-excess	line-conditioned excess
LEL	local evaporation line
LMWL	local meteoric water line
Mg	magnesium
Mn	manganese
MST	microbial source tracking
Na	sodium
NH_4^+-N	ammonium nitrogen
NLS	National Land Survey of Finland
nMDS	non-metric multidimensional scaling
NO_3^-	nitrate
$(NO_3^- + NO_2^-)$-N	combined nitrate and nitrite nitrogen
ORP	redox potential
P	phosphorus
PCA	principal component analysis
PCoA	principal coordinate analysis
PCR	polymerase chain reaction
PERMANOVA	permutational multivariate analysis of variance
PO_4^{3-}	phosphate
rDNA	ribosomal deoxyribonucleic acid
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
SiO_2	silica

SO_4^{2-}	sulphate
SYKE	Finnish Environment Institute
TC	total carbon
N or TN or N_{tot}	total nitrogen
P or TP or P_{tot}	total phosphorus
qPCR	quantitative real-time polymerase chain reaction
WHO	World Health Organization

List of original publications

This thesis is based on the following publications, which are referred to throughout the text as ‘the first study’, ‘the second study’ and ‘the third study’, or by their Roman numerals:

I 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. *Microbiology Spectrum*, *9*(3), e00179-21. https://doi.org/10.1128/Spectrum.00179-21

II 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*, 864, 161199. https://doi.org/10.1016/j.scitotenv.2022.161199

III Lyons, K. J., Yapiyev, V., Lehosmaa, K., Ronkanen, A-K., Rossi, P. M. & Kujala, K. (Manuscript). Assessing the influence of surface water intrusion on shallow groundwater wells using water quality measurements, stable water isotopes, and microbial community analysis.

The author’s contribution to the publications:

I Processed and analysed the data with advice and input from doctoral supervisors. Took primary responsibility for writing the manuscript text and preparing the figures and tables for publication with input from the co-authors.

II Collected some of the samples from the field. Processed and analysed the data with advice and input from doctoral supervisors. Took primary responsibility for developing and finalising the manuscript text, figures, and tables based on an early draft by the co-first author, with input from the other co-authors.

III Conducted most of the field work and a large part of the laboratory work. Processed and analysed the data with advice and input from doctoral supervisors. Took primary responsibility for writing the manuscript text and preparing the figures and tables for publication with input from the co-authors.

Contents

1 Introduction

1.1 Groundwater and groundwater wells

Groundwater – that is, water held within the cracks and pores of fully saturated soil, sand, gravel, or rock – is estimated to provide domestic water to about half of the world's population (Margat & van der Gun, 2013). Hence, safe supplies of groundwater are crucial for maintaining good public health and a well-functioning society (Figueras & Borrego, 2010; Macler & 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 & Kendy, 2005; Wada et al., 2010) and water quality degradation (Burri et al., 2019). Thus, research is needed to better understand these threats and to make recommendations for the better management of groundwater resources.

Groundwater can be extracted and investigated via wells, which are dug, driven, or drilled into the ground. Small-diameter, non-pumping wells are typically used for monitoring purposes, whereas pumping wells used for potable water supply tend to vary in scale and design (Cushman & Tartakovsky, 2017; Misstear et al., 2017). Studying wells can bring insights into groundwater quality, quantity, and ecology, and how these are affected by weather events and climate change (Burri et al., 2019; Griebler et al., 2014; Taylor et al., 2013; Treidel et al., 2011; Wada et al., 2010). All of the wells studied in this work are pumping wells used for potable water supply in Finland, where groundwater depletion is not currently a major problem (Lavapuro et al., 2008). Here, the safeguarding of human health mostly involves understanding and minimising chemical and microbiological risks to well water quality and the negative health consequences thereof (Isomäki et al., 2006, 2008; Lahermo, 2002; Pitkänen et al., 2011, 2015; Wallin, 2016).

1.2 Geology and hydrogeology of Finland

The geology of Finland can perhaps be summarised as "the very old meeting the very new". The Precambrian crystalline basement underlying Finland is part of the Fennoscandian Shield, which also covers much of southern Norway, most of Sweden, as well as regions of Russia directly to the east of Finland (Donner, 1996). In Finland, this (largely granitic) basement consists of Archean, Svecokarelian, and

Svecofinnian rocks formed between 3.5 and 1.75 billion years ago (Kitterød et al., 2022). Atop this ancient basement lies a variety of superficial elements, which have much newer origins; namely, glacial tills (46.5% of total area), peats (14.9%), surface waters (10%), sands and gravels (7.5%), clays (5%), and silts (3%); with exposed bedrock making up the final 13.6% (Kristensen & Bøgestrand, 1996; Luoma & Ikonen, 2020).

Of most relevance to this thesis are the Quaternary deposits of sand and gravel, formed during the deglaciation of the Fennoscandian Ice Sheet about 12000–9000 years ago (Katko et al., 2006; Luoma & Ikonen, 2020). These coarse-grained glaciofluvial deposits (eskers and ice-marginal formations) have high exploitation potential for groundwater abstraction and are the most economically significant sources of groundwater in Finland (Karro & Lahermo, 1999; Katko et al., 2006; Lavapuro et al., 2008). These glaciofluvial aquifers are mostly long, narrow, and small (only a few km^3 in area) (Katko et al., 2006). They are also typically unconfined (covered only by a thin soil layer) and shallow (having a water table at most 10–50 m below the surface), making them vulnerable to contamination (Isomäki et al., 2008; Karro & Lahermo, 1999; Luoma & Ikonen, 2020; Väisänen, 2004).

Based on data from the period 1971–2020, the average depth to groundwater in Finland has been estimated at 4.5 m (Luoma & Ikonen, 2020). Given the generally thin overlying soil layer, the groundwater table tends to respond quite rapidly to snowmelt and precipitation inputs (often within days), giving rise to characteristic seasonal fluctuations (Soveri, 1985). The exact pattern of fluctuations varies depending on location. However, in much of the country, the rather large-scale input of snowmelt water in springtime is the most important source of groundwater recharge and leads to an annual maximum in groundwater levels; levels then tend to fall gradually throughout the summer, before rising again in response to autumn precipitation (R. Mäkinen et al., 2008; Okkonen et al., 2010).

1.3 Use of shallow wells for potable water supply in Finland

1.3.1 Prevalence and management

In Finland, as in many other parts of the world, groundwater is often favoured for potable water supply over surface water, because it tends to be of more stable temperature and quality, to have greater protection from contamination, and to therefore not require the same degree of treatment (Katko et al., 2006; Katsanou & Karapanagioti, 2017; Ratnayaka et al., 2009). Groundwater in Finland is typically supplied to users via (i) wells managed by a municipality or a municipality-owned company, (ii) wells managed by a local water co-operative (also known as ‘community wells’), or (iii) wells belonging to a private household (Arvonen et al., 2017; Katko et al., 2006). Based on data from the Finnish Ministry of Agriculture and Forestry, naturally occurring groundwater provides about 43% of total potable water supply, and artificial groundwater about 16%, giving a combined total of about 60% (Maa- ja metsätalousministeriö, 2019). However, these figures omit the contribution of private household wells, meaning that the actual combined total supplied by groundwater and artificial groundwater may be as high as 75% (Katko et al., 2006).

Some of the wells investigated for this thesis are managed by a municipality or a municipality-owned company (n = 10). Most, however, are ‘community wells’ managed by a local water co-operative (n = 23) (Isomäki et al., 2006; Katko et al., 2006; Pitkänen et al., 2011). These co-operatives are consumer-owned entities whose goal is not to make a profit, but to produce good-quality water services at affordable rates, with any financial surpluses being fed back into general maintenance and improvement of the services (Pietilä & Vihanta, 2018). As is typical of small community water supplies around the world (Hofkes et al., 2002), small groundwater co-operatives in Finland have historically been managed by volunteers, who do not always have the necessary training to ensure the safe operation of the water supply (Isomäki et al., 2006). In 2008, however, an attempt was made to improve the situation by introducing a requirement for all operators to pass a test of competency (Isomäki et al., 2006, 2008).

1.3.2 Relevant EU and Finnish legislation

As a member state of the European Union, Finland is obligated to observe the requirements and recommendations of EU directives such as the Water Framework Directive (2000/60/EC), the Groundwater Directive (2006/118/EC), and the 2020 recast of the Drinking Water Directive (2020/2184). The latter directive (the DWD) outlines “a complete risk-based approach to water safety, covering the whole supply chain from the catchment area, abstraction, treatment, storage and distribution to the point of compliance”, based on the ‘Guidelines for Drinking Water Quality’ and ‘Water Safety Plan’ approach developed by the World Health Organization (WHO) (WHO, 2011). The DWD applies to all water suppliers providing an average of ≥10 m^3 of water/day or serving ≥50 people. Small suppliers providing an average of <10 m^3 of water/day or serving <50 people may be exempted ("unless the water is supplied as part of a commercial or public activity"). Member states can also choose to exclude (mid-sized) water suppliers providing an average of 10–100 m^3 of water/day or serving 50–500 people, from the risk assessment and risk management requirements ("in order to reduce the potential administrative burden"). According to the DWD, water intended for human consumption should fulfil various requirements to be considered "wholesome and clean". This includes meeting the minimum microbiological and chemical requirements described in Parts A, B, and D of Annex I. (Additional "indicator parameters" are given in Part C of Annex I, largely to be used for monitoring purposes.) These requirements and recommendations are mirrored in two Finnish decrees: (i) 'The Decree of the Ministry of Social Affairs and Health on the quality and control of domestic water and risk management of water installations in buildings' (17.11.2015/1352) (which applies to larger suppliers; namely, those providing ≥10 m^3 of water/day or serving ≥50 people), and (ii) 'The Decree of the Ministry of Social Affairs and Health on the quality requirements and control studies of domestic water for small units' (401/2001) (which applies both to private household wells and also smaller water suppliers; namely, those providing <10 m^3 of water/day or serving <50 people). Most of the wells investigated for this thesis clearly come under the first of these decrees (n = 21), and a few clearly under the second (n = 3), with the remaining wells either matching both definitions (n = 4) or lacking the data necessary for classification (n = 5). (No private household wells were investigated here.) Overall, the water quality requirements and recommendations of both Finnish decrees are rather similar, although some leniency is given to private household wells (e.g. for iron, manganese, and counts of coliform bacteria). Both decrees require more frequent monitoring of well water for parameters such as turbidity, colour, smell, taste, pH, Fe, Mn, *Escherichia coli*, intestinal enterococci, and coliform bacteria, whereas less frequent monitoring is required for various aromatic hydrocarbons, heavy metals, cyanides, pesticides, radon, and many other parameters. For both monitoring types, the monitoring frequency ranges widely and depends on the volume of water distributed by the well and the number of water users, with water from larger wells being more frequently monitored (e.g. some parameters being measured dozens of times per year), and water from the smallest wells being less frequently monitored (e.g. some parameters being measured less than once per year).

1.3.3 Risks and influences

Previous studies of wells in Finland have identified a variety of risks relating to chemical and microbiological water quality, as well as issues relating to well structure and water supply management. Low pH levels, and high concentrations of iron and manganese are common problems (Isomäki et al., 2006; Pitkänen et al., 2015; Wallin, 2016), and road salt used for de-icing in winter can increase chloride concentrations (Gustafsson & Nystén, 2000; Salminen et al., 2011). Nitrogen species may also enter the groundwater from agricultural sources (Burkart & Stoner, 2008; Isomäki et al., 2008). Faecal contamination has previously been detected in some wells – as indicated by the presence, in well water samples, of intestinal enterococci, the bacterium *Escherichia coli*, or the protozoan parasite *Giardia intestinalis* (Pitkänen et al., 2011, 2015). The most common on-site hazards found to increase the risk of microbial contamination at wells in Finland appear to include (i) unwanted intrusion of precipitation and floodwater into the well due to flaws in the well structure, (ii) close proximity of the water table to the land surface (i.e. presence of only a thin protective soil layer above the water table), and (iii) nearby agricultural activity (Pitkänen et al., 2011). Additional risks may include (iv) lack of disinfection facilities, (v) leakage from nearby sewerage systems, (vi) uncontrolled bank filtration of water from a nearby river, lake, or pond, or (vii) insufficient well maintenance and training of personnel (Isomäki et al., 2006, 2008; Pitkänen et al., 2011). Data from the website of the Finnish Institute for Health and Welfare (THL) indicate that contaminated groundwater is the primary cause of waterborne outbreaks in Finland, with about 89% (93 out of 104) of the reported waterborne outbreaks from the period 1998–2022 originating from groundwater, and 54% (56 out of 104) from “private or communal groundwater wells” (THL, 2022). Thus, further work is needed to understand and mitigate risks to groundwater quality in Finnish groundwater wells.

1.4 Surface water intrusion from unintended bank filtration

1.4.1 Intended and unintended bank filtration

Despite appearing to be separate entities, groundwater and surface water bodies (such as lakes, ponds, rivers and streams) are often hydraulically connected (Winter et al., 1998), and this is also true in Finland (Kløve et al., 2017). Thus, the water level and water quality of wells located next to a surface water body can be influenced by fluctuations in surface water level and quality; and similarly, surface water can be influenced by groundwater via a process known as baseflow. This need not necessarily be a problem for potable water supply. In fact, this surface water–groundwater interaction can be exploited as an efficient and cost-effective way of improving the quality of surface water for the purposes of potable water supply; a process known as ‘induced bank filtration’ (Maliva & Missimer, 2012; Ray, 2002; Ray et al., 2002, 2003; Tufenkji et al., 2002). When water is extracted from one or more wells, the local groundwater level drops, which causes increased flow of surface water through the subsurface towards the well(s). As the surface water flows through the subsurface, various abiotic and biotic factors enact natural attenuation processes, by which the concentrations of various microbes and pollutants in the surface water are reduced – thus producing water of better quality than the initial surface water (Hiscock & Grischek, 2002).

The kind of induced bank filtration outlined above is intended and desirable. However, surface water intrusion may also occur in situations where bank filtration is unintended (e.g. a well is established in a sand-and-gravel aquifer near a gravel extraction pit, and the pit is later abandoned, filling with water to form a pond). In such cases, especially, it may be beneficial to investigate the potential influences of surface water intrusion on well water quality, as these influences may not yet have been considered with regard to risk management. Both kinds of bank filtration can be conceptually distinguished from (i) the infiltration of precipitation water through the unsaturated layer to the groundwater table (i.e. the most important source of recharge for most Finnish aquifers), and (ii) the direct ingress of water into a well due to flooding, or due to precipitation or surface runoff entering a structurally comprised well (Andrade et al., 2018).

In 1986, the US Environmental Protection Agency (USEPA) amended the Safe Drinking Water Act (SDWA) such that all states would henceforth be required to identify groundwater bodies strongly influenced by nearby surface water (Nnadi & Fulkerson, 2002). In the US, these groundwater bodies are termed ‘groundwater under the direct influence of surface water’ (variously abbreviated as GUI, GUDI, GWDI, GWUDI, or GWUDISW) and are required to follow the same treatment guidelines as surface water sources (Chin & Qi, 2000). More specifically, the USEPA regulation 40 CFR 141.2 defines GUI as any water beneath the surface of the ground which shows: (i) a significant presence of insects or other macro-organisms, algae, organic debris, or large-diameter pathogens such as *Giardia lamblia* or *Cryptosporidium*; or (ii) significant and rapid shifts in turbidity, temperature, conductivity, or pH which closely correlate to climatological or surface water conditions (Chaudhary et al., 2009; Chin & Qi, 2000). This is a useful definition in the context of this thesis, which examines surface water intrusion from unintended bank filtration as one of the potential factors influencing the quality of well water in Finland.

1.4.2 Impacts of surface water intrusion on groundwater

Groundwater ecosystems are often characterized by stable temperatures, low levels of nutrients and organic carbon, and a lack of light; whereas surface waters are typically characterized by daily and seasonal temperature fluctuations, higher levels of nutrients and organic carbon, and sunlight exposure, which enables the growth of phototrophic organisms (Brunke & Gonser, 1997; Griebler & Lueders, 2009). In addition, surface waters typically have higher prokaryotic cell counts than groundwaters (Alfreider et al., 1997; Brugger et al., 2001; Van Driezum et al., 2018; Vargha et al., 2023; Wang et al., 2022; Zhou et al., 2012), as well as lower prokaryotic diversity (Fillinger et al., 2021; Ji, Wang, et al., 2022; Ji, Zhang, et al., 2022). Thus, intrusion of surface water into the groundwater environment constitutes an ecological disturbance that may bring about simultaneous or sequential abiotic and biotic changes in the aquifer (Fillinger et al., 2021; J.-H. Lee et al., 2018; Lin et al., 2012; Stegen et al., 2016). Abiotic physicochemical changes in the groundwater caused by intruding surface water may include altered stable water isotope ratios ($\delta^{18}O$ and $\delta^{2}H$); changes in temperature, pH, electrical conductivity (EC), redox potential (ORP) and turbidity; or changes in the concentrations of major ions (e.g. Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, Cl^{-}), dissolved oxygen (DO), dissolved organic carbon (DOC), nitrate (NO_3^{-}), sulphate (SO_4^{2-}), iron (Fe) and manganese (Mn) (Bourg & Bertin, 1993; Fillinger et al., 2021; Massmann et al., 2004).

Attempting to distinguish the relative importance of abiotic and biotic effects in shaping groundwater prokaryotic communities can prove challenging, as aquifer biogeochemistry and microbial ecology are often interlinked (Lin et al., 2012). For example, a direct abiotic change (e.g. an increased abundance of nutrients and organic carbon) may lead to secondary biotic changes (e.g. increased heterotrophic microbial metabolism and/or selection of certain microbial groups) (Foulquier, Mermillod-Blondin, et al., 2011), and a direct biotic change (e.g. direct transport of surface water–derived organisms into the aquifer) may lead to secondary abiotic changes (e.g. physicochemical effects brought about by the activities of the altered aquifer community) (Korbel et al., 2022); not to mention the fact that several abiotic and/or biotic changes may occur simultaneously. One previous study of a riverbank filtration system found that “changes in environmental conditions caused by the infiltrating river water” (i.e. species sorting) had a greater effect on the aquifer microbial community composition than “influx of river-derived [microbes]” (i.e. mass effects) (Fillinger et al., 2021). Overall, however, the effects of surface water intrusion on groundwater quality and groundwater ecosystems are likely to be highly context dependent.

1.5 Overview of methods relevant to this thesis

Many methods exist to study the quality of water extracted from groundwater wells as well as to detect the presence of surface water intrusion. For example, measurements of physicochemical water quality parameters (e.g. temperature, pH, EC, DO, ORP, turbidity, and the concentrations of various chemical species and nutrients) can provide useful basic information about the well water and about the kinds of influences the surrounding area might have on the well (Edmunds & Shand, 2008; Fitts, 2013; WHO, 2022). Microbiological indicators (e.g. *Escherichia coli*, coliform bacteria, enterococci, heterotrophs, coliphages, and microbial source tracking (MST) markers) can be used to detect water quality changes and faecal contamination events (V. Harwood et al., 2019; V. J. Harwood et al., 2014; WHO, 2022). The determination of stable water isotope compositions can be used to relate water samples to different phases of the hydrological cycle (Han et al., 2019; Hunt et al., 2005; N. Kortelainen, 2007, 2011; Rosenberry et al., 2015; Yang et al., 2023). More specifically, due to isotopic fractionation processes mainly associated with evaporation, lakes and oceans tend to be enriched with respect to the heavy isotopes ^{2}H and ^{18}O, whereas snowmelt waters tend to be depleted (J. Gat, 2010; J. R. Gat, 1996; Hunt et al., 2005). 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. Another method involves the analysis of microbial communities from water samples (e.g. via amplicon sequencing of the 16S rRNA gene) to gain insights into the composition and functional capabilities of microbial communities in the studied water bodies and wells (Chik et al., 2020; D. R. Clark et al., 2018; Kim et al., 2015). For example, surface waters often contain more phototrophs, and groundwaters more anaerobes; and exchange of these organisms can occur in groundwater–surface water interactions (Korbel et al., 2022). In addition to the manual methods described above, continuous monitoring methods are also increasingly being used to detect physicochemical and/or microbiological water quality changes in a variety of contexts, though use of these methods in groundwater wells is still somewhat rare (Oppus, Guico, et al., 2020; Oppus, Guzman, et al., 2020). The studies described in this thesis make use of all of the above methods, as combining several methods and data types can lead to a more holistic understanding than any one method used alone (Fig. 1).

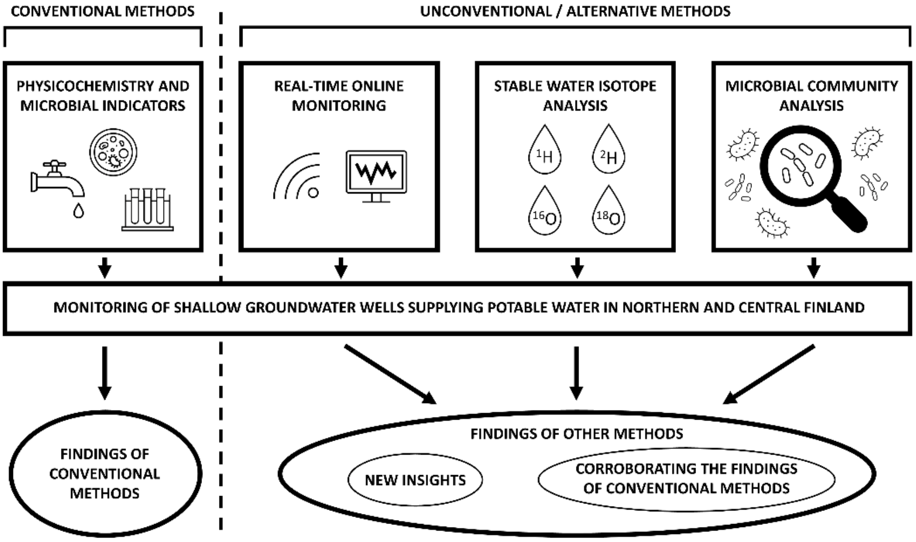

Fig. 1. **Overview of methods used for the work described in this thesis, including conventional methods such as periodic manual measurements of physicochemical parameters and microbiological indicators, as well as unconventional (alternative) methods such as (i) real-time online monitoring of groundwater physicochemistry, (ii) stable water isotope analysis, and (iii) microbial community analysis. Unconventional methods were applied with the aim of building a more holistic understanding of the factors influencing physicochemical and microbiological well water quality. Publicly available data on land use and weather conditions, obtained from the National Land Survey of Finland (NLS) and Finnish Meteorological Institute (FMI) respectively, were also considered, though these are not shown in the figure (Reprinted under CC BY 4.0 license from Publication II © 2022 Authors).**

1.6 Objectives of the research described in this thesis

The objective of the first study described in this thesis was (i) to investigate the physicochemical, microbiological, and isotopic characteristics of water from twenty-eight shallow wells used for potable water supply in northern and central Finland, and (ii) to explore how the well water quality at these sites might be influenced by the area surrounding each well (e.g. land use, roads, and surface water intrusion). Although there have been many previous studies of microbial communities in the deep groundwater of Finland (Bell et al., 2018; Bomberg et al., 2015, 2016; Itävaara et al., 2011; Kutvonen et al., 2015; Miettinen et al., 2015; Nyyssönen et al., 2012; Pedersen et al., 2014; Purkamo et al., 2013, 2018, 2020), the composition and diversity of communities in shallow groundwater have remained largely unexplored. One or two previous studies from Finland analysed a small number of groundwater samples, but never as the primary focus (Länsivaara, 2020; Tiwari, 2020; Tiwari et al., 2021). Thus, the first study described in this thesis was likely the first major study to observe and assess the microbial communities in water from shallow Finnish wells, and one of the first such studies in the Nordic countries (Knobloch et al., 2021).

Having completed this initial snapshot study, a second study was conducted, (i) to investigate the seasonal variation in the physicochemical, microbiological, and isotopic characteristics of water from four wells, and (ii) to assess whether less commonly used monitoring methods – namely, real-time online monitoring, stable water isotopes, and 16S rRNA amplicon sequencing – could be used to enhance water quality monitoring in shallow wells. 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 & Callegari, 2009; Storey et al., 2011; Velasquez-Orta et al., 2017). Thus, there is a need to continuously improve groundwater monitoring strategies, and explore alternative analytical methods, to ensure safe supplies of potable water.

In the first two studies, surface water intrusion from nearby surface water bodies emerged as a likely factor influencing well water quality. Hence, the third study aimed to further investigate the influence that different kinds of nearby surface water bodies – namely, a river, a pond, and a lake – have on well water quality, based on physicochemical, microbiological, and isotopic data collected approximately every two weeks for one year. Unlike the first two studies, this study produced 16S rRNA amplicon sequencing data from surface water as well as groundwater, thus enabling comparisons between the microbial communities of both. Such comparisons have not always been made in previous work (Fiedler et al., 2018).

1.7 Potential wider relevance of this thesis

Shallow groundwater is the preferred source for community water supply in many parts of the world (Foster et al., 2021; Hofkes et al., 2002; Liddle et al., 2016; Pritchard et al., 2008), and although local conditions differ widely, similar planning, management, and health-related issues are often observed (Hofkes et al., 2002; Howard et al., 2003; Isomäki et al., 2008; Mahmoud et al., 2022; Zaryab et al., 2017). Thus, although this thesis is centred on the study of wells extracting groundwater from shallow sand-and-gravel aquifers in Finland, this does not necessarily mean that the findings presented here are relevant only to the Finnish context.

In addition to the frequently made geological and hydrogeological comparisons between Finland and its nearest neighbours in the Nordic and Baltic regions (Holmlund et al., 2016; Kitterød et al., 2022; Knutsson, 2008), comparisons between Finland and certain regions of Canada are often made too, based on similarities between conditions left by the retreat of the Laurentide Ice Sheet (which covered most of Canada about 23000 to 10000 years ago) and those left by the retreat of the Fennoscandian Ice Sheet in Finland (Brennand, 2000; P. U. Clark & Walder, 1994): namely, thin soil covers overlying glacial till and glaciofluvial deposits (including sand-and-gravel eskers) (Brennand, 2000; Gale, 1993). Similarities of climate, widespread forest and peat cover, and heavy reliance on groundwater for potable water supply are also notable (Rivera, 2014).

Finland is also not the only country in which contaminated groundwater is the primary cause of waterborne outbreaks. The same appears to be true of the entire Nordic region, as well as the US, the UK, and Canada (G. F. Craun et al., 2003, 2010; M. F. Craun et al., 2006; Guzman-Herrador et al., 2015; Ligon & Bartram, 2016; Murphy et al., 2017). In light of these facts, further action is clearly needed to identify and understand the factors influencing well water quality, not only in Finland, but worldwide – and especially, perhaps, in rural areas, where financial and human resources are often less plentiful, but where investments in potable water provision have been shown to be highly cost beneficial (Hunter et al., 2009).

2 Materials and methods

2.1 The study regions and wells

This thesis is a compilation of findings from three studies of shallow groundwater wells used as sources of potable water in northern and central Finland. Most of the wells are in the North Ostrobothnia region (n = 27), with a smaller number in North Savo (n = 5) and Lapland (n = 1). Based on data from Statistics Finland for the year 2022, all three of these regions have below-average population densities (11.3, 14.3 and 1.9 inhabitants/km^2 compared to a national average of 18.3), as well as above-average proportions of the population living in sparsely populated rural areas (7.6%, 8.7% and 29.2% compared to a national average of 4.7%). Based on data from the nearest weather monitoring stations of the Finnish Meteorological Institute, daily air temperatures at the studied wells tend to fluctuate between about −25 and +25 °C throughout the year (with an annual mean of about +1 to +4 °C). Annual precipitation is about 550–750 mm, and there are about five to six months of snow cover annually, with the main period of snow cover typically lasting from November/December until April/May. All wells are located between about 62.5 and 66°N. Groundwater wells are considered critical infrastructure in Finland and therefore, to protect the wells, more precise coordinates cannot be reported here. The smallest of the studied wells serve less than ten cubic metres of water per day to tens of users, whereas the largest wells serve hundreds of cubic metres of water per day to networks that serve thousands of users. Many of the smaller wells studied here employed no water treatment during the period of investigation, and several others had only alkalisation to raise pH (Sallanko et al., 2013). Twenty-eight wells were investigated in the first study: twenty-two in North Ostrobothnia, five in North Savo, and one in Lapland. Four wells were investigated in the second study: three in North Ostrobothnia and one in North Savo. Six wells were investigated in the third study: all in North Ostrobothnia. All four wells from the second study and one of the wells from the third study were chosen from the initial twenty-eight. Thus, the total number of investigated wells was thirty-three.

Twenty of the twenty-eight wells in the first study (wells 3–22 in this thesis) are located in a single large municipality in North Ostrobothnia and were chosen largely by convenience, to provide a glimpse of physicochemical and microbiological groundwater quality in rural shallow wells and to reveal the extent of variation across a single municipality (Fig. 2). Wells 1 and 2 are also located in

North Ostrobothnia (in two other municipalities) and were selected as potentially problematic sites, based on a history of water quality issues. Well 23, in Lapland, was investigated by request of the well managers at that site, and wells 24–28, in North Savo, were selected due to a history of water quality issues, as well as geographical proximity to our collaborators at the Finnish Institute of Health and Welfare (THL). All of the wells studied for this thesis are located in areas with shallow groundwater tables (mostly <10 m below the surface), and all but one of the wells are located in rural areas. Well 26, by contrast, is located in a town with about 20000 residents.

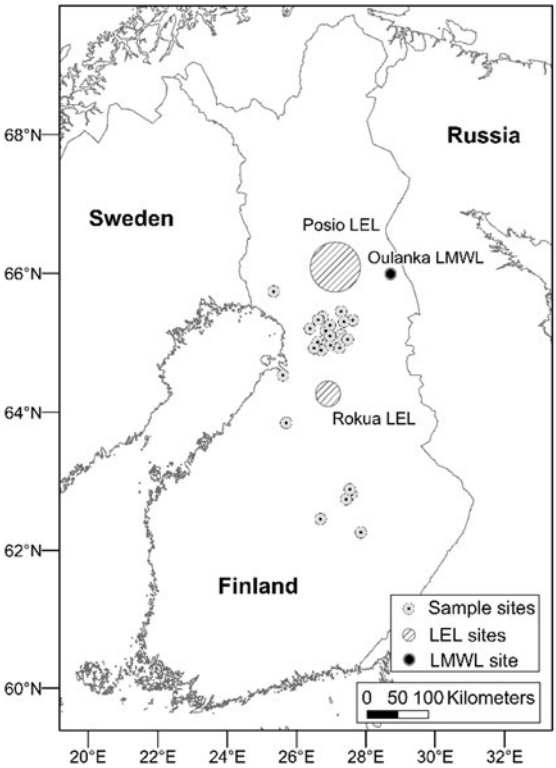

Fig. 2. Map showing the well locations for the first study, as well as the regions in which water samples were collected for stable water isotope analyses in earlier studies: from precipitation (black point; Oulanka LMWL, (Rossi et al., 2015)) and surface water (lined areas; Rokua LEL, (Isokangas et al., 2015), and Posio LEL, (Nora et al., 2019)). Abbreviations: LMWL, local meteoric water line; LEL, local evaporation line (Reprinted under CC BY 4.0 license from Publication I © 2021 Authors).

The four wells in the second study (Fig. 3) (wells W, X, Y and Z in this thesis) were chosen from the initial twenty-eight based on the findings of the first study and information from the well managers about site-specific problems, which raised concerns about potential risks to water quality at these sites (e.g. high Fe levels, suspected surface water intrusion, impacts from nearby land use). Well W is located about 31 m from a ditch, about 66 m from a small stream, and about 167 m from a river; well X is located about 60 m from a river; well Y is located in an area with several ponds (at distances of about 104 m, 161 m, 200 m and 209 m) including one used for swimming (at a distance of about 282 m); and well Z is located about 71 m from a lake.

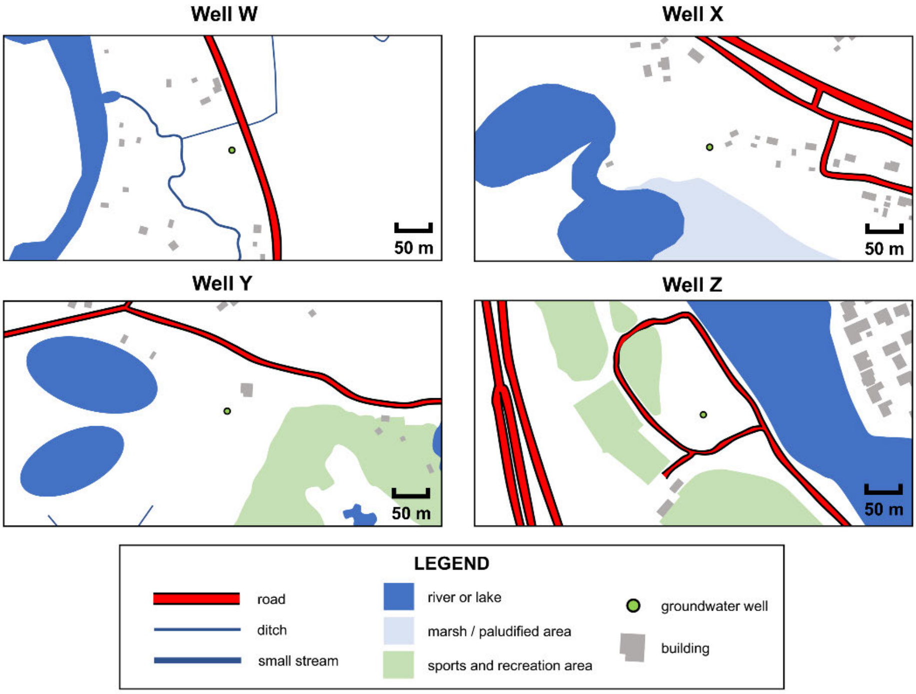

Fig. 3. Maps of the four wells investigated in the second study. The borders of nearby surface water bodies have been artificially smoothed at the request of the well managers, to preserve the confidentiality of the well locations (Reprinted under CC BY 4.0 license from Publication II © 2022 Authors).

In the third study, three sparsely populated study areas in the North Ostrobothnia region were investigated (Areas A, B, and C in this thesis) (Fig. 4). These three study areas were chosen because they each contain a groundwater well which is suspected to be under the influence of surface water intrusion based on the findings of the first two studies and other work (Yapiyev et al., 2023); these wells are hereafter known as the ‘suspect’ wells. In each of the three areas, a surface water body (located near the suspect well and thought to be the source of the intrusion) was also investigated, as was a ‘comparison’ well (located farther from the surface water body and therefore thought to be free from its influence). Hence, there were nine sampling/measurement sites in total: six groundwater wells and three surface water bodies. For Areas A, B, and 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 lake; the distance between the suspect well and the studied surface water body is about 60 m, 180 m, and 220 m.

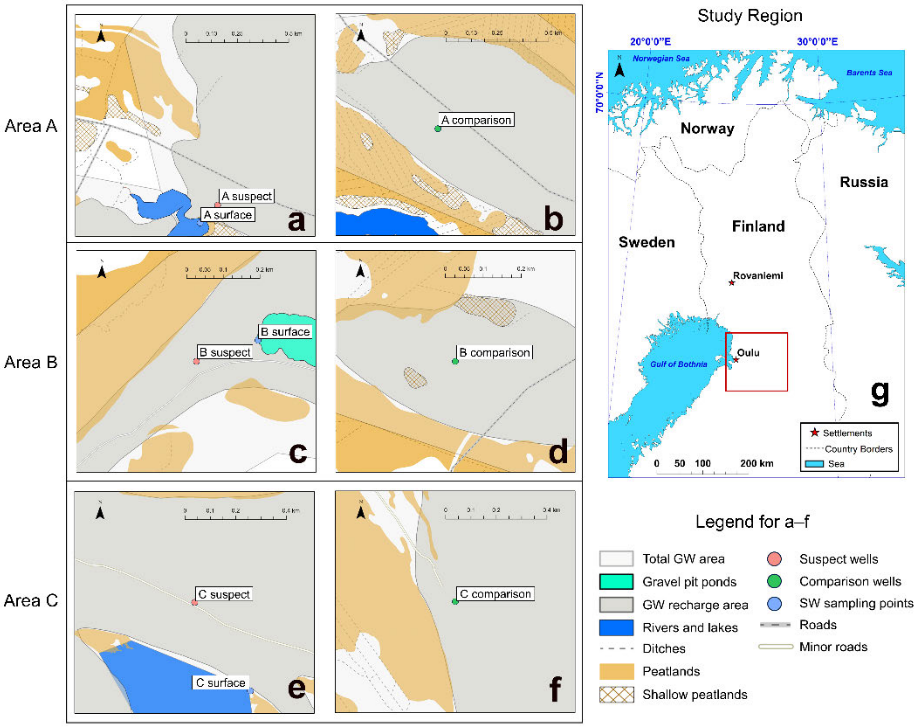

Fig. 4. Maps showing the locations of the nine sampling/measurement sites investigated in the third study (a–f) and the approximate study region, shown as a red square (g). The borders of nearby surface water bodies have been artificially smoothed at the request of the well managers, to preserve the confidentiality of the well locations. Abbreviations: suspect = suspect well (suspected to be under the influence of the nearby surface water body); comparison = comparison well (thought to be free of surface water influence); surface = surface water body; GW = groundwater; SW = surface water

2.2 Collection of land use data (I), weather data (II), and water level data (III)

In the first study, land-use data were extracted from maps produced by the National Land Survey of Finland (NLS) using their ‘MapSite’ tool. For each of the wells, measurements were made of distance to, and length or area of, nearby roads, fields, marshes, surface water sources, and buildings. ‘Nearby’ land use was defined as land use occurring within an area of about 1 km^2 around the well. In addition to the map-based analyses, informal site inspections were conducted during sampling visits to examine the immediate surroundings for potential risk factors, and well managers were probed for information on nearby land use and details of any past or suspected problems (e.g. high turbidity, surface water intrusion, faecal indicators). Regrettably, the presence or absence of some important potential risk factors, such as agricultural grazing, slurry pits, septic tanks and drain fields etc., was not formally determined.

In the second study, 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 monitoring stations of the Finnish Meteorological Institute (FMI). As wells W and X are in the same inland municipality, a single FMI monitoring station proved to be the nearest to both wells. For wells W, X, Y and Z, respectively, distance to the nearest FMI monitoring stations was about 16.8 km, 0.8 km, 14.5 km, and 21.1 km. Some of the sampling for the second study was arranged to coincide with periods of increased recharge (as described in following sections).

In the third study, water level timeseries data were collected from one of the wells and the three studied surface water bodies. Water level timeseries data for the river in Area A were collected from the Hertta database of the Finnish Environmental Institute (SYKE), whereas data for the Area A well, for the gravel pit pond in Area B, and for the lake in Area C were collected from Levelogger® 5 water level dataloggers (Solinst Canada Ltd., Ontario, Canada). Water level data for the river in Area A were directly available in N2000 height system values. Water level data from the Leveloggers were barometrically compensated using a method provided by Solinst and pressure data from the nearest FMI stations. These water level values were then converted into the N2000 height system with the help of elevation measurements from a Trimble R10 GNSS Receiver (Trimble, CO, USA). Thus, all water level data for the third study are reported in the same height system, as metres above sea level (m.a.s.l.) values, and are therefore directly comparable. Water level data in the third study were not used to determine the timepoints for water sampling and measurements, rather samples were collected at more or less regular intervals throughout the sampling period (as described in following sections).

2.3 Overview of sampling and measurement methods

This section is intended as a brief overview of sampling and measurement methods used in each of the three studies, for the reader's ease of reference. A table-based summary of all sampling, measurements, and data collection is shown in Table 1. More details regarding sampling and measurement methods can be found in the following sections.

In the first and second studies, methods included (i) collecting groundwater samples (for on-site and off-site physicochemical measurements, as well as stable water isotope analyses, and culture-based counts of *Escherichia coli*, coliform bacteria, heterotrophic bacteria, and spores of sulphite-reducing clostridia), (ii) filtering about 200 L of groundwater through a dead-end ultrafiltration (DEUF) capsule (to enable counts of coliphages, qPCR of microbial source tracker (MST) markers and Gram-negative bacteria, as well as amplicon sequencing of the V3–V4 regions of 16S rRNA and the rRNA gene), and (iii) collecting surface water samples from nearby surface water bodies (for stable water isotope analyses). For the first study (of twenty-eight wells), all of the above methods were implemented at a single timepoint during October–November 2018 (as this is often a period with increased groundwater recharge from autumntime rainfall), whereas for the second study (of four wells), all of the above methods were implemented at several timepoints during March–November 2019. The second study also involved the establishment of real-time online monitoring systems for the continuous monitoring of groundwater physicochemistry in the studied wells. The first two studies involved collaboration with the Finnish Institute for Health and Welfare (THL). The third study was conducted without this collaboration and had a specific focus on how well water quality is influenced by intrusion from nearby surface water bodies. In the third study, no culture-based counts of microbiological indicators, qPCR analyses, or real-time online monitoring of groundwater physicochemistry were performed (due to practical constraints). Instead, wells were sampled more frequently than in the first two studies (twice-monthly during October 2021–October 2022), and nearby surface water bodies were analysed at the same frequency and using the same methods as for the wells (unlike in the first two studies): in other words, groundwater and surface water samples were collected for (i) physicochemical measurements, (ii) stable water isotope analyses, and (iii) amplicon sequencing of the V3–V4 region of the 16S rRNA gene.

Table 1. Summary of sampling, measurements, and data collection.

Study	Sampling timepoints	Methods
I	one (Oct–Nov 2018)	On-site and off-site physicochemical measurements (GW) Stable water isotope analysis (GW & SW) Culture-based counts of microbiological indicators (GW) qPCR-based analyses of MST markers (GW) 16S rRNA amplicon sequencing of the V3–V4 region (GW) Land-use data (from NLS)
II	multiple (Mar–Nov 2019)	On-site and off-site physicochemical measurements (GW) Stable water isotope analysis (GW & SW) Culture-based counts of microbiological indicators (GW) qPCR-based analyses of MST markers (GW) 16S rRNA amplicon sequencing of the V3–V4 region (GW) Real-time online monitoring of physicochemistry (GW) Weather data (from FMI)
III	multiple (Oct 2021–Oct 2022)	On-site and off-site physicochemical measurements (GW & SW) Stable water isotope analysis (GW & SW) 16S rRNA amplicon sequencing of the V3–V4 region (GW & SW) Water level timeseries data (SW) (partly from SYKE)

GW = groundwater; SW = surface water; NLS = the National Land Survey of Finland; FMI = the Finnish Meteorological Institute; SYKE = the Finnish Environment Institute

2.4 Water sampling

Water sampling for the first study was carried out during October–November 2018 at twenty-eight shallow wells used for potable water supply. Each well was sampled at a single timepoint only. During sampling, raw (untreated) groundwater from each well was collected aseptically from a sampling tap into sample containers. Groundwater samples and samples from nearby surface waters were also collected for stable water isotope analyses in 15-mL high-density polyethylene tubes, which were rinsed with the sampled water before filling. The groundwater samples from most wells consisted of raw groundwater, which had not undergone any water treatment. The water samples from wells 11, 13, and 21, however, were alkalised due to the presence of alkalisation material in the wells. Sample containers were placed in cool boxes and transported to the laboratory within 24 h after sample collection.

Water sampling for the second study was carried out during eight monthly sampling timepoints at four shallow wells used for potable water supply (March–November 2019; excluding July, the summer holiday month for researchers and well managers alike). During sampling, raw (untreated) groundwater samples were collected at each of the study sites from a sampling tap. 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, (iii) a 1 L polypropylene (PP) bottle, for cultivation-based analyses of microbiological indicators, and (iv) a 15-mL high-density polyethylene tube, for analyses of stable water isotopes. All bottles and tubes were rinsed with the sample before filling. 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. 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. Surface water samples for stable isotope analyses were taken in 15-mL high-density polyethylene tubes during summertime and autumntime from surface water sources near wells W and X (within 300 m of the well in each case). These two sampling periods were chosen because springtime snowmelt and autumn rains tend to cause the largest peaks in groundwater recharge and are therefore periods of interest for changes in groundwater quality (R. Mäkinen et al., 2008; Okkonen et al., 2010). All samples were transported from the study sites to the laboratory in cool boxes and either processed or frozen within 24 h.

Water sampling for the third study was carried out twice-monthly for twelve months (October 2021–October 2022) at six groundwater wells and three surface water bodies. During each sampling round, untreated groundwater was collected from a groundwater sampling tap in each of the six wells following an initial flame-sterilisation step (i.e. spraying 70% ethanol onto the tap from a spray bottle and lighting the ethanol with a handheld lighter); and surface water was collected from the three surface water bodies via aseptic methods which varied depending on the time of year. Namely, 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 (<10 m from the non-flooding-period shoreline). 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.

2.5 Dead-end ultrafiltration (DEUF) sampling (I and II)

At each well in the first study, a large volume of groundwater (about 200 L) was filtered through a dead-end ultrafiltration (DEUF) capsule (ASAHI Rexeed-25A, Asahi Kasei Medical Co., Ltd., Tokyo, Japan) to concentrate the otherwise highly diluted microbes for further analysis, as described earlier (Inkinen et al., 2019). The flow rate during filtering was adjusted to around 1 L/18 s (3.33 L/min), which enabled 200 L of water to be filtered in about 1 h. DEUF capsules were placed in cool boxes and transported within 24 h to the laboratories for microbiological analyses.

The same DEUF method was employed at each well in the second study, though at two timepoints each, arranged to coincide with the snowmelt period in springtime (late April–early May 2019) and a period of rain in autumntime (late October 2019). A third DEUF sampling timepoint was arranged for well W only (in early October 2019), after a period of heavy autumntime rain. DEUF capsules were placed in cool boxes and transported within 24 h to the laboratories for microbiological analyses.

2.6 Physicochemical water quality measurements

2.6.1 Manual on-site measurements

For the first study, groundwater temperature, pH, dissolved oxygen (DO), redox potential, and electrical conductivity (EC) were measured on site using portable field meters, according to the manufacturer's instructions (SenTix 940, FDO 925, SenTix ORP 900, TetraCon 325; WTW, Weilheim, Germany). For the second study, the same parameters as for the first study were measured from groundwater on site with the WTW Multi 350i (at wells W, X and Y) and WTW Multi 3430 (at well Z) handheld meters and associated sensors (WTW, Weilheim, Germany). For the third study, again the same parameters were measured on site – however, this time from both surface water as well as groundwater – using a MultiLine® Multi 3630 IDS portable multi-parameter meter and associated sensors (WTW, Weilheim, Germany). For all studies, field redox potential values were converted to standard hydrogen electrode (ORP) values by temperature-based adjustment.

2.6.2 Manual off-site measurements

For the first study, analyses of total nitrogen (N_{tot}), ammonium nitrogen (NH_4^+-N), nitrite nitrogen (NO_2^--N), phosphorus (P), combined nitrate and nitrite nitrogen ($NO_3^-+NO_2^-$)-N, chloride (Cl^-), calcium (Ca), potassium (K), magnesium (Mg), sodium (Na), and silica (SiO_2) were performed in an accredited commercial laboratory according to international standards for chemical water quality. Iron (Fe) and sulphate (SO_4^{2-}) concentrations were determined colourimetrically via the phenanthroline method and the barium gelatine method, respectively (Fortune & Mellon, 1938; Tabatabai, 1974). Total carbon (TC) values were determined using a Sievers 900 portable TOC analyser. Turbidity (EN 27027:1994) was determined using a Hach Ratio XR turbidity meter. UV absorbance values of unfiltered and 0.45-µm-filtered water samples were determined at 254 nm (UV_{254}) using a UV-1800 spectrophotometer (Shimadzu, Japan) according to the manufacturer's instructions. For the second study, temperature, pH, EC, turbidity, UV absorbance at 254 nm, Fe, and Mn were measured from the groundwater samples in the laboratory.

For the third study, both surface water and groundwater samples were analysed. The turbidity of 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 sulphate (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 (P), total nitrogen (N), 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).

2.7 Real-time online groundwater monitoring (II)

In the second study, a real-time online monitoring system was installed in each of the four groundwater wells, to continuously assess well water physicochemistry. At wells W and Z, this system consisted 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 X and Y, a YSI 600 OMS V2-1 multiparameter optical monitoring sonde (YSI Inc., Yellow Springs, OH, USA) and a DL-12 data logger. 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. Groundwater temperature, EC, turbidity, and groundwater level were monitored at all wells, and pH, redox potential, and DO were monitored at wells W and Z. Measurements of all physicochemical parameters were taken by the YSI sonde, and groundwater level was monitored by a pressure sensor. In all cases, measurements were taken every 5 min during the monitoring period. Real-time monitoring proceeded for twelve months (February 2019– February 2020), although not all parameters were measured for this entire duration due to practical and planning contraints. At each site, measurements from the YSI sonde were transferred to the DL-12 data logger via the SDI-12 protocol. 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.

2.8 Stable water isotope analyses

To identify anomalies which may indicate surface water intrusion into the groundwater wells in the first study, stable water isotope (δ^2H, $\delta^{18}O$) analyses were conducted on untreated groundwater samples from each of the wells and on water samples taken from nearby surface water sources, such as rivers and lakes. The isotope ratios $^2H/^1H$ and $^{18}O/^{16}O$ were determined using cavity ring-down spectroscopy with a Picarro L2130-i analyser. All isotope ratios are expressed in δ notation relative to Vienna Standard Mean Ocean Water 2 (VSMOW2) with precision for δ^2H 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 compositions 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 compositions of local surface water sources. All wells are located with 450 km of the Oulanka, Rokua and Posio regions (more detailed figures are given in later sections). Deuterium excess (d-excess = $\delta^2H - 8\ \delta^{18}O$) was determined for the well water samples in order to study the effect of evaporative fractionation on the samples potentially resulting from surface water intrusion (Dansgaard, 1964). The isotope methods for the second study were identical to the first study.

For the third study, methods were identical to the first and second studies, with several additions. This study also incorporated the recently published precipitation-weighted least square regression (PWLSR) MWL for northern Finland, based on long-term data from Rovaniemi (location shown in Fig. 4g) (Yapiyev et al., 2023). Line-conditioned excess (lc-excess*) was also applied to identify what changes the isotope composition of water samples had undergone on the land surface (Landwehr & Coplen, 2004).

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

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

In the case of surface water intrusion in a suspect well a two-component, end-member mixing analysis was conducted (Christophersen & Hooper, 1992; F. Liu et al., 2004; Phillips et al., 2005), using lc-excess* as a tracer:

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

where R_{surface} is the fraction of surface water in the suspect well; and lc-excess* $_{\mathrm{(gw)}}$, lc-excess* $_{\mathrm{(source)}}$, and lc-excess* $_{\mathrm{(surface)}}$ 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.

2.9 Nucleic acid extractions

In the first two studies, raw groundwater from each well was filtered through dead-end ultrafiltration (DEUF) capsules on site (as described in Section 2.4). For both of these studies, DEUF capsules were eluted in the laboratory as described previously (Inkinen et al., 2019), except that the secondary concentration of DEUF eluates of 35 to 250 mL was performed by filtration through 0.22 µm Millipore Express PLUS membrane filters (Merck KGaA, Darmstadt, Germany). The filters were stored at −75 °C or lower prior to nucleic acid extraction. Total nucleic acids were extracted from the membrane filters using Chemagic DNA Plant Kit (Perkin Elmer, Waltham, MA, USA) as described previously (Brester et al., 2020). Total RNA was further purified using 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) for use in the 16S rRNA sequencing analysis. The total RNA was stored at −75 °C or lower, while the cDNA and the DNA extracts were stored at −20 °C until use.

Different methods were used for the third study, for several reasons. The large volume requirements and time-consuming nature of the DEUF method used in the first two studies were considered impractical for the higher (twice-monthly) sampling frequency of the third study; and 0.1 µm filters were deemed more suitable than 0.22 µm filters for ensuring that very small cells were not lost during the filtering step (Luef et al., 2015). Water samples collected in the Whirl-Pak® bags were first filtered in a Kojair BioWizard laminar flow hood (Kojair Tech Oy, Tampere, Finland). Specifically, samples were filtered via vacuum filtration through sterile porcelain Buchner filtering funnels equipped with 0.1 µm sterile polyethersulfone membrane filters (Sartorius, Göttingen, Germany). 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).

2.10 Microbiological indicator analyses (I and II)

For the first and second study, *Escherichia coli* and coliform bacteria were analysed from untreated groundwater samples according to standard methods using membrane filtration with LES Endo agar medium and Chromocult coliform agar medium (SFS 3016 and ISO 9308-1). Spores of sulphite-reducing clostridia (SSRC) were enumerated from water samples after heat treatment of membranes for 15 min at 75 °C and incubation for 2 days on tryptose sulphite cycloserine (TSC) agar (ISO 6461-2). Heterotrophic bacteria were enumerated from water samples by the spread-plate technique on Reasoner's 2 agar (R2A) medium (Difco, USA) and incubated at 22 ± 2 °C for 7 days (Kauppinen et al., 2019; Reasoner & Geldreich, 1985). Polyethylene glycol (PEG) precipitation of DEUF filtrate (200 to 500 mL) after secondary concentration via Millipore Express PLUS membrane filtration (as described above) was performed as described previously (Kauppinen et al., 2019), and analyses of somatic coliphages and F-specific coliphages were performed immediately from PEG precipitates using a double agar layer (DAL) procedure (USEPA Method 1601; with excess precipitate being stored at −75 °C or lower). In the second study, the Colilert-18® method (for the detection of coliforms and *E. coli*) was applied monthly to samples from well Z only (Fricker et al., 1997), as part of a more extensive monitoring programme at that site. Also in the second study, 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).

Gene copy (GC) numbers of two microbial source tracking (MST) markers – GenBac3 (targeting Bacteroidales bacteria as general indicators of faecal contamination) (Dick & Field, 2004; Siefring et al., 2008), and HF183 (targeting human-associated *Bacteroides* bacteria as indicators of human-derived faecal contamination) (Bernhard & Field, 2000a, 2000b; Converse et al., 2009; Haugland et al., 2010) – as well as gene copy numbers of Gram-negative bacteria (a general indicator of bacteria abundance) in the samples (including extraction and filtration blanks) were measured from cDNA and DNA extracts with TaqMan chemistry as described previously (Kärkkäinen et al., 2010; Tiwari et al., 2019). The qPCR assays were performed as described previously (Pitkänen et al., 2013), by processing 16 µL of RNA in a cDNA synthesis (reverse transcription [RT]). The qPCRs were performed using a QuantStudio 6 real-time PCR system (Applied Biosystems) in 20 µL volume using the TaqMan Environmental PCR master mix (Life Technologies), all with primers and probes at final concentrations 0.2 µM (IDT Technologies, Inc.) (Table 2). The cycling conditions included 95 °C for 10 min of enzyme activation and pre-denaturation followed by 40 cycles at 95 °C for 15 s of denaturation and at 60 °C for 60 s of annealing. Standard curves were generated using artificial gene fragments (gBlocks, IDT Technologies, Inc.) containing the sequences for each of the targeted genes. In qPCR, undiluted and 10-fold diluted cDNA and DNA samples were used as qPCR templates to detect PCR polymerase inhibition. For samples in which PCR inhibition was detected, qPCR data were generated using the results from diluted samples. Background signals, if detected in the filtration blanks, were subtracted from all the results to generate the final qPCR and RT-qPCR data per assay. The limit of detection (LOD) was set as 3 copies per reaction. The final qPCR, equivalent LOD (eLOD), and equivalent limit of quantification (eLOQ) values were calculated after considering the volume/mass of the processed sample, factors associated with the different processing steps of the RNA and DNA manipulations, and the dilutions used for each sample analysed.

Table 2. Primers and probes used for qPCR and RT-qPCR.

Assay	Primers and probes	Reference
HF183	F primer (HF183F): ATCATGAGTTCACATGTCCG	Bernhard & Field, 2000a
	R primer (BFDRev): CGTAGGAGTTTGGACCGTGT	Converse et al., 2009
	probe (BFDFAM): CTGAGAGGAAGGTCCCCCACATTGGA	Converse et al., 2009
GenBac3	F primer: GGGGTTCTGAGAGGAAGGT	Siefring et al., 2008
	R primer: CCGTCATCCTTCACGCTACT	Dick & Field, 2004
	probe: CAATATTCCTCACTGCTGCCTCCCGTA	Dick & Field, 2004
Gram-negative bacteria	F primer (Gram-F): GGGTTAAGTCCCGCAACGA	Kärkkäinen et al., 2010
	R primer (Gram-R): CATTGTAGCACGTGTGTAGCCC	Kärkkäinen et al., 2010
	probe (Gramneg-P): Cy5-TGACGTCAAGTCATCATGGCCCTTACG-BHQ3	Kärkkäinen et al., 2010

2.11 Microbial community analyses

2.11.1 16S rRNA amplicon sequencing

Sequencing methods were the same for the first two studies. Subsamples of the nucleic acids (DNA and cDNA) were sent to Macrogen Inc. (Seoul, South Korea) for amplicon generation and subsequent sequencing. Bacterial 16S rRNA genes were amplified from DNA (targeting all bacteria present) and cDNA (traditionally considered to target only metabolically active bacteria) using the primers Bakt_341F (5’-CCTACGGGNGGCWGCAG-3’) and Bakt_805R (5’-GACTACHVGGGTATCTAATCC-3’), which target the V3–V4 variable region of the 16S rRNA gene (Herlemann et al., 2011). Amplicons were sequenced as 2 × 300 bp paired-end reads using the Illumina MiSeq platform. Some samples were sequenced in duplicate to check for reproducibility. In both studies, negative controls from different sample processing steps were also sequenced (in the first study: a tube control for sampling/elution, elution solution, membrane filtration, nucleic acid extraction; in the second study: 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 Z were lost in transit to the sequencing company and therefore not sequenced.

For the third study, DNA extractions were sent (with a negative 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, 2 × 250 bp paired-end reads were generated via sequencing on the Illumina NovaSeq™ 6000 platform. A total of 198 DNA extractions were sent for sequencing. Of these, the extractions from 129 water samples and 1 negative control were successfully sequenced, and a total of >13000000 reads were generated (mean reads per sample was >100000).

2.11.2 Pre-processing of sequencing data

In all three studies, the sequencing data were processed and analysed via the QIIME 2™ bioinformatics pipeline (Bolyen et al., 2019). In the first study, the DADA2 denoise-paired QIIME 2™ plugin was used, with the parameters --p-trim-left-f 9, --p-trimleft-r 9, --p-trunc-len-f 290, and --p-trunc-len-r 250, to trim sequences (to remove bad quality reads with quality score of <20) and to denoise and merge trimmed reads to produce a table of amplicon sequence variants (ASVs) (Callahan et al., 2016). The DADA2 denoise-paired QIIME 2™ plugin was also used in the second study, but with the parameters --p-trim-left-f 17, --p-trim-left-r 21, --p-trunc-len-f 294, and --p-trunc-len-r 216; these parameters were chosen with the same rationale as for the first study, but the exact numbers differ slightly between studies due to differences in sequencing data quality. In the third study, the sequencing company (Novogene) trimmed the raw reads to remove adapters and sequencing primers. Reads, as received from Novogene, were imported into QIIME 2™ (Bolyen et al., 2019), and all remaining adapters, amplicon primers, and unknown nucleotides (Ns) 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 the DADA2 denoise-paired QIIME

2™ plugin with no additional trimming, thereby generating a table of ASVs (Callahan et al., 2016).

2.11.3 Taxonomic classification of ASVs

In all three studies, taxonomic classification of the ASVs was performed via the q2-feature-classifier plugin in QIIME 2™ (Bokulich et al., 2018) using a naive Bayes classifier trained on the V3–V4 variable region of representative 16S rRNA sequences from the SILVA rRNA database (Quast et al., 2012). In the first study, these representative 16S rRNA sequences were derived by clustering 16S rRNA sequences from the SILVA rRNA database (release 132) into operational taxonomic units (OTUs) based on 99% sequence identity (Quast et al., 2012). The default confidence cutoff of 70% was used in assigning taxonomic labels, as this is designed to balance precision and recall in classifying 16S rRNA sequences (Bokulich et al., 2018). Non-target sequences such as archaeal, mitochondrial, and chloroplastic sequences were filtered out, so that only bacterial sequences remained. In the second and third studies, the same method was used, although this time archaeal, mitochondrial, and chloroplastic sequences were not filtered out, and sequences from a newer release of SILVA were used to train the classifier: namely, 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).

In the first study, taxonomic classifications of bacterial DNA- and cDNA-based communities were screened for taxa relevant to drinking water safety, e.g. species such as *Escherichia coli*, *Clostridium perfringens*, and *Pseudomonas aeruginosa* and genera such as *Klebsiella*, *Aeromonas*, *Arcobacter*, *Enterococcus*, *Legionella*, *Mycobacterium*, *Yersinia*, and *Listeria*. Also in the first study, relative abundance values for the bacterial DNA- and cDNA-based communities were calculated in R, and a heatmap showing relative abundance values of selected taxa was generated using the pheatmap function from the ‘pheatmap’ R package (Kolde, 2019).

2.11.4 Alpha and beta diversity analyses

For analyses of alpha diversity in the first study, the ASV table was rarefied to a sampling depth of 2504, which excluded four samples with sequence counts below this threshold (three of these were controls, and one was a duplicate DNA sample). Remaining duplicates were checked for consistency and merged. Alpha diversity metrics for bacterial DNA- and cDNA-based communities – namely, Faith's phylogenetic diversity, Pielou's evenness, observed ASVs, and Shannon's diversity index – were calculated based on the rarefied ASV table using the QIIME 2™ diversity plugin (Halko et al., 2011; Sørensen, 1948). For the second study, the same alpha diversity metrics, as well as beta diversity metrics and principal coordinates analysis (PCoA) of the Bray-Curtis dissimilarity matrix were calculated and derived using the QIIME 2™ diversity plugin and a sampling depth of 18332. In the third study, alpha and beta diversity metrics were also calculated via the QIIME 2™ diversity plugin, using a sampling depth of 18187.

For each alpha diversity metric in the third study, 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.

In the first study, non-metric multidimensional scaling (nMDS) plots using weighted UniFrac distance matrices (calculated from the rarefied ASV table in QIIME 2™) were created in R using the metaMDS function in the 'vegan' package (Oksanen, 2020). Metadata variables (e.g. turbidity, dissolved oxygen, nitrate, ammonium) were fitted to the nMDS plots using the 'envfit' function. In the second and third studies, nMDS plots were prepared in the same manner, except using the Bray-Curtis dissimilarity metric instead (Oksanen, 2020). Bray–Curtis appears to be the most widely used beta diversity metric, and so there are advantages to its use in terms of comparability between studies (Kers & Saccenti, 2022). For all studies, 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 figures were made with the 'ggplot2' R package (Wickham, 2016). For the third study, 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' package (Oksanen, 2020), and pairwise comparisons were performed by pairwise PERMANOVA using the 'pairwise.adonis2' function of the 'pairwiseAdonis2' 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.

2.12 Correlation analysis (I)

In the first study, correlations were sought between physicochemical data, microbial indicator data, land use data, and bacterial alpha diversity data via calculation of Spearman rank-based correlation coefficients (r_s) using the ‘rcorr’ function from the ‘Hmisc’ R package (Harrell, 2020). Spearman was used here because Shapiro-Wilk tests carried out in R showed that many parameters were not normally distributed. Correlograms were produced using the ‘corrplot’ function from the ‘corrplot’ R package (Wei, 2017). Only statistically significant Spearman correlation coefficients ($p < 0.05$) are reported here, with the following ranges being used for discussion of correlations: very strong (>0.8 or <−0.8), strong (between 0.6 and 0.8, or between −0.6 and −0.8), moderate (between 0.4 and 0.6, or between −0.4 and −0.6), weak (between 0.3 and 0.4).

2.13 Principal component analysis (PCA) of physicochemical water quality data (III)

Principal component analysis (PCA) of the physicochemical water quality data was conducted in R. The raw data were first cross-checked against notes made during sample collection and information obtained from the well managers. Values which were missing due to sampling/measurement errors 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 & Husson, 2016). Plotted PC axes were very stable indicating that the uncertainty of the imputed data was low. Overall, about 92% of data points were obtained directly from sampling and measurements, and about 8% were imputed. 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).

2.14 Sequencing data availability

The 16S rRNA amplicon sequencing data for the studies described in this thesis have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under primary accession numbers PRJEB41020 (Publication I), PRJEB52434 (Publication II), and PRJEB66003 (Publication III).

3 Results and discussion

3.1 Snapshot study of well water quality and environmental influences (I)

3.1.1 Wells, physicochemistry and possible influences

Most of the twenty-eight wells investigated for the first study were shallow dug wells (well depth <11 m) serving <75 m^3 of water/day to <500 users, under the management of a co-operative (Table 3). There were, however, also some shallow tube wells, wells managed by municipalities, and wells supplying water to distribution networks serving many thousands of users. During the period of investigation, fourteen of the wells employed no water treatment whatsoever, eight had only alkalisation to raise pH, and six had ultraviolet disinfection (Table 3).

The water from most of the wells was oligotrophic (median N_{tot}: 0.12 mg/L; median P: 6.9 µg/L), cool (median temperature: 6.3 °C), oxic (median DO: 9 mg/L) and slightly acidic (median pH: 6.4) (Table 4). These characteristics are quite typical for Finnish groundwaters (Isomäki et al., 2008). Some wells had high pH (up to 8.1) because of alkalinisation material in the wells (e.g. well 21) or because of naturally high levels of Ca and Mg in the water (e.g. well 23).

Well 1 had the highest water temperature (8.7 °C), turbidity (2.98 NTU), Fe (2.3 mg/L) and P (100 µg/L), and the lowest ORP (71.5 mV) and DO (0.42 mg/L). Such turbidity and Fe values exceed the target levels outlined in Finnish law (1 NTU and 0.2 mg/L, respectively). Reddish-brown staining and biofilm were observed on pipes at this site, suggesting that the high Fe levels were causing high turbidity and growth of iron-oxidizing bacteria. This well is in a clay-rich coastal area of a kind that, in Finland, is often associated with high Fe and Mn, and low DO (Fältmarsch et al., 2008; Isomäki et al., 2008). High P in this well might be explained by the fact that low DO can cause Fe oxides in soils and aquifers to dissolve and release adsorbed P into the water (Domagalski & Johnson, 2012).

Well 2 had the highest N (8.1 mg/L), NH_4^+-N (86 µg/L), NO_2^--N (14 µg/L), (NO_3^-+NO_2^-)-N (8 mg/L), and K (22.7 mg/L), as well as the largest ‘nearby field area’ (fields make up about 72% of the 1 km^2 around the well), suggesting that high input of these nutrients may be coming from nearby agriculture (Burkart & Stoner, 2008). Wastewater leakages from septic tanks and sewage systems are another potential source of N in groundwater (Isomäki et al., 2008). However, the HF183

MST marker (targeting human-associated *Bacteroides* bacteria as indicators of human-derived faecal contamination), was not detected in this study, thereby somewhat diminishing (but not completely ruling out) this possibility.

Well 25 had the highest SO_4^{2-} (171 mg/L) and SiO_2 (9.86 mg/L). The SO_4^{2-} may have come from sulphur-containing fertilizers used in nearby agriculture (the well had the fourth largest ‘nearby field area’) or from the weathering of rocks and minerals, which are also potential sources of SiO_2 (Piispanen & Nykyri, 1997; Porowski et al., 2019).

Several wells also showed relatively high concentrations of Na and Cl^-, potentially indicating the influence of road salt used for de-icing in winter (Gustafsson & Nystén, 2000; Salminen et al., 2011). The strongest example of this was well 26, which had the highest Na (17.8 mg/L), Cl^- (43.7 mg/L), and EC (344 µS/cm), as well as the largest ‘total nearby road length' (8074 m within an area of about 1 km^2 around the well). Sodium chloride (NaCl) is the main type of road salt used in Finland (Salminen et al., 2011).

Thus, based on the observations above, the physicochemical water quality in several wells may be influenced by several factors in the surrounding environment (such as soil type, agriculture, weathering of rocks and minerals, and road salt). However, these data alone are only suggestive of these influences rather than providing conclusive evidence of them. Further work would be needed to confirm these influences and to identify their sources. Regrettably, the risk assessment approach adopted here was rather informal and unsystematic, and the presence or absence of some important potential risk factors, such as agricultural grazing, slurry pits, septic tanks and drain fields etc., was not formally determined. A more thorough, systematic and standardised survey of nearby risks would likely provide more certainty as to the sources and magnitudes of the potential influencing factors discussed above, and therefore act as a more secure basis for site-specific risk mitigation.

Table 3. General characteristics of wells 1–28.

Well	Type	Well depth (m)	GW depth (m)	Users[a]	Yield (m^3/d)	Management	Treatment
1	Tube	≥8	1.5	7000[b]	NA	municipal	UV, ALK, CH
2	Dug	6	3	6400[b]	250	municipal	UV, ALK
3	Dug	6	2.5	190	32.9	co-operative	none
4	Dug	~3	1.5	10	1	co-operative	none
5	Dug	3	1.5	40	24	co-operative	ALK

Well	Type	Well depth (m)	GW depth (m)	Users[a]	Yield (m^3/d)	Management	Treatment
6	Tube	8	4	150	16	co-operative	none
7	NA	NA	NA	<50	<5.5	co-operative	none
8	Dug	6	2	<100	13.7	co-operative	none
9	Dug	5	2	100–200	11	co-operative	none
10	Dug	5.5	2	105	8.9	co-operative	none
11	Dug	6	4.2	100	<11	co-operative	ALK[c]
12	Dug	7	2	~170	12	co-operative	none
13	Dug	5	2.5	~280	11	co-operative	ALK[c]
14	Dug	NA	2	50	8.3	co-operative	none
15	Dug	7	3	150	125	co-operative	none
16	Tube	11	3	150	65	co-operative	ALK
17	Dug	4	1	50–60	16.4	co-operative	none
18	Dug	6	3.5	155	71.2	co-operative	none
19	Dug	7.5	3	4200[b]	650	co-operative	UV, ALK
20	Dug	4	~2	80	27.4	co-operative	none
21	Dug	3.5	1	<50	10	co-operative	ALK[c]
22	Tube	7	>0.5[d]	28	5	co-operative	none
23	Tube	7	4	20000[b]	NA	municipal	UV, ALK
24	Tube	9.2	2	1000	170	municipal	UV, ALK
25	Tube	7.5	3	>200	38	municipal	ALK
26	Dug	9	5–10	20000[b]	600–1000	municipal	UV, ALK
27	Tube	NA	NA	500	120	co-operative	ALK
28	Tube	NA	NA	2000	400	co-operative	ALK

GW = groundwater; ALK = alkalisation; UV = ultraviolet disinfection; CH = chemical purification; [a] = many of these are rough estimates rather than exact values; [b] = water served from several wells to the same distribution network; [c] = raw water samples have been alkalized; [d] = artesian spring

Table 4. Summary of physicochemical water quality parameters for wells 1–28.

	Median	Max	Min
Temperature (°C)	6.3	8.7	5.1
pH	6.4	8.1	5.7
EC (µS/cm)	51	344	23
DO (mg/L)	8.9	12.9	0.4
ORP (mV)	330	391	72
Turbidity	0.32	2.98	0.15
UV_{254} (unfiltered water)	0.011	0.323	0.000
UV_{254} (0.45-µm-filtered water)	0.074	2.075	0.009
TC (ppm)	9.4	36	4.4
Fe (mg/L)	<0.07	2.3	<0.07
SO_4^{2-} (mg/L)	4.05	171.13	1.23

	Median	Max	Min
N_{tot} (µg/L)	120	8100	<50
NH_4^+-N (µg/L)	2.5	86	<5
NO_2^--N (µg/L)	<2	14	<2
P (µg/L)	6.9	100	<3
(NO_3^-+NO_2^-)-N (µg/L)	69	8000	2.5
Cl^- (mg/L)	1.6	43.7	0.6
Ca (mg/L)	4.43	30	1.46
K (mg/L)	1.08	22.7	<0.5
Mg (mg/L)	1.3	12	0.5
Na (mg/L)	3.1	17.8	1.8
SiO_2 (mg/L)	6.28	9.86	3.97

EC = electrical conductivity; DO = dissolved oxygen; ORP = redox potential; UV_{254} = ultraviolet absorbance at 254 nm; TC = total carbon; SO_4^{2-} = sulphate; N_{tot} = total nitrogen; NH_4^+-N = ammonium nitrogen; NO_2^--N = nitrite nitrogen; (NO_3^-+NO_2^-)-N = combined nitrate and nitrite nitrogen; Cl^- = chloride; SiO_2 = silica

3.1.2 Stable water isotope compositions

Groundwater samples from most wells had stable water isotopes compositions in the vicinity of the Oulanka LMWL (Rossi et al., 2015), with $\delta^{18}O$ between −14.22 and −12.82, δ^2H between −102.33 and −94.09, and d-excess between 7.26 and 11.44 (Table 5 and Fig. 5). These compositions are typical of Finnish groundwaters where local precipitation (especially snowmelt) is the main source of recharge (N. M. Kortelainen & Karhu, 2004). By contrast, surface water samples mainly followed the Rokua LEL (Isokangas et al., 2015).

Groundwater samples from wells 1 and 2 deviated from most others, following the Rokua LEL rather than the Oulanka LMWL, with $\delta^{18}O$ of −12.22 and −12.41, δ^2H of −90.04 and −90.84, and d-excess of 7.68 and 8.46, respectively.

Groundwater samples from wells 18, 20, and 21 also deviated from most others, following the Posio LEL rather than the Oulanka LMWL (Nora et al., 2019), with d-excess values varying between 3.63 and 4.59 (a different range to the other wells). Wells 18 and 20 were relatively unremarkable based on physicochemical data, although well 21 had above-median pH, Ca, EC, and TC.

The deviation of these five groundwater samples from the LMWL indicates possible surface water intrusion at these five wells (i.e. wells 1, 2, 18, 20, and 21) (Hunt et al., 2005). Regrettably, however, water level data were not collected in this study (neither from wells nor surface waters), and thus the hydraulic gradients at the studied sites at the time of sampling were not determined. A more detailed multivariate assessment of the intrusion risks at the studied sites would require knowledge of the hydraulic gradients and would ideally also include data regarding subsoil type and thickness, as well as other hydrogeological parameters. Repeated sampling campaigns at the suspected sites would also be desirable to better understand the temporal dynamics of the intrusion, which remained unexplored in this snapshot study.

Table 5. Stable water isotope data for wells 1–28 and distances[a] to comparison regions.

Well	δ^2H (‰)	δ^{18}O (‰)	d-excess (‰)	km to Oulanka	km to Rokua	km to Posio
1	−90.04	−12.22	7.68	249	65	196
2	−90.84	−12.41	8.46	307	68	255
3	−99.97	−13.71	9.70	136	127	83
4	−95.84	−13.27	10.30	144	122	92
5	−96.96	−13.24	8.94	164	108	111
6	−97.34	−13.37	9.59	154	97	101
7	−94.09	−12.82	8.48	148	105	95
8	−100.17	−13.91	11.12	137	115	84
9	−97.97	−13.60	10.84	172	83	119
10	−97.30	−13.23	8.51	170	86	117
11	−102.33	−14.22	11.44	164	84	111
12	−97.28	−13.07	7.26	162	83	110
13	−95.60	−13.27	10.54	177	73	125
14	−95.58	−13.27	10.62	182	76	129
15	−99.63	−13.80	10.73	108	134	57
16	−100.23	−13.79	10.06	108	141	55
17	−99.16	−13.79	11.12	118	126	66
18	−95.73	−12.54	4.59	139	104	87
19	−100.40	−13.80	9.98	150	98	97
20	−96.80	−12.67	4.56	158	84	107
21	−95.68	−12.41	3.63	144	99	92
22	−98.18	−13.31	8.30	140	103	90
23	−96.41	−13.15	8.77	189	183	147
24	−96.51	−13.38	10.53	431	203	385
25	−94.48	−12.90	8.74	441	235	400
26	−96.06	−13.26	10.01	381	171	338
27	−95.76	−13.21	9.89	391	177	347
28	−98.52	−13.55	9.91	374	163	331

[a]distances reported here are approximate and rounded to the nearest kilometre

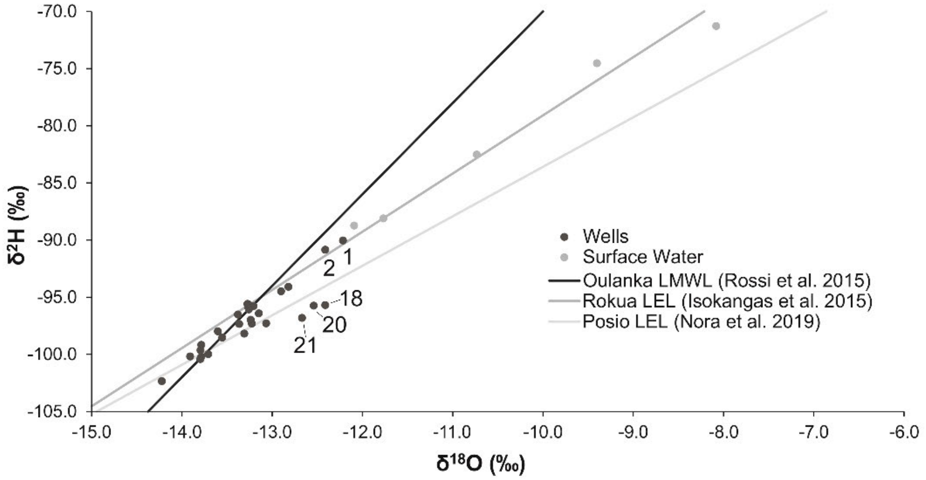

Fig. 5. Stable water isotope results from the twenty-eight wells compared to surface water samples and rainfall. Data for Oulanka local meteoric water line (LMWL) and Rokua and Posio local evaporation lines (LEL) were extracted from previous studies (Isokangas et al., 2015; Nora et al., 2019; Rossi et al., 2015). Wells with deviating d-excess values and/or alignment with LEL lines are marked with well numbers (Reprinted under CC BY 4.0 license from Publication I © 2021 Authors).

3.1.3 Microbiological indicators

Detected microbial loads were low in most wells. Coliform bacteria were detected in exactly half of the wells (n = 14), but mostly at low levels (<20 CFU/L). Microbiological indicators detected in at least two wells are shown in Table 6.

E. coli, somatic coliphages, and the HF183 MST marker (targeting human-associated *Bacteroides* bacteria as indicators of human-derived faecal contamination) were not detected in any of the wells, and F-specific coliphages and spores of sulphite-reducing clostridia (SSRC) were detected in only one well each: 0.04 PFU/L of F-specific coliphages in well 2, and 1 CFU/100 mL of SSRC in well 1. *E. coli*, coliphages, and SSRC are potential indicators of faecal contamination (Marcheggiani et al., 2008; Roslev et al., 2004; Singh et al., 2022).

The GenBac3 MST marker (targeting Bacteroidales bacteria as general indicators of faecal contamination) was detected in DNA extracts from wells 1, 2 and 15, and cDNA extracts from wells 1, 2, 7, 8, 15, 16, 20, 21, 24, and 26 (Table 6), indicating that these ten wells may have been previously subjected to some degree of non-human faecal contamination. However, as *E. coli* – often considered evidence of recent contamination – was not detected in any of the wells, the contamination indicated by the GenBac3 MST marker may be historical rather than recent (WHO, 2022).

Well 2 had the highest indication of Gram-negative bacteria abundance (2.5 × 10^5 and 3.7 × 10^7 GC/100 mL in DNA and cDNA extracts respectively), the highest levels of the GenBac3 marker (8 and 260 GC/100 mL in DNA and cDNA extracts respectively), the second highest heterotrophic plate count (1200 CFU/mL), the second highest coliform count (80 CFU/L), and was the only well in which F-specific coliphages were detected (0.04 PFU/L). These observations, combined with the potential agricultural influence and possible indications of surface water intrusion described in previous sections, suggest that further investigation of water quality risks may be warranted at this site.

Well 21 had the highest heterotrophic plate count (16000 CFU/mL) and the second highest indication of Gram-negative bacteria abundance based on cDNA extracts (9.7 × 10^7 GC/100 mL).

Water intruding from a nearby surface water body could be one potential source of the microbiological indicators detected in wells 2 and 21, as surface water tends to have higher prokaryotic counts than groundwater (Alfreider et al., 1997; Brugger et al., 2001; Van Driezum et al., 2018; Vargha et al., 2023; Wang et al., 2022; Zhou et al., 2012), and is a known source of microbial contamination in groundwater (P. D. Hynds et al., 2014; Sasakova et al., 2018; Van Driezum et al., 2018). Other sources are also possible, however, such as (i) the infiltration (or preferential flow) of precipitation water through the unsaturated layer to the groundwater table, and/or (ii) the direct ingress of water into a well due to flooding, or due to precipitation or surface runoff entering a structurally comprised well (Andrade et al., 2018).

Table 6. Selected[a] microbiological indicators for wells 1–28.

Well	Coliforms[b] (CFU/L)	Coliforms[c] (CFU/L)	Heterotrophs (CFU/mL)	GenBac3 DNA (GC/100 mL)	GenBac3 cDNA (GC/100 mL)	Gram⁻ DNA (GC/100 ml)	Gram⁻ cDNA (GC/100 ml)
1	4	1	77	7	29	9.6 × 10^4	5 × 10^5
2	80	80	1200	8	260	2.5 × 10^5	3.7 × 10^7
3	n.d.	n.d.	20	n.d.	n.d.	2.2 × 10^3	n.d.
4	2	n.d.	420	n.d.	n.d.	7.2 × 10^4	4.7 × 10^6
5	n.d.	n.d.	25	n.d.	n.d.	5.9 × 10^3	2.1 × 10^4
6	n.d.	n.d.	210	n.d.	n.d.	9.4 × 10^3	n.d.
7	n.d.	n.d.	160	n.d.	49	4.2 × 10^4	4 × 10^5
8	n.d.	20	45	n.d.	67	1.5 × 10^4	1.2 × 10^6

Well	Coliforms[b] (CFU/L)	Coliforms[c] (CFU/L)	Heterotrophs (CFU/mL)	GenBac3 DNA (GC/100 mL)	GenBac3 cDNA (GC/100 mL)	Gram$^-$ DNA (GC/100 ml)	Gram$^-$ cDNA (GC/100 ml)
9	20	20	160	n.d.	n.d.	5.8×10^4	2.4×10^6
10	n.d.	n.d.	77	n.d.	n.d.	1.6×10^4	2.9×10^5
11	n.d.	n.d.	220	n.d.	n.d.	3.9×10^4	1.3×10^6
12	10	1	140	n.d.	n.d.	3×10^3	2.9×10^5
13	n.d.	n.d.	110	n.d.	n.d.	1.6×10^4	1.4×10^6
14	n.d.	n.d.	640	n.d.	81	1×10^5	5.6×10^6
15	4	4	270	4	183	7.8×10^3	n.d.
16	n.d.	n.d.	730	n.d.	n.d.	1.7×10^3	n.d.
17	260	210	450	n.d.	n.d.	1.1×10^4	1.1×10^6
18	n.d.	n.d.	10	n.d.	n.d.	2.5×10^3	n.d.
19	n.d.	1	5	n.d.	n.d.	8.9×10^3	1.1×10^6
20	1	n.d.	340	n.d.	1	7.5×10^3	1.4×10^5
21	13	n.d.	16000	n.d.	1	6.3×10^4	9.7×10^6
22	n.d.	n.d.	380	n.d.	n.d.	2.2×10^4	5.3×10^6
23	n.d.	n.d.	10	n.d.	n.d.	5.6×10^3	2.4×10^5
24	n.d.	n.d.	20	n.d.	12	1.4×10^3	n.d.
25	n.d.	n.d.	30	n.d.	0	1.7×10^3	n.d.
26	4	n.d.	35	n.d.	1	2.1×10^4	1×10^6
27	n.d.	3	10	n.d.	n.d.	4×10^4	3×10^6
28	n.d.	n.d.	10	n.d.	n.d.	1.5×10^3	n.d.

[a]only parameters detected in at least two wells are shown, others are summarised in the text; [b]SFS 3016 method; [c]SFS-EN ISO 9308-1 method; n.d. = not detected

3.1.4 Alpha diversity metrics

Bacterial 16S rRNA amplicons were sequenced from DNA and cDNA extracted from groundwater samples, processed, and alpha diversity metrics calculated; namely, (i) observed ASVs, (ii) Faith's phylogenetic diversity, (iii) Pielou's evenness, and (iv) Shannon's diversity index (Table 7). Median values for these metrics, in the order given above, were 742, 79, 0.89, and 8.52 for DNA; and 489, 57, 0.88, and 7.68 for cDNA, respectively.

Wells 2 and 21 were exceptional here again, being characterized by relatively low values of all four alpha diversity metrics in both DNA and cDNA libraries. Values for well 2 were 321, 43, 0.65, and 5.39 for DNA; and 184, 22, 0.67, and 5.06 for cDNA; in other words, about 57%, 46%, 27%, and 37% below the median for DNA; and about 62%, 61%, 24%, and 34% below the median for cDNA. Values for well 21 were 404, 48, 0.81, and 6.98 for DNA; and 100, 11, 0.67, and 4.48 for cDNA; in other words, about 46%, 39%, 9% and 18% below the median for DNA; and about 79%, 81%, 24%, and 42% below the median for cDNA.

These relatively low bacterial alpha diversity values, combined with the relatively high microbial counts described in the previous section, suggest that, at the time of sampling, the microbial communities at these two wells were dominated by a limited subset of organisms. Surface water intrusion is one possible explanation for this, given that possible indications of intrusion were detected in both wells, and that the microbial communities of surface water tend to be less diverse than those of groundwater (Fillinger et al., 2021; Ji, Wang, et al., 2022; Ji, Zhang, et al., 2022).

Table 7. Alpha diversity metrics for wells 1–28.

Well	Observed ASVs		Faith's PD		Pielou's evenness		Shannon	
	DNA	cDNA	DNA	cDNA	DNA	cDNA	DNA	cDNA
1	906	227	92	25	0.88	0.79	8.65	6.11
2	321	184	43	22	0.65	0.67	5.39	5.06
3	844	718	79	70	0.91	0.92	8.86	8.70
4	782	411	86	54	0.94	0.89	9.02	7.66
5	736	807	78	80	0.87	0.92	8.32	8.93
6	747	483	78	56	0.89	0.84	8.50	7.49
7	584	376	60	42	0.85	0.89	7.78	7.58
8	971	688	94	71	0.92	0.88	9.17	8.26
9	1153	537	104	63	0.95	0.89	9.65	7.97
10	818	809	89	78	0.90	0.91	8.74	8.80
11	823	450	87	51	0.92	0.87	8.94	7.64
12	752	174	79	21	0.83	0.42	7.94	3.15
13	735	494	75	57	0.90	0.83	8.55	7.43
14	692	174	81	23	0.91	0.80	8.54	5.92
15	578	286	60	33	0.69	0.57	6.33	4.63
16	500	414	57	46	0.83	0.84	7.44	7.29
17	942	518	94	64	0.93	0.91	9.18	8.19
18	538	572	60	64	0.82	0.86	7.41	7.88
19	836	658	84	70	0.92	0.91	8.96	8.48
20	817	831	83	79	0.90	0.90	8.68	8.73
21	404	100	48	11	0.81	0.67	6.98	4.48
22	492	408	63	47	0.69	0.78	6.15	6.80
23	837	819	95	89	0.93	0.91	9.01	8.77
24	666	592	68	58	0.89	0.89	8.38	8.22
25	587	648	68	67	0.87	0.88	8.03	8.21
26	809	957	84	88	0.92	0.93	8.84	9.21

Well	Observed ASVs		Faith’s PD		Pielou’s evenness		Shannon	
	DNA	cDNA	DNA	cDNA	DNA	cDNA	DNA	cDNA
27	451	442	55	51	0.84	0.88	7.40	7.70
28	685	340	79	43	0.87	0.71	8.20	5.98
Median	742	489	79	57	0.89	0.88	8.52	7.68
Max	1153	957	104	89	0.95	0.93	9.65	9.21
Min	321	100	43	11	0.65	0.42	5.39	3.15

3.1.5 16S rRNA amplicon sequencing data

The most commonly identified bacterial taxa in DNA – based on mean relative abundance values of phyla, with Proteobacteria split to the class level – were Patescibacteria (44%), Gammaproteobacteria (12%), Omnitrophicaeota (8%), Deltaproteobacteria (7%), and Bacteroidetes (5%). Members of the Patescibacteria superphylum have previously been shown to have high relative abundance in DNA-derived 16S rRNA gene amplicons from groundwater environments (Bruno et al., 2017; Herrmann et al., 2019; Kumar et al., 2017; Schwab et al., 2017). These bacteria have particularly small cell sizes and are not easily cultivated (Luef et al., 2015). The most commonly identified taxa in cDNA were Gammaproteobacteria (32%), Deltaproteobacteria (12%), Patescibacteria (8%), Bacteroidetes (8%), and Entotheonellaeota (4%).

In twenty of the twenty-eight wells, Patescibacteria were dominant in DNA and Gammaproteobacteria were dominant in cDNA. Five wells, however, clearly showed the opposite pattern: wells 2, 12, and 15 had >10% higher relative abundances of Gammaproteobacteria than Patescibacteria in DNA; and wells 18 and 28 had >15% higher relative abundances of Patescibacteria than Gammaproteobacteria in cDNA. Of these five wells, wells 2 and 18 are suspected to be under the influence of surface water intrusion, as discussed in previous sections.

When the most abundant bacterial classes were considered, DNA-derived and cDNA-derived 16S amplicons largely clustered apart (Fig. 6), indicating general differences between ‘present’ (DNA) and metabolically ‘active’ (cDNA) bacterial communities; these differences are often observed in other studies (Herrmann et al., 2020; Inkinen et al., 2019; Schwab et al., 2017). However, some of the cDNA libraries (i.e. those from wells 1, 2, 12, 15, 21), formed a small outlier group during clustering (far left of Fig. 6), with noticeably higher relative abundances of

Gammaproteobacteria and lower relative abundances of Patescibacteria than the other wells. Of these, wells 1, 2, and 21 are suspected to be under the influence of surface water intrusion, as discussed in previous sections.

Thus, there appears to be some overlap between (i) wells with a different dominant taxon to most other wells (wells 2, 12, 15, 18, 28), (ii) wells with a different overall bacterial composition to most other wells (wells 1, 2, 12, 15, 21), and (iii) wells with indications of surface water intrusion in stable water isotope data (wells 1, 2, 18, 20, 21). This suggests that surface water intrusion may play some role in shaping the groundwater bacterial communities at several wells, either by (i) changing the environmental conditions in the aquifer, (ii) direct influx of surface water microbes, or (iii) some combination of both (Fillinger et al., 2021).

It must be noted, however, that surface water intrusion is unlikely to be the sole influencing factor on these wells, as indications of surface water intrusion were not detected in the stable water isotopes from wells 12, 15 and 28. Regrettably, surface water microbial communities were not sequenced in this initial snapshot study, meaning that direct influx of surface water microbes into the aquifer could not be assessed. In addition, alternative pathways of microbial well contamination, such as (i) direct ingress at the wellhead, (ii) infiltration (including rapid preferential flow), and (iii) annular flow (i.e. flow occurring through the well annulus, either originating from the ground surface or the subsurface) (P. Hynds et al., 2014), were not considered; and neither were the configurations and intactness of the well structures, nor the strainer depths. Thus, there are major limitations on the conclusions that can be drawn. Five of the wells discussed in this section (wells 2, 12, 15, 20 and 21) are dug wells, whereas two are tube wells (wells 1 and 28). Based on data collected for wells 1, 2, 12, 15, 20 and 21, it seems likely that well depth is <10 m at these sites, and groundwater depth is ≤3 m. (Well depth and groundwater depth for well 28 were not recorded.) Overall, however, the collected metadata appear insufficient for fully unpacking the influence of surface water intrusion on bacterial communities at these wells.

Geogenic influences on well water quality also appeared to affect the microbes found in water samples. Namely, some bacterial taxa with certain metabolic lifestyles were found at higher relative abundance in wells with suitable physicochemical properties. For example, well 1, which had the highest Fe, also had one of the highest relative abundances of the *Gallionella* genus of iron-oxidizing bacteria in 16S libraries from DNA (30%) and cDNA (6%). And well 25, which had the highest SO_4^{2-}, also had the highest relative abundances of the

Beggiatoaceae family of sulphur-oxidizing bacteria in 16S libraries from DNA (1%) and cDNA (11%).

As evidence of agreement between qPCR and 16S rRNA amplicon sequencing data, well 2, which had the highest indication of Bacteroidales in qPCR data (8 and 260 GC/100 mL of the GenBac3 MST marker in DNA and cDNA extracts respectively), also had the highest relative abundances of Bacteroidetes – the phylum containing the order Bacteroidales – in 16S libraries from DNA (51%) and cDNA (35%).

None of the health-related taxa screened for in this study (e.g. *E. coli*, *C. perfringens*, *Legionella*, *Klebsiella* etc.) were detected in the 16S data, though that does not entirely prove their absence. It is also worth noting that due to the compositional nature of 16S data, the absolute abundances of all identified groups discussed above remain unknown (Gloor et al., 2017).

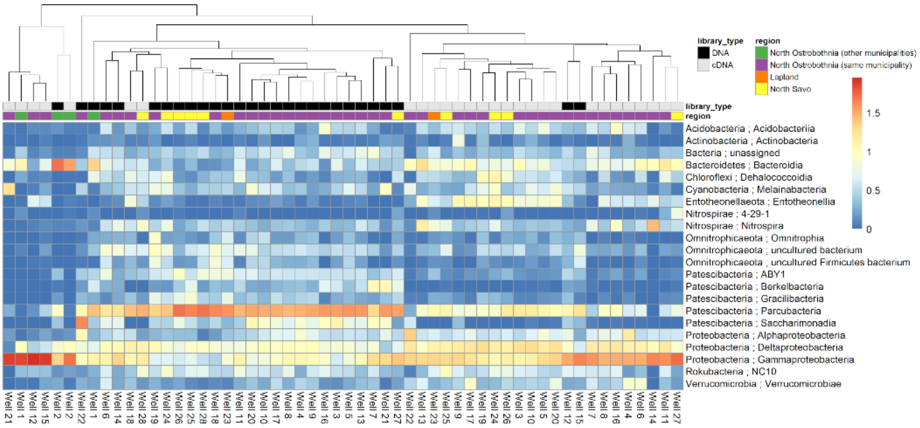

Fig. 6. Heatmap showing differences in bacterial communities based on taxonomic classifications of DNA- and cDNA-derived 16S amplicons generated in QIIME 2™ using the SSU SILVA 132 majority taxonomy. The heatmap was generated in R (using the 'pheatmap' package) (Kolde, 2019) from the log-transformed relative abundance values of bacterial classes which had a relative abundance of 5% or more in at least one library. Columns were clustered using average linkage hierarchical clustering based on the Bray–Curtis dissimilarity matrix of the data set (using the 'vegan' R package) (Oksanen, 2020) (Reprinted under CC BY 4.0 license from Publication I © 2021 Authors).

3.1.6 Spearman rank-based correlations

Indications of surface water intrusion associated with lower alpha diversity

For the DNA-derived 16S data, d-excess had positive correlations with Pielou's evenness ($r_s = 0.57$, $p = 0.001$), Shannon's diversity ($r_s = 0.51$, $p = 0.006$), and Faith's phylogenetic diversity ($r_s = 0.38$, $p = 0.044$) (Fig. 7). Lower d-excess values in groundwater wells indicate greater deviation from the meteoric water line, and therefore a higher likelihood of surface water intrusion (J. R. Gat & Gonfiantini, 1981). Thus, indications of intrusion (i.e. lower d-excess values) appeared to be weakly to moderately associated with lower alpha diversity in the wells (i.e. Pielou's evenness, Shannon's diversity and Faith's phylogenetic diversity). These correlations agree with findings from previous work showing that microbial communities in groundwater react sensitively to surface water intrusion in the context of riverbank filtration, with losses of resident taxa indicated by declining alpha diversity (Fillinger et al., 2021).

Redox potential had positive correlations with all four alpha diversity metrics (r_s range: 0.42 to 0.52). In other words, more reducing conditions appeared to be somewhat associated with lower alpha diversity. This may point to surface water intrusion again, as intrusion of surface water into the nearby subsurface in riverbank filtration systems has previously been shown to produce reducing conditions (Bourg & Bertin, 1993; Tufenkji et al., 2002) as well as lower alpha diversity (Fillinger et al., 2021), whereas microbial communities in undisturbed groundwater are often highly diverse (Karwautz & Griebler, 2022).

For the cDNA libraries, turbidity, heterotrophs, and Bacteroidales DNA (GenBac3) had moderate negative correlations with all four alpha diversity metrics (r_s ranges: −0.44 to −0.40; −0.40 to −0.54; and −0.45 to −0.41, respectively). These are also parameters that might be expected to increase as a result of surface water intrusion, although other explanations are also possible.

Microbial communities catalyse important biogeochemical processes in groundwater, such as the turnover of carbon and other nutrients, and also play a role in pollutant attenuation (Griebler & Lueders, 2009; Long et al., 2016). Thus, disturbances and changes to their community composition via surface water intrusion have important implications for drinking water quality and safety, and thereby for the proper management of groundwater wells.

Overall, however, a correlation-based approach based only on a small number of well water samples can only provide rather simplistic and limited interpretations regarding the potential effects of surface water intrusion on groundwater microbial communities. A more comprehensive, multivariate approach involving well water, surface water and aquifer material samples, as well as hydrogeological measurements and modelling, would likely yield more conclusive results.

Possible influences from road salt and agriculture

‘Total length of nearby roads’ was positively correlated with N ($r_s = 0.48$, $p = 0.009$), Na ($r_s = 0.47$, $p = 0.012$) and Cl^- ($r_s = 0.46$, $p = 0.014$), the first of which may originate from vehicle emissions and the latter two from road salt (Bettez et al., 2013; Salminen et al., 2011). ‘Total nearby field area’ was positively correlated with Na ($r_s = 0.54$, $p = 0.003$), UV_{254} absorbance of unfiltered water ($r_s = 0.54$, $p = 0.003$), N ($r_s = 0.52$, $p = 0.004$), Cl^- ($r_s = 0.50$, $p = 0.007$), K ($r_s = 0.48$, $p = 0.010$), NO_2^--N ($r_s = 0.42$, $p = 0.027$), and $(NO_3^- + NO_2^-)$-N ($r_s = 0.40$, $p = 0.034$), many of which may be linked to fertilizer use (Chiy & Phillips, 2002; Huhtanen et al., 2000; Scherer, 2005).

The same limitations of the correlation-based approach apply here as for the possible associations between surface water intrusion and alpha diversity metrics discussed above. A more thorough, systematic and standardised survey of nearby risks should consider factors such as which of the roads near each well are salted and which are not, and also assess the primary usage types of nearby fields (e.g. agricultural grazing, crop growing, etc.).

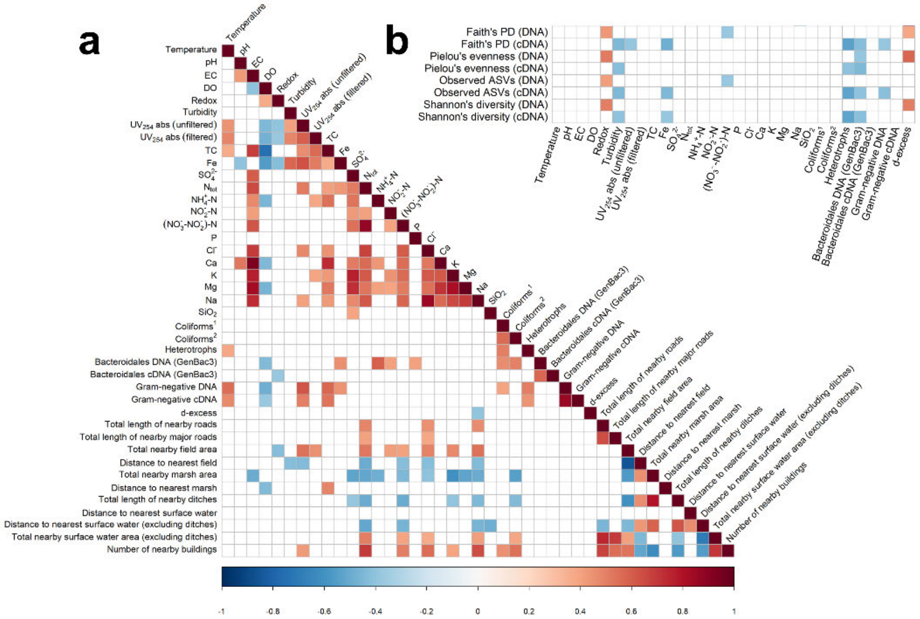

Fig. 7. Correlograms showing (a) correlations between physicochemical data, microbiological data, and environmental data and (b) correlations between physicochemical data and alpha diversity metrics. Both correlograms were constructed using a Spearman rank-based correlation coefficient matrix and associated p-values. Only the statistically significant correlations ($p < 0.05$) are shown. Red colours are positive correlations. Blue colours are negative correlations. In each case, the intensity of the colour indicates the strength of the correlation. Spearman rank-based correlation coefficients were calculated using the rcorr() function from the 'Hmisc' R package (Harrell, 2020), and correlograms were produced using the corrplot() function from the 'corrplot' R package (Wei, 2017). [1]SFS 3016 method, [2]ISO 9308-1 method (Reprinted under CC BY 4.0 license from Publication I © 2021 Authors).

3.1.7 Summary

This study revealed the physicochemical and microbiological characteristics of water from twenty-eight shallow groundwater wells in northern and central Finland. The findings highlighted several potentially problematic wells on the basis of physicochemical data, stable water isotopes, microbiological indicators, and 16S rRNA amplicon libraries. For example, well 1 had turbidity and Fe values exceeding the target levels outlined in Finnish law; wells 1, 2, 18, 20, and 21 had indications of potential surface water intrusion in stable water isotope data; and well 2 had relatively high levels of several microbiological indicators.

Possible influences on well water quality from surface water intrusion, road salt, and agriculture were discussed with the aid of the collected data and Spearman rank-based correlations. This discussion included a possible association between indications of surface water intrusion and reduced alpha diversity in groundwater bacterial communities – an observation which would be in agreement with previous work on a riverbank filtration system (Fillinger et al., 2021). However, given the rather small sample size (n = 28) and rather simplistic correlation-based analysis presented here, these findings are largely preliminary in nature, and further work would be needed to confirm the associations and suspected site-specific influences and trace them unequivocally to their sources. A more thorough, systematic and standardised survey of nearby environmental risks would go a long way towards increasing certainty about the sources of influencing factors, and towards providing a more secure basis for site-specific risk mitigation. Future work should also involve more detailed consideration of the hydrogeological setting and well characteristics, including local hydraulic gradients, well depth, casing type, structural integrity, strainer depth etc.

Although there have been many previous studies of microbial communities in the deep groundwater of Finland (Bell et al., 2018; Bomberg et al., 2015, 2016; Itävaara et al., 2011; Kutvonen et al., 2015; Miettinen et al., 2015; Nyyssönen et al., 2012; Pedersen et al., 2014; Purkamo et al., 2013, 2018, 2020), the composition and diversity of communities in shallow groundwater have remained largely unexplored. This was likely the first major study to observe these communities in Finland (Länsivaara, 2020; Tiwari, 2020; Tiwari et al., 2021), and one of the first such studies in the Nordic countries (Knobloch et al., 2021).

This study showed that snapshot sampling of stable water isotopes from well water can be used to identify wells which may be under the influence of surface water intrusion. The downside of the snapshot sampling approach was that temporal variations in physicochemical and microbiological characteristics of the well water were not considered. Another limitation of the study was that the physicochemical and microbiological characteristics of nearby surface water bodies were not assessed. The following studies described in this thesis address these two issues and further explore the question of surface water intrusion – to better assess its risk to well water quality and safety, and to better understand its potential influence on groundwater microbial communities.

3.2 1-year study of temporal variation with real-time online monitoring (II)

3.2.1 *Periodic manual sampling and measurements*

Physicochemical and hydrological parameters

Much of the overall physicochemistry of the four wells studied here was rather typical of Finnish groundwaters (Isomäki et al., 2008). However, some extreme physicochemical values and unexpected patterns of variation were observed throughout the monitoring period (Table 8). In periodic manual measurements (i.e. on-site and lab-based measurements), well W had the lowest median EC (27 µS/cm) and the highest median ORP (357 mV); well Y had the highest median temperature (8.7 °C), turbidity (11.9 NTU), UV_{254} absorbance (1.650), Fe (2.76 mg/L) and Mn (0.15 mg/L), and the lowest median DO (3 mg/L); and well Z had the highest median EC (381 µS/cm) and pH (7.1). The median turbidity, Fe, and Mn of well Y exceed the target levels outlined in Finnish law (1 NTU, 0.2 mg/L, and 0.05 mg/L respectively). Unpublished data from the North Savo Centre for Economic Development, Transport and the Environment suggest that high EC values in well Z may be due to high Cl^- in the groundwater (although Cl^- was not specifically measured here).

Table 8. Summary of manual physicochemical measurements for wells W, X, Y and Z.

	Median	Max	Min
Well W			
Temperature (°C)	6.3	6.5	5.1
pH	6.2	6.2	5.5
EC (µS/cm)	27	40	25
DO (mg/L)	9.5	12.7	5.8
ORP (mV)	357	530	212
Turbidity (NTU)	0.03	0.19	0.01
UV_{254} absorbance	0.047	0.181	0.037
Fe (mg/L)	<0.02	0.03	<0.02
Mn (mg/L)	<0.006	0.010	<0.006
Well X			
Temperature (°C)	5.8	6.1	5.4
pH	6.6	6.9	6.5
EC (µS/cm)	90	235	32

	Median	Max	Min
DO (mg/L)	13.4	16.5	10
ORP (mV)	266	420	235
Turbidity (NTU)	0.08	0.51	0.01
UV_{254} absorbance	0.021	0.055	0.017
Fe (mg/L)	<0.02	0.05	<0.02
Mn (mg/L)	<0.006	0.007	<0.006
Well Y			
Temperature (°C)	8.7	10.2	6.9
pH	6.2	6.4	6.1
EC (µS/cm)	93	101	82
DO (mg/L)	3.0	8.6	1.0
ORP (mV)	96	110	82
Turbidity (NTU)	11.90	26.26	5.64
UV_{254} absorbance	1.650	1.820	1.416
Fe (mg/L)	2.76	3.70	2.45
Mn (mg/L)	0.150	0.180	0.141
Well Z			
Temperature (°C)	6.7	6.9	5.9
pH	7.1	7.4	6.6
EC (µS/cm)	381	467	364
DO (mg/L)	14.12	14.51	12.86
ORP (mV)	NA	NA	NA
Turbidity (NTU)	0.05	0.32	0.01
UV_{254} absorbance	0.060	0.084	0.054
Fe (mg/L)	<0.02	0.02	<0.02
Mn (mg/L)	0.013	0.017	<0.006

EC = electrical conductivity, DO = dissolved oxygen, ORP = redox potential, Fe = iron, Mn = manganese; NA = this parameter was not measured at all for the well in question

Microbiological water quality indicators

Heterotrophic bacteria, coliform bacteria (including *E. coli*), somatic coliphages, and spores of sulphite-reducing clostridia (SSRC) were measured during eight monthly sampling timepoints (March–November 2019; excluding July) (Fig. 8). *E. coli*, coliphages, and SSRC are potential indicators of faecal contamination (Marcheggiani et al., 2008; Roslev et al., 2004; Singh et al., 2022).

E. coli bacteria were not detected in any groundwater samples, the Colilert-18® method applied to samples from well Z also returned non-detects at every timepoint, and somatic coliphages were detected only once in the study (1 PFU/1.1 L, well Y, May). Low levels of SSRC (1 CFU/ L) were detected in well X in May, and higher levels in well Y at almost all timepoints, the highest being in August and September (both 20 CFU/ L).

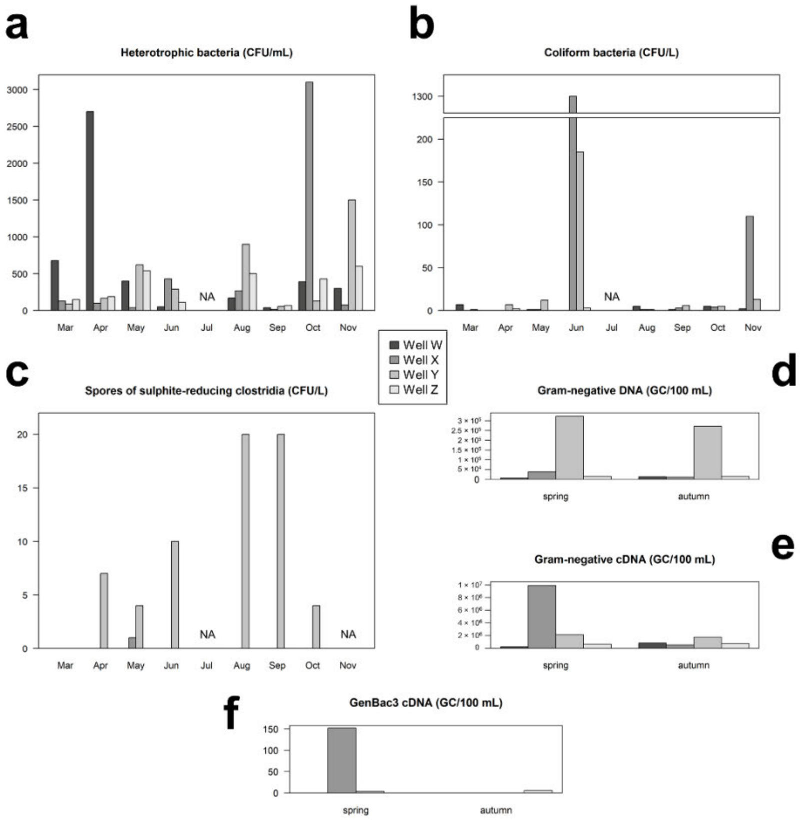

Fig. 8. Microbiological water quality indicators in untreated groundwater samples. a: Counts of heterotrophic bacteria (CFU/mL), determined by culturing on Reasoner's 2A (R2A) agar at 22 °C for 7 days. No samples were taken in July. b: Counts of coliform bacteria (CFU/L). Both SFS 3016 and ISO 9308-1 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 (Reprinted under CC BY 4.0 license from Publication II © 2022 Authors).

3.2.2 Real-time online monitoring

Springtime snowmelt is the main source of groundwater recharge

At all wells, the annual peak in groundwater level was observed soon after the springtime snowmelt period, indicating snowmelt to be the primary source of recharge in these wells (Fig. 9a), as is often the case in Finland (Soveri, 1981). Time delays between snow depth reaching zero and the annual peak in groundwater level were about 3–7 weeks. Well W exhibited several noticeable groundwater level peaks after heavy rains in summer, and another peak in autumn (presumably indicating groundwater recharge from rain or snow events near the beginning of another period of snow accumulation). Spring and autumn recharge are typical of many Finnish groundwater formations (Soveri, 1985).

Operational and maintenance-related variation

The groundwater level in well X is shown in Fig. 9b. This well is 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. 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 from late April–early June. However, pumping was continued at 10% of the typical rate (without supplying the water to users) 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 X was on a daily on/off cycle, causing predictable daily fluctuations of groundwater level.

The electrical conductivity (EC) and turbidity in well Y are shown in Fig. 9c. Although the real-time online monitoring largely proceeded in a satisfactory manner in all wells, the high Fe in well Y (median of 2.76 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 Y, an Fe precipitate accumulated around the sensor during the monitoring period, leading to the occurrence, at times, of extremely high turbidity readings which likely 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 measurements. 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 Fe precipitate (Fig. 9c). 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). These phenomena are described here as examples of potential challenges encountered during real-time monitoring, and potential sources of artificial variation.

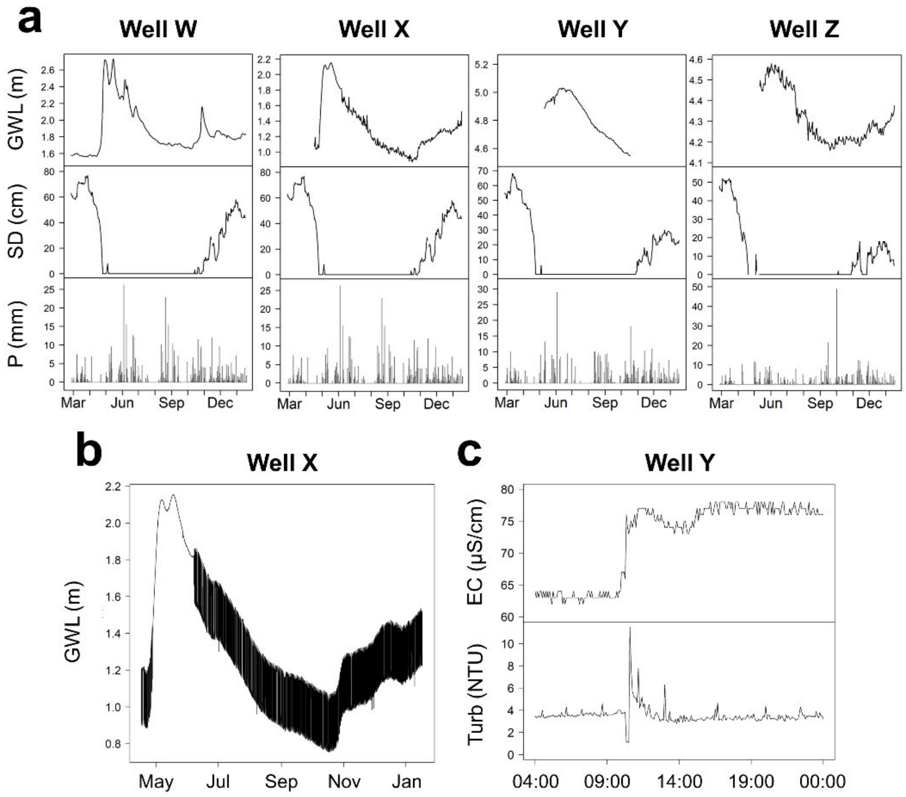

Fig. 9. True and artificial variation in the real-time online monitoring data. a: Daily data for groundwater level, snow depth and precipitation at the four studied wells. Snow depth and precipitation data are derived from the nearest monitoring station of the Finnish Meteorological Institute (FMI). As wells W and X are in the same inland municipality, a single FMI monitoring 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 Y 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 X. Normal pumping was halted between late April and early June. At other times, pumping followed 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 Y from 17 October 2019. Pipe flushing increased the EC and caused a momentary increase in the turbidity measurements (Reprinted under CC BY 4.0 license from Publication II © 2022 Authors).

3.2.3 Stable water isotope analyses

For well Y, clear signs of evaporation (i.e. enrichment of the heavy isotopes ^{2}H and ^{18}O) were observed in the isotopic compositions of most of the groundwater samples (Fig. 10), including all samples taken between late summer and autumn (late August–late November 2019). The greater alignment of the samples with the Rokua LEL than with the Oulanka LMWL suggests that surface water which has been subjected to evaporation has entered the well. Such strong isotopic enrichment was not observed in samples from any of the three other wells. Nevertheless, isotope compositions from these other three sites are discussed briefly below.

For well W, the isotopic compositions of well water samples varied along the LMWL, thereby showing no obvious signs of surface water intrusion. In addition, however, the isotopic compositions of most samples collected from surface waters near well W – ditches, a small stream, and a river – also varied along the LMWL; the two surface water samples with the greatest signs of evaporation were collected in late September (δ^{2}H = −89.41 and δ^{18}O = −12.19) and early October (δ^{2}H = −90.38 and δ^{18}O = −12.33). Because most of the surface water samples from this site plotted along or near the LMWL, the possibility of surface water intrusion at well W could be neither confirmed nor refuted by the stable water isotope analyses. Further investigation using additional methods (e.g. measurements of physicochemical parameters such as EC and temperature from both surface waters and well water, hydraulic gradient determinations, and rainfall–runoff modelling) may be necessary to assess the influence of surface water intrusion at this site.

For well X, the isotopic compositions of most groundwater samples varied along the LMWL, as at well W. Two well water samples with slightly greater enrichment were collected in late March (δ^{2}H = −100.19 and δ^{18}O = −13.61) and late October (δ^{2}H = −98.33 and δ^{18}O = −13.43); however, these indications of enrichment were small in comparison to the enrichment seen in well Y. The isotopic compositions of two of the surface water samples followed the LMWL, but at least one other deviated somewhat, along the Posio LEL. Overall, it was not possible to confirm nor refute the possibility of surface water intrusion at well X using isotopes, because during springtime flooding (when surface water intrusion into the well may be most likely) the isotopic compositions of the groundwater and the nearby river both appear to be influenced by snowmelt. Hence, similarly to well W, further investigation and possibly other methods may be necessary to assess the influence of surface water intrusion at this site.

For well Z, the isotopic compositions of groundwater samples varied along the LMWL (with less spread than other wells), and no clear signs of evaporation were observed. Prior to this study, the groundwater quality in well Z was thought to be influenced by a nearby lake. Although surface water samples were not collected from the lake during this study, isotopic compositions of lake samples taken after the study, in August 2021, exhibited strong signs of evaporation (unpublished data from the North Savo Environmental Centre), meaning that stable water isotope analysis should be a good method to detect possible surface water intrusion at this site. Thus, based on this 2021 finding, and results presented here, it would appear that surface water intrusion is largely absent at this well. However, future work with simultaneous sampling of isotopes from surface water and groundwater would be needed to confirm this.

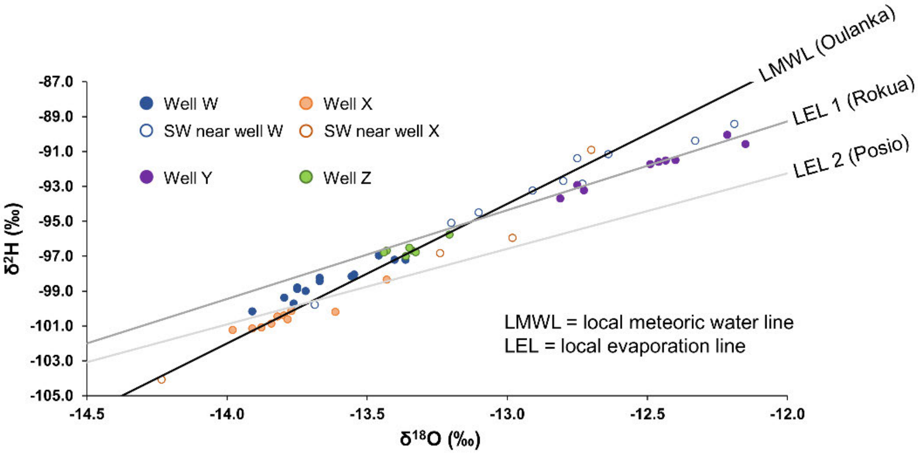

Fig. 10. Stable water isotope results from the wells compared to surface water samples from well W and X surroundings and to rainfall. Data for Oulanka local meteoric water line (LMWL) and Rokua and Posio local evaporation lines (LEL) were taken from previous studies (Isokangas et al., 2015; Nora et al., 2019; Rossi et al., 2015). Samples near LMWL indicate similarity to the isotopic composition of meteoric water, and samples near the LELs indicate similarity to the isotopic composition of surface water that has been subjected to evaporation (i.e. enriched in the heavy isotopes ^{2}H and ^{18}O). Abbreviation: SW = surface water (Reprinted under CC BY 4.0 license from Publication II © 2022 Authors).

3.2.4 16S rRNA amplicon sequencing data

High-level observations

Overall, at the phylum level, DNA-derived 16S rRNA amplicons had high median relative abundances of Patescibacteria (40%), whereas cDNA-derived amplicons had high relative abundances of Proteobacteria (30%); these comprised mostly Parcubacteria (29%) and Gammaproteobacteria (26%) at the class level, respectively. This matches with findings from the snapshot study described above.

Alpha diversity metrics were calculated based on the 16S data: namely, (i) observed ASVs, (ii) Faith's phylogenetic diversity, (iii) Pielou's evenness, and (iv) Shannon's diversity index. Values for well Y (which appeared to have indications of surface water intrusion in stable water isotope data) were generally lower than for the other wells, especially when considering cDNA libraries (Table 9).

Table 9. Alpha diversity metrics for wells W, X, Y and Z.

Well	Observed ASVs		Faith's PD		Pielou's evenness		Shannon	
	DNA	cDNA	DNA	cDNA	DNA	cDNA	DNA	cDNA
Well W								
Spring	1515	1459	137	126	0.87	0.85	9.19	8.90
Autumn 1	1681	1611	153	135	0.89	0.89	9.54	9.46
Autumn 2	1734	1641	155	137	0.89	0.89	9.60	9.48
Well X								
Spring	2039	1075	178	114	0.89	0.67	9.74	6.77
Autumn	1723	1181	165	116	0.90	0.79	9.65	8.10
Well Y								
Spring	1251	672	140	75	0.78	0.62	8.01	5.82
Autumn	1617	618	162	72	0.75	0.49	7.95	4.55
Well Z								
Spring	NA	NA	NA	NA	NA	NA	NA	NA
Autumn	1561	1791	146	152	0.90	0.88	9.52	9.51
Max	2039	1791	178	152	0.90	0.89	9.74	9.51
Min	1251	618	137	72	0.75	0.49	7.95	4.55

Specific taxa of interest

For well W, phylum-level community composition was quite similar for all DNA samples regardless of season, and the same was true of all cDNA samples (Fig. 11).

In well X, however, the springtime samples had relatively high relative abundances of chloroplast sequences from the freshwater microalgal genus *Epipyxis* in DNA (8%) and cDNA (27%), which were responsible for the bulk of the Cyanobacteria bars shown in Fig. 11. The well X springtime samples also had chloroplast sequences from the freshwater algal genus *Neotessella* in DNA (5%) and cDNA (3%). Algae are known surface water indicators (Gollnitz et al., 2003; Moulton-Hancock et al., 2000; Robertson & Edberg, 1997; Vasconcelos & Harris, 1992), and are not often found in groundwater as they typically require light for photosynthesis. However, well X is located about 60 m from a river, and there is a known risk of surface water flowing towards the well in the springtime each year. One possible route for algae entering the groundwater could be slow overland flow of river water followed by preferential flow through the subsurface to the groundwater table. Whatever the route of entry, the detection of algae or chloroplast sequences in the groundwater could be a useful indicator of surface water intrusion at this site in springtime. Support for this idea comes from a study showing 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 a very similar climate to Michigan, so it is not entirely unreasonable to imagine that the relative abundance of chloroplast sequences in some Finnish lakes and rivers might also reach a peak in springtime, around the time of ice melt. High levels of chloroplast sequences in surface water around the time of the springtime ice melt, coupled with intrusion of surface water into groundwater, could make detection of chloroplast sequences in nearby groundwaters more likely during springtime. Regrettably, this hypothesis could not be properly evaluated in this study as surface water samples were not collected. 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 previous studies involved 16S rRNA gene amplicon sequencing and used 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 might be better 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). The springtime cDNA from well X also had moderate to high relative abundances of the Gram-negative bacterial genera *Polaromonas* (13%), *Polynucleobacter* (6%), and *Zoogloea* (5%); and the autumntime cDNA had a high relative abundance of the nitrite-oxidizing bacterial genus, *Nitrospira* (11%).

In well Y, springtime and autumntime cDNA-derived libraries were dominated very strongly by Proteobacteria (spring: 75%, autumn: 80%). Springtime DNA had relatively high levels of the class Clostridia (23%), of which a considerable portion was the *Blautia* genus (11%), a mammalian gut bacterium which may indicate faecal contamination (Eren et al., 2015; McLellan & Eren, 2014). Members of the class Clostridia are also anaerobic, so their presence in largely aerobic groundwater (the median DO for well Y was 3 mg/L) may indicate recent intrusion. However, the DO for well Y also sometimes dropped below 1, which may indicate that anaerobic conditions sometimes prevail in this well. All amplicon libraries from well Y, both from DNA and cDNA, had high relative abundances of the *Gallionella* genus of iron-oxidizing bacteria (range: 17–73%). This was not entirely surprising as this well also had low DO, high turbidity, and high Fe and Mn, 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 & De Vet, 2015). A recent study found that *Gallionella* correlated with non-operational groundwater wells in eastern Russia (Braun et al., 2016). This corresponds rather well with the history of well Y 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).

In well Z, autumntime DNA had a high relative abundance of the Parcubacteria superphylum (32%), which 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 Z, and thus a seasonal comparison could not be made.

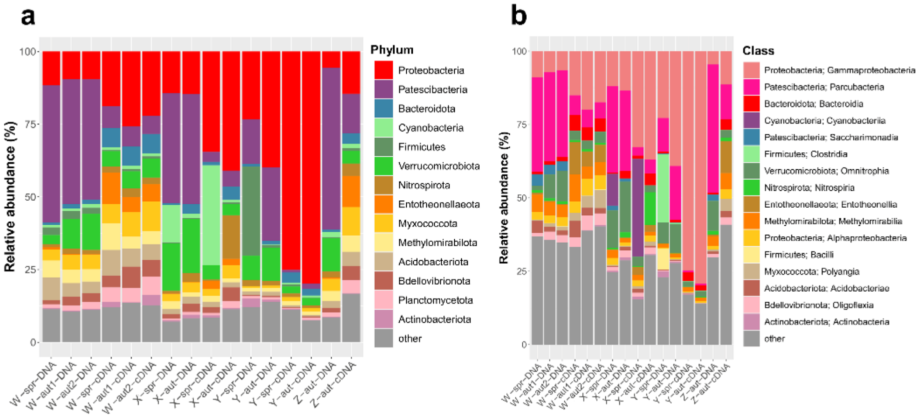

Fig. 11. Phylum- (a) and class-level (b) taxonomic classifications for 16S rRNA amplicon sequencing data. Only the most abundant microbial taxa are shown (those which were present at ≥5% in at least one true sample or control sample). Abbreviations: W, X, Y, Z = well W, well X, well Y, well Z; spr = 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 (Reprinted under CC BY 4.0 license from Publication II © 2022 Authors).

3.2.5 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 & Kennedy, 2020; Oppus, Guico, 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 & Kennedy, 2020; J.-Y. 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 (perhaps seasonal) 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 (D. R. Clark et al., 2018). These methods provide a big-picture overview of the communities and their diversity, and 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, flow cytometry, or other methods could also be applied to measure absolute abundances (Converse et al., 2009; Haugland et al., 2010; Jian et al., 2020; Kärkkäinen et al., 2010; Pitkänen et al., 2013; Props et al., 2017). Water suppliers could avail of these methods via private companies, 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.

3.2.6 Study limitations and recommendations

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 early warning system (EWS). Hence, researchers should 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) ensuring 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, and (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 W and X (not Y and Z; although a lake near well Z was sampled for isotopes by the North Savo Environmental Centre in August 2021), and (ii) the surface water sources near wells W and X 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 Z, 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.

Fourthly, many useful parameters for investigating the potential influence of surface water intrusion on nearby groundwater wells were not assessed in great detail here; for example, groundwater and well depths, configurations and intactness of the wells, hydraulic gradients, subsoil types and thicknesses. Future work should consider these to obtain more conclusive evidence about whether or not intrusion is, in fact, occurring, and to better understand the spatiotemporal nature of the interactions between surface water and groundwater at the studied sites.

Finally, given that threats to groundwater quality can vary in a site-specific manner (as shown in this and many other studies) future groundwater quality monitoring programmes should identify and focus on parameters most likely to prove useful in detecting potential site-specific water quality problems.

3.2.7 Summary

Real-time online monitoring measurements demonstrated potential usefulness for monitoring physicochemical well water quality, though more frequent sensor calibration and cleaning may improve reliability, and the potential costs and benefits of applying the method should be considered before implementation. Stable water isotope analyses identified potential indications of surface water intrusion at one of the studied wells but were less informative at the other sites. 16S rRNA amplicon sequencing data was used to identify potential influences on the wells from surface water intrusion and faecal contamination.

Overall, well Y had (i) the most problematic physicochemical water quality, including high median turbidity, Fe, and Mn values which exceeded the target levels outlined in Finnish law; (ii) positive detections of somatic coliphages and SSRC, which could indicate faecal and/or viral pollution; (iii) potential indications of surface water intrusion in stable water isotope data; (iv) generally lower alpha diversity metrics compared to other wells; and (v) high relative abundance of *Gallionella* (a genus of iron-oxidising bacteria) and *Blautia* (a genus of mammalian gut bacteria) in 16S data, suggesting potential risks to water quality from high Fe concentrations and faecal contamination, respectively.

Well X had (i) detection of SSRC (at low levels) in May; and (ii) high relative abundances of chloroplast sequences from the freshwater algal genera *Epipyxis* and *Neotessella* during springtime, which may indicate surface water intrusion during the snowmelt period.

The characteristics of wells W and Z were relatively unremarkable, and no indications of surface water intrusion or faecal contamination were detected.

3.3 1-year study of surface water intrusion impacts (III)

3.3.1 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. 12a). In Area B, samples from the suspect well overlapped completely with samples from the nearby gravel pit pond, suggesting strong physicochemical similarity (Fig. 12b). And in Area C, samples from the suspect well grouped somewhat closely to samples from the nearby lake, suggesting moderate physicochemical similarity (Fig. 12c). It must be noted, however, that the PCA analysis was conducted independently for each study area, meaning that the principal components in each plot are not identical (despite identical input parameters). Thus, direct comparisons between study areas should be treated with some caution.

The lack of physicochemical similarity between suspect well A and the nearby river was somewhat surprising, as (i) this well had the shortest distance to surface water of the three suspect wells (~60 m vs. ~180 m and ~220 m), and (ii) 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, suggesting increased subsurface flow of river water towards the well (Hunt et al., 2005; Winter et al., 1998), and presumably therefore an increased risk of changes in well water quality. Nevertheless, no obvious fluctuations of physicochemical parameters 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 in February 2017 found logical reductions in well water temperature, EC, and the concentrations of major ions; and increases 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 pumping rate. This may have prevented any major effects of river water intrusion on well water physicochemistry during the springtime snowmelt period in May. The well is not usually pumped at all during late April–early June due to the potential risks to well water quality. During this study, however, pumping was continued at 10% of the typical rate (without supplying the water to users) so that the interaction between surface water and groundwater could be examined. Another factor that could play an important role is the hydraulic gradient between river and well, something which was not determined during this study.

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. 12b and 12c). 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. 12a). However, isotope data (described in the next section) do not support this idea, hence the importance of using a multi-method approach.

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. 12a, 12b and 12c). Samples from suspect well B also exhibited more variation than any of the other wells (as indicated by a larger ellipsis), presumably due to surface water influence (Fig. 12b).

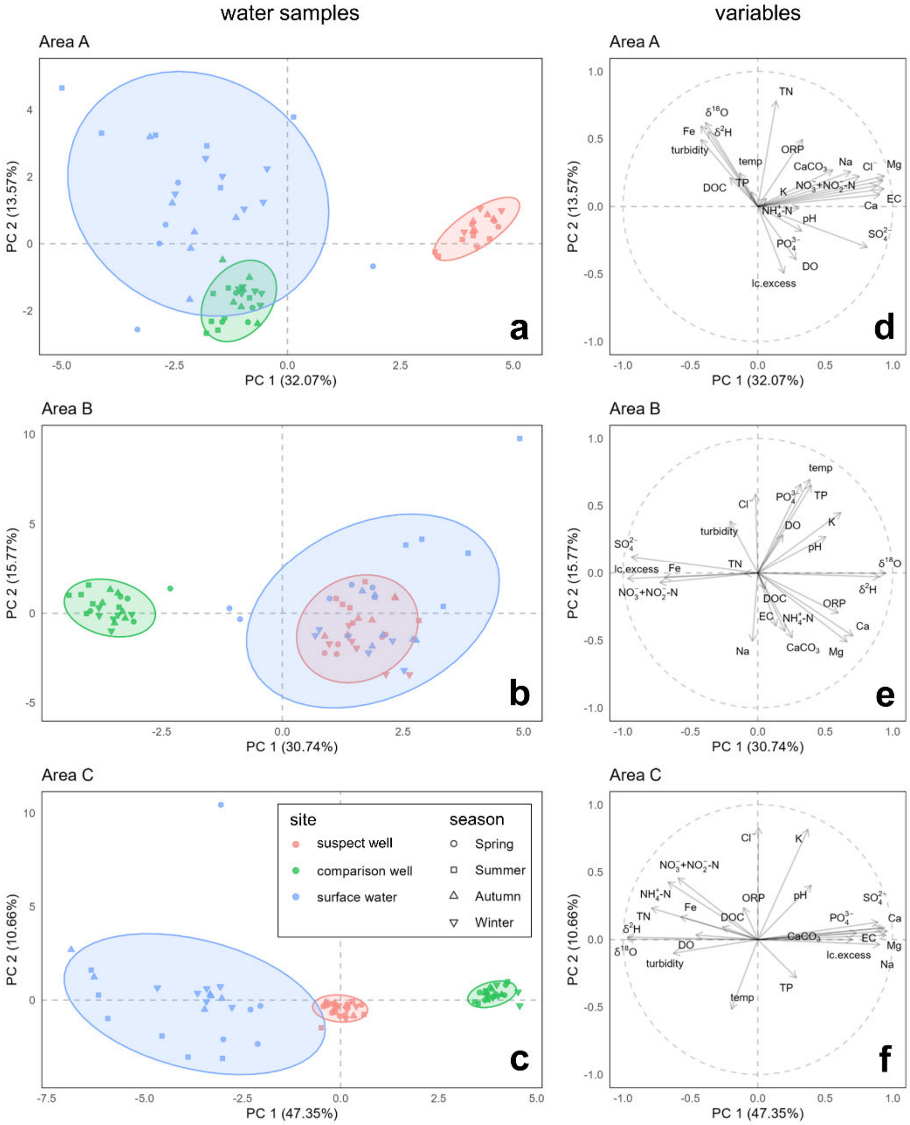

Fig. 12. Principal component analysis (PCA) plots for individuals (i.e. water samples) (a, b, c), and variables (d, e, f) from study areas A, B, and C, respectively. Seasons are winter (January–March), spring (April–June), summer (July–September) and autumn (October–December). The ellipses in plots a, b, and c enclose approximately 95% of the data points in each group.

3.3.2 Suspect wells B and C showed strong and moderate indications of surface water intrusion in isotope data

Stable water 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. 13b and 13c). Stable water isotope data from suspect well B overlapped strongly with data from the nearby gravel pit pond (Fig. 13b), whereas data from suspect well C plotted approximately midway between data from the comparison well and data from the nearby lake (Fig. 13c). In other words, suspect wells B and C showed strong and moderate indications of surface water intrusion respectively. It must be noted, however, that the terms ‘strong’ and ‘moderate’ are used loosely here, as intrusion is not a constant and consistent process, and even when indications of intrusion are detected, further investigation is needed to determine when, where and how intrusion occurs.

Data from all comparison wells and suspect well A plotted along the LMWL (Fig. 13). These wells had similar isotopic values and small standard deviations. Of all studied sites, compositions from the lake were the most enriched in ^{2}H and ^{18}O (Fig. 13c).

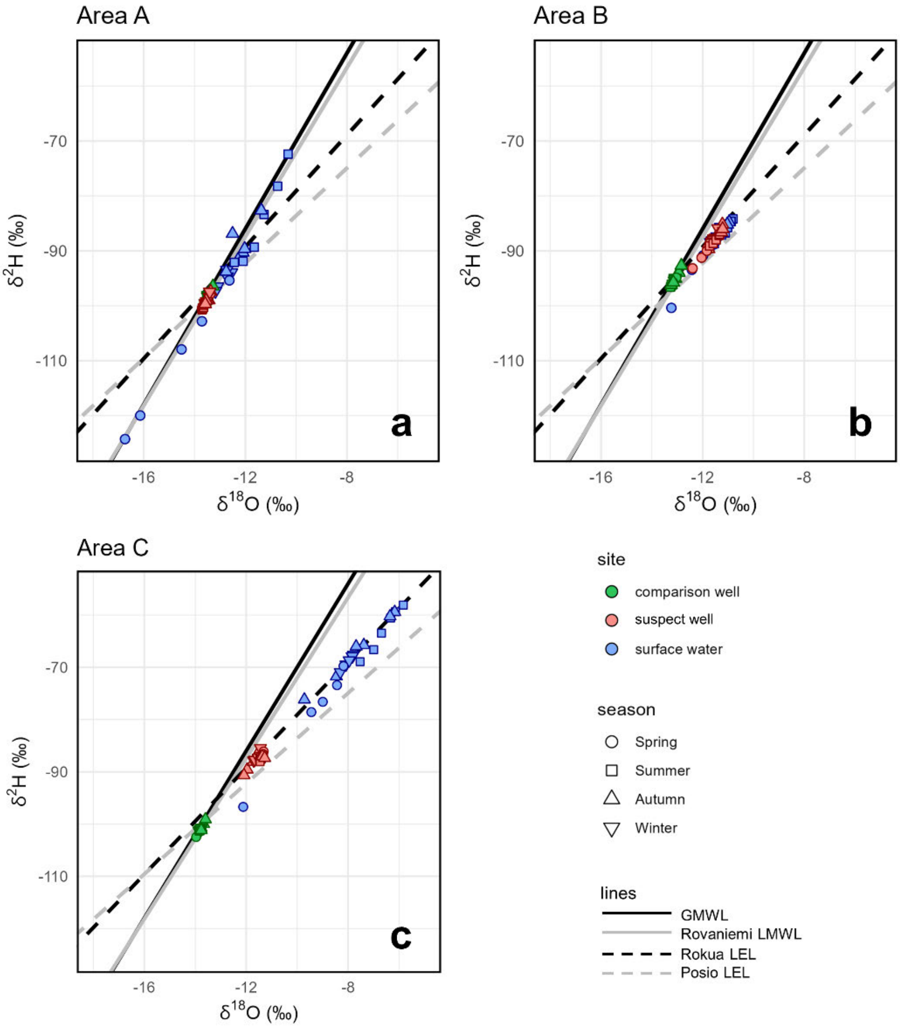

Fig. 13. The stable water isotope compositions, in dual isotope ($\delta^{18}O$–$\delta^{2}H$) space, of collected samples from wells and surface water bodies in Area A (a), Area B (b), and Area C (c). Abbreviations: GMWL = global meteoric water line; LMWL = local meteoric water line; LEL = local evaporation 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).

Values of lc-excess for precipitation and groundwater typically fluctuate around 0, whereas the lc-excess values of water bodies undergoing evaporation are negative

(Y. Liu et al., 2020). Mean lc-excess* values for suspect wells B and C were both below −3.3, providing further evidence of surface water intrusion at these wells (Table 10). Mean lc-excess* for suspect well B and the nearby gravel pit pond were strikingly similar (−3.48 and −3.99), suggesting a strong surface water influence on suspect well B. Mean lc-excess* for suspect well C and the nearby lake were less similar (−3.35 and −11.23), suggesting a more moderate surface water influence on suspect well C. 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 at these wells, despite noticeable physicochemical similarity between the river and comparison well A in the PCA plot (Fig. 12a). Of all studied sites, compositions from the lake had the lowest mean values for δ^{2}H, δ^{18}O, and lc-excess* indicating strong enrichment of ^{2}H and ^{18}O due to open-water evaporation (Table 10).

Table 10. Mean δ^{2}H, δ^{18}O, and lc-excess* values for the nine studied sites.

Area	Site	Mean δ^{2}H ± SD (‰)	Mean δ^{18}O ± SD (‰)	Mean lc-excess* ± SD
A	Suspect well A	−99.44 ± 1.00	−13.57 ± 0.11	−0.20 ± 0.21
	Comparison well A	−98.53 ± 0.60	−13.49 ± 0.06	+0.08 ± 0.19
	River	−94.04 ± 11.60	−12.70 ± 1.49	−1.19 ± 1.44
B	Suspect well B	−88.07 ± 1.85	−11.56 ± 0.29	−3.48 ± 0.47
	Comparison well B	−95.39 ± 0.95	−13.11 ± 0.11	+0.25 ± 0.16
	Gravel pit pond	−87.72 ± 3.56	−11.44 ± 0.55	−3.99 ± 0.80
C	Suspect well C	−87.45 ± 1.06	−11.51 ± 0.20	−3.35 ± 0.75
	Comparison well C	−101.00 ± 0.60	−13.80 ± 0.07	+0.04 ± 0.16
	Lake	−68.87 ± 8.29	−7.84 ± 1.40	−11.23 ± 2.48

Suspect wells B and C had higher δ^{18}O and lower lc-excess* across the entire sampling period than suspect well A (Fig. 14). lc-excess* values in these two wells were lowest in summer, most likely as a result of higher evaporation rates (and thus lower lc-excess* values) in the nearby surface waters during that time. By contrast, lc-excess* values in suspect well A did not exhibit any noticeable seasonal fluctuation.

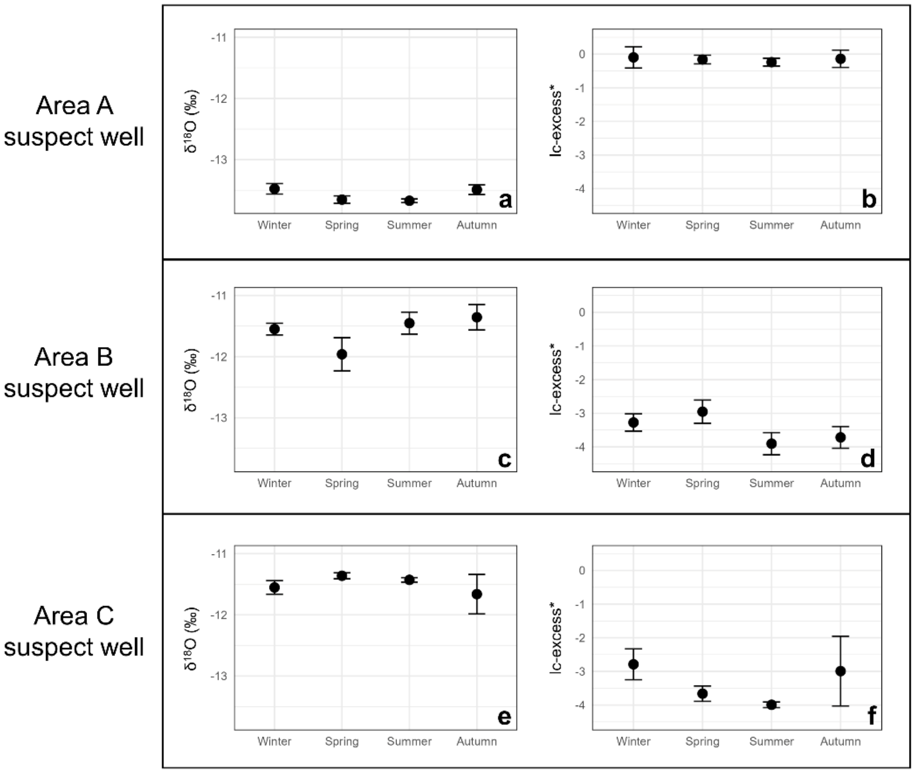

Fig. 14. Mean values for seasonal δ^{18}O and lc-excess* in 'suspect' wells from Area A (a and b), Area B (c and d) and Area C (e and f). Seasons are winter (January–March), spring (April–June), summer (July–September) and autumn (October–December). Error bars are one standard deviation.

It appeared, based on isotope mixing (Eq. 2), that about 80–95% of the water in suspect well B comes from the nearby gravel pit pond (Fig. 15a), whereas about 25–40% of the water in suspect well C comes from the nearby lake (Fig. 15b). In suspect well B, the fraction of surface water in the well ($R_{surface}$), was very similar in the winter, spring, and autumn (0.94, 0.91, and 0.92), and slightly lower in the summer (0.79) (Fig. 15a). In suspect well C, $R_{surface}$ was highest in springtime (0.39) while approximately equal in the other seasons (0.27–0.28) (Fig. 15b).

Thus, the relationship between suspect well B and the nearby gravel pit pond might be best described as a bank filtration system, whereby much of the water extracted from the well consists not of groundwater but rather of surface water from the nearby pond. 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. The interaction to the pond effectively provides extra water storage for the well, and therefore better drought resilience. In addition, given the extra water storage provided by the pond, it may be that the pumping rate of the well could be increased somewhat without fear of depleting the groundwater reserves. If pumping rate is increased, however, care should be taken to ensure that well water quality is not negatively affected by this change. A final point to consider is that since the gravel pit pond is a result of human activity rather than a natural ecosystem, increasing the pumping rate at this well may be less problematic as compared to increasing the pumping rate at a well located near a naturally occurring river or lake, which may have higher biodiversity and/or be under environmental protection.

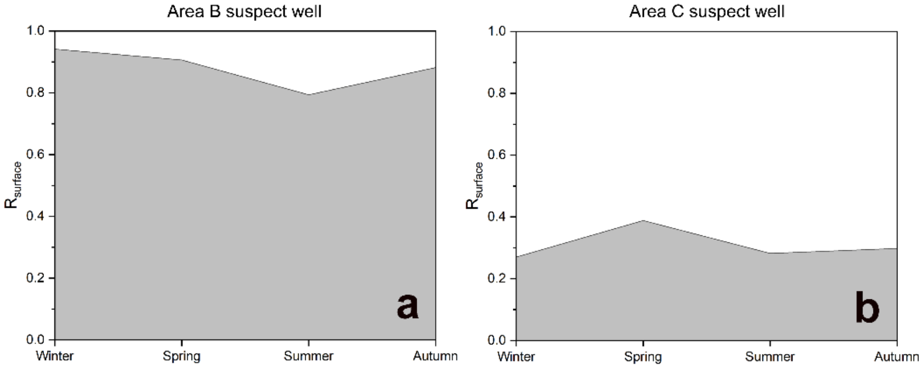

Fig. 15. The fraction of surface water in the 'B suspect' well (a) and the 'C suspect' well (b). Seasons are winter (January–March), spring (April–June), summer (July–September) and autumn (October–December).

3.3.3 *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. 16). This wave was likely caused by the intrusion of warmer water from the nearby gravel pit pond (Anderson, 2005; Foulquier, Malard, 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, the water from this well might also be said to fit the definition of 'groundwater under the direct influence of surface water' (GUI) given in USEPA regulation 40 CFR 141.2, 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 & Qi, 2000). None of the other wells exhibited similar temperature fluctuations, having temperatures mostly between 4 and 9 °C year-round. 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. 16). If the intrusion of warmer water from the pond is the cause of increased temperature in the well, then the travel time between pond and well during the summer–early autumn period may be approximately 6–10 weeks. Estimates of travel times can vary depending on the tracer used (Hunt et al., 2005). However, based on the rough estimate provided above, travel times in this instance are almost certainly on the order of weeks to months (rather than years). Assuming a travel time of about 6–10 weeks, and a distance between well and pond of about 180 m, the velocity of surface water intruding towards the well during the summer–early autumn period may be about 2–4 m/day. This estimate should be treated with caution, however, as it does not consider the nature of the aquifer material, effective porosity, hydraulic gradient, or hydraulic conductivity.

A noticeable drop in $\delta^{18}O$ and $\delta^{2}H$ values was observed in all surface waters around the time of the springtime snowmelt (during April–May) (Fig. 16). A similar drop was observed in suspect well B after a 3-week delay, suggesting a shorter travel time between pond and well during this period (possibly indicating a greater degree of overland flow as opposed to interflow). The drop in $\delta^{18}O$ and $\delta^{2}H$ values in suspect well B at this time is likely due to the interaction 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. The $\delta^{18}O$, $\delta^{2}H$, and lc-excess* values from suspect well B were overall very similar to values from the nearby gravel pit pond, suggesting a close groundwater–surface water interaction, with values being near identical from November–March and slightly less similar from May–October (probably due to open water evaporation in the pond during this period). $\delta^{18}O$, $\delta^{2}H$, and lc-excess* values were relatively stable in the other wells during the study period, except for a gradual decline in lc-excess* in suspect well C (from about −2.3 to −4.7), which may indicate long-term, moderate intrusion from the nearby lake.

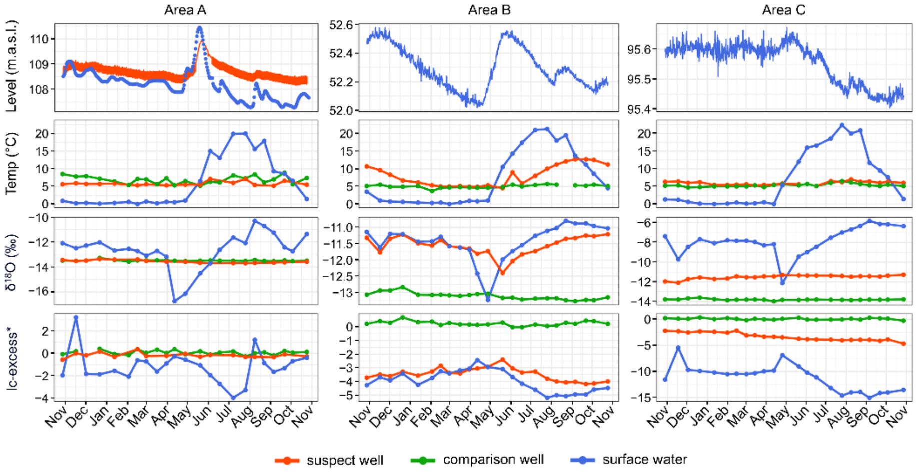

Fig. 16. Temporal variation in water level, water temperature, $\delta^{18}O$, and lc-excess* values for all sites throughout the study period. Note that, except in the case of temperature, the y-axis ranges are different for each study area (to better show the variation). No water level data were collected for wells in study areas B and C.

3.3.4 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 (described above), prokaryotic communities in the well water samples had significantly higher values of all alpha diversity metrics than communities from the nearby gravel pit pond ($p < 0.05$) (Fig. 17). This might be considered surprising, given that much of the water from this well likely comes from surface water, and the prokaryotic communities of surface waters tend to have lower alpha diversity than those of groundwater (Fillinger et al., 2021; Ji, Wang, et al., 2022; Ji, Zhang, et al., 2022). However, a recent study of a riverbank filtration system in a sand and gravel aquifer in Germany, which involved collection of water samples both before and after restarting the system after several years of disuse, found that restarting the system led to a reduction in well water alpha diversity in the first seven days, but a subsequent return of alpha diversity values to near-original levels after forty-nine days (Fillinger et al., 2021). Thus, perhaps a decline in well water prokaryotic diversity can only be expected near the onset of intrusion, with diversity returning to near-original levels over time. This would explain why, despite the indications of intrusion in suspect well B, prokaryotic diversity in this well was more similar to that of the comparison well than to that of the gravel pit pond. Another point worth considering is that suspect well B is situated ~180 m from the gravel pit pond whereas the wells that saw the most obvious drops in alpha diversity in the riverbank filtration 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 & Grischek, 2002). No significant differences in prokaryotic alpha diversity were observed between sampling/measurement sites in the other two study areas.

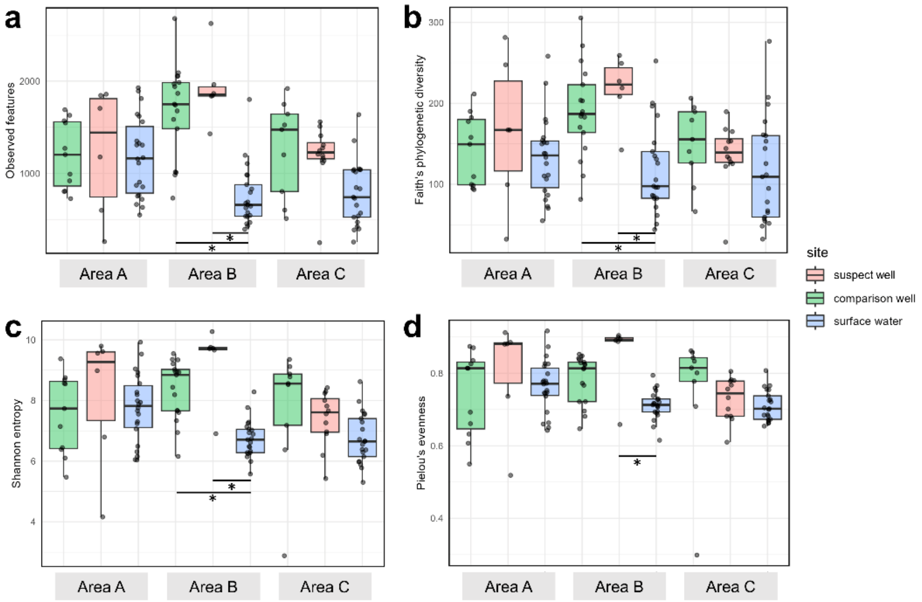

Fig. 17. Box and whisker plots for all wells and surface water bodies, showing values of alpha diversity metrics – (a) Pielou's evenness, (b) Faith's phylogenetic diversity, (c) Observed features, (d) Shannon entropy – as calculated from 16S rRNA gene amplicon sequencing data in QIIME 2™. Asterisks indicate significant differences as determined by adjusted p-values < 0.05 in Dunn's tests.

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

Despite strong indications of surface water intrusion at suspect well B (described above), prokaryotic communities from the suspect well and gravel pit pond in Area B were clearly separated in nMDS plots (Fig. 18). Communities from the suspect well grouped more closely to communities from the comparison well, suggesting similarity between the prokaryotic communities from both wells, despite the suspect well communities being strongly exposed to surface water intrusion, and the comparison well communities being free of this influence. The similarity between the suspect well and comparison well observed here was in agreement with alpha diversity findings described in the previous section.

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 prokaryotic communities. Prokaryotic 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 $\delta^{18}O$, 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.

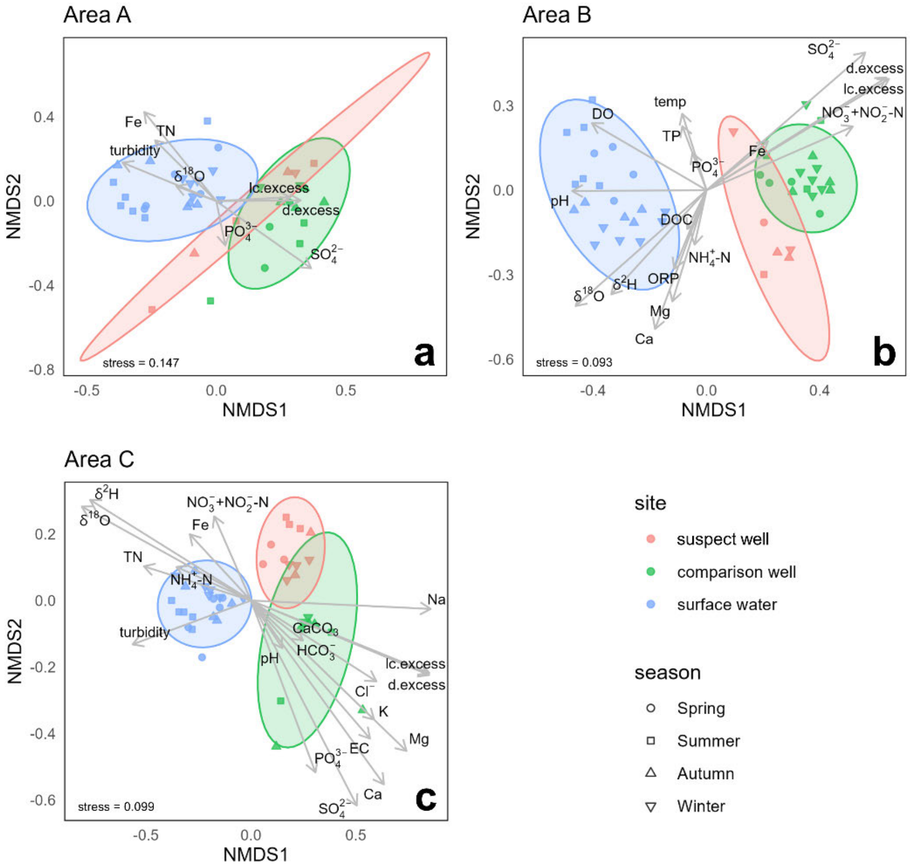

Fig. 18. Non-metric multidimensional scaling (nMDS) plots of Bray–Curtis dissimilarity, for study areas A, B, and C (a, b, c), as calculated from 16S rRNA gene amplicon sequencing data via the metaMDS() function of the ‘vegan’ R package. Water quality variables were fit to plots using the envfit() function 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.

3.3.6 Site-specific differences and temporal variation in dominant prokaryotic groups

Marked increase in the relative abundance of the Crenothrix *genus in suspect well C around the time of the springtime snowmelt*

The relative abundance of the *Crenothrix* genus in samples from suspect well C was low between November 2021 and April 2022, having a median during this period of about 1% (Fig. 19). This was followed, however, by a marked increase in relative abundance to about 25% in late May (shortly after the springtime snowmelt), and levels rose further to a peak of about 37% in late August before falling to about 14% by the final sampling timepoint in late October.

Virtually all of the *Crenothrix* ASVs at this well were classified as *Crenothrix polyspora* at the species level. Although species-level taxonomic annotations from the SILVA database are not curated, and may be somewhat unreliable for that reason (Robeson et al., 2021), the detection of *C. polyspora* in this context would not be entirely unusual. *C. polyspora* is a distinctive filamentous methanotrophic (methane-oxidizing) bacterium with a complex life cycle, and has been proposed as a biological indicator for methane in drinking water wells (Stoecker et al., 2006). Intrusion of surface water from the nearby lake is one possible source of this methane, as intrusion appears to occur (based on physicochemical and isotopic data from this study), and freshwater lakes and especially lake sediments are typically ‘hot spots’ of methane production (Bastviken et al., 2004). Transport of this methane via surface water intrusion might have caused increased abundance of *Crenothrix* spp. in suspect well C. Another possibility is that the increased abundance of *Crenothrix* spp. is driven by methane from a source other than the lake, such as a nearby peatland or from deeper groundwater (Bomberg et al., 2015; Kotiaho et al., 2010). Unfortunately, these possibilities could not be investigated further in this study, because methane concentrations were not measured. Another possible explanation for the high relative abundances of *Crenothrix* detected in suspect well C could be the transport of *Crenothrix* cells (or at least *Crenothrix* DNA) into the well from the nearby lake. *Crenothrix* spp. are known to be capable of growing as planktonic species in the lake water column, and to be a stable part of the indigenous microbial community in some lakes (Oswald et al., 2017). *Crenothrix* spp. were not detected in very high relative abundances in the lake water in our work, but that does not entirely discount this possibility given that absolute abundances remain unknown.

Candidatus Moranbacteria, a group belonging to the Candidate Phylum Radiation (CPR) (Brown et al., 2015; Wrighton et al., 2012), also appeared at high relative abundances in suspect well C (≥10% at seven timepoints). The CPR is a large bacterial lineage of over 70 phyla thought to account for more than a quarter of microbial diversity (Brown et al., 2015; X. He, 2023). These largely uncultured, small-celled symbiotic bacteria are often detected in groundwater, and have small genomes and somewhat (if not highly) restricted metabolic capabilities (C. He et al., 2021; Hug et al., 2016). Little is known as of yet about the members of Candidatus Moranbacteria specifically, although they have also been detected in a thermokarst lake (Vigneron et al., 2020), a sulphide spring biofilm (Valentin-Alvarado et al., 2024), and activated sludge (Fujii et al., 2022).

Other site-specific observations

High relative abundances of the *Gallionella* genus of iron-oxidizing bacteria were detected in samples from comparison well B throughout the entire study period (median: ~23%) (Fig. 19), probably related to the fact that this well had the highest Fe levels of all studied wells (median: ~0.5 mg/L). Dominant groups in the surface waters included *Polynucleobacter* (most abundant in the river in Area A), the hgcl clade (most abundant in the gravel pit pond in Area B, and the lake in Area C), and several others (Fig. 19).

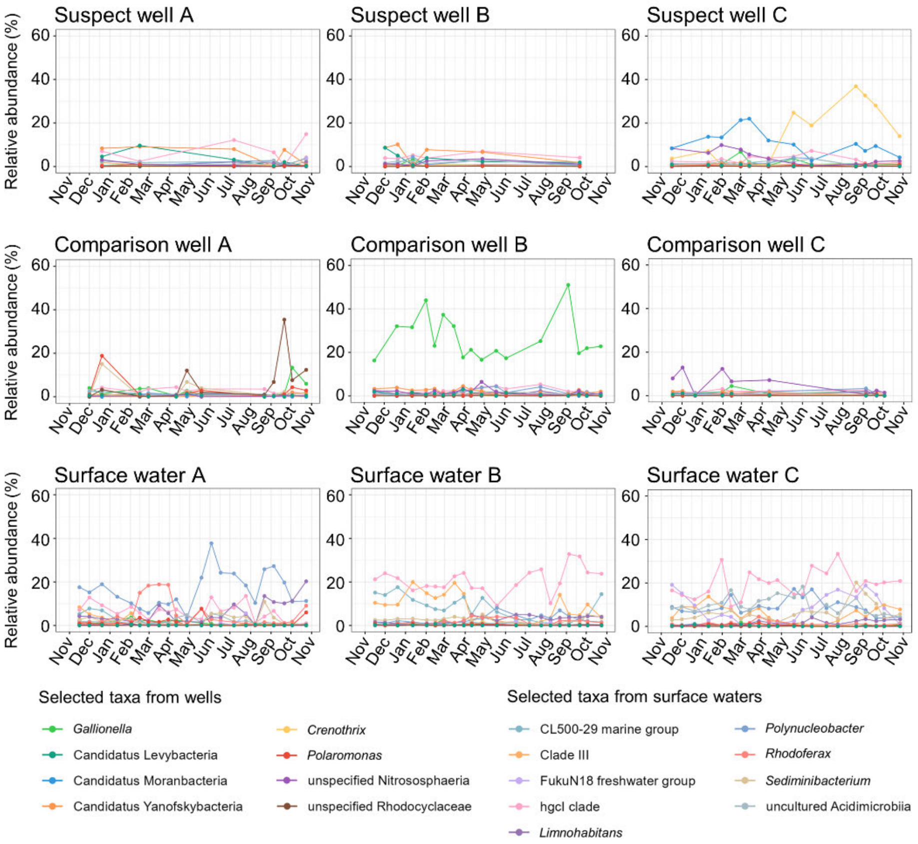

Fig. 19. Selected taxa from wells and surface waters. Taxa were selected for the plots on the basis of their having either (i) high relative abundance in samples from one or more wells, or (ii) high relative abundance in one or more surface waters. All of the taxa listed at the bottom of the figure are shown in all of the plots. Note that the y-axis range of the plots is set to 0–60% rather than 0–100%.

Archaea and cyanobacteria as groundwater and surface water indicators

A recent study of surface water–groundwater interactions in the Murray Darling Basin of New South Wales, Australia found higher relative abundances of archaea in groundwater samples and hyporheic zone samples than in surface water samples, as well as higher relative abundances of cyanobacteria and lower relative abundances of archaea in samples from piezometers under surface water influence, suggesting that these two groups could be useful indicators for assessing the extent of surface water influence in piezometers and wells (Korbel et al., 2022).

Here, relative abundances of archaea showed a similar pattern, with surface water samples tending to have very low relative abundances, and well water samples typically having higher relative abundances (Fig. 20). However, the comparison wells (supposedly free from surface water intrusion) did not appear to have consistently higher relative abundances of archaea than the suspect wells. The highest relative abundances of archaea, of which most were Crenarchaeota, were observed in samples from the two Area C wells. Crenarchaeota were also the most common archaeal group detected across all sites. Smaller amounts of the Halobacterota group were almost exclusive to samples from suspect well C (often comprising between 3% and 8% of the total prokaryotic communities).

The overall highest relative abundances of cyanobacteria were observed in surface water B during the period from about April to August (Fig. 20). These levels did not appear to lead to an increase of cyanobacterial relative abundances in suspect well B, however, despite the indications of intrusion discussed in earlier sections. Suspect well A had relative abundances of about 5% or more at several timepoints, which may indicate some degree of intrusion from the nearby river. The single high relative abundance value observed in comparison well C is likely due to measurement error, as all other values were very low. Many of the well water samples from suspect wells A and B were not successfully sequenced, therefore limiting more in-depth investigation of the aforementioned indicators.

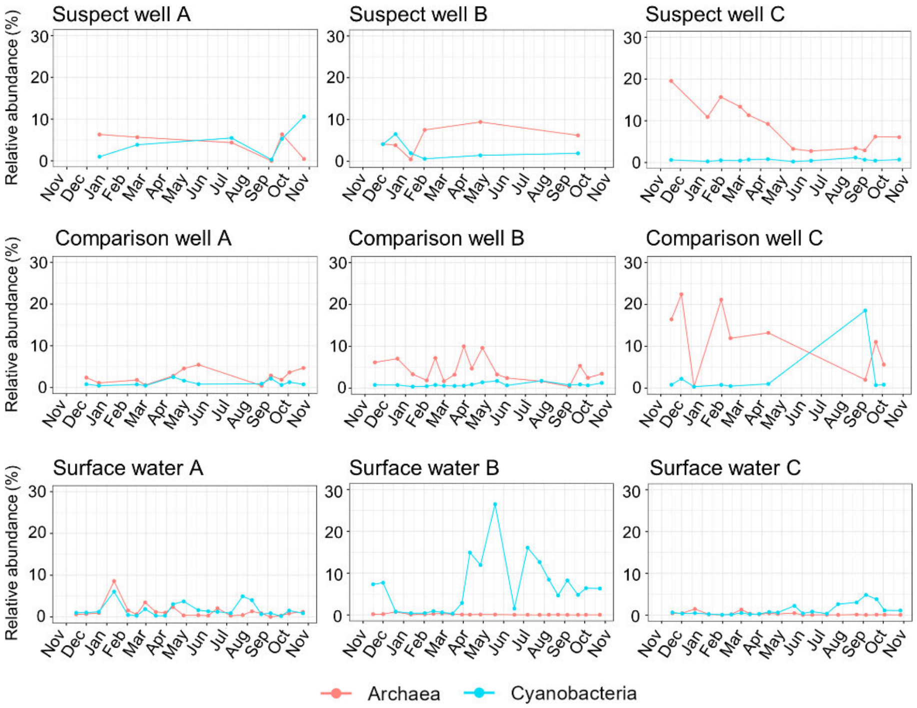

Fig. 20. Relative abundances of archaea and cyanobacteria in the well water and surface water samples. Note that the y-axis range of the plots is set to 0–30% rather than 0–100%.

Temporal variance of relative abundance data

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 (Fig. 21), although the communities of suspect well B might be said to appear rather stable at five of the six successfully sequenced timepoints (excluding January). Sequencing was more successful for suspect well C. At this well, the Nitrososphaeria group of Crenarchaeota, which appeared at about 5–10% relative abundance from November 2021 to March 2022 largely disappeared by May onwards, returning only at the end of the study period in October 2022. Relative abundances of Patescibacteria, which are commonly observed in groundwaters, also appeared to decline at the beginning of the summer (Herrmann et al., 2019; Tian et al., 2020). It should be noted, however, that given the compositional nature of relative abundance values, apparent increases and decreases can be artificially generated by changes in the proportions of other groups (Gloor et al., 2017). Hence, strong conclusions cannot be drawn on the basis of these data alone. Nevertheless, 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 likely to be a strong influence on the groundwater communities at this well.

Overall, direct transfers of microbial groups from surface water to groundwater during periods of supposed surface water intrusion could not be confirmed by the methods used here. Such transfers may occur, although other techniques, such as those that more specifically target indicator species or that assess absolute abundances, may be required to better assess their importance. A previous study of a riverbank filtration system concluded that, influx of surface-water derived taxa played a relatively minor role in shaping groundwater microbial communities, and the same appeared to be true here (Fillinger et al., 2021).

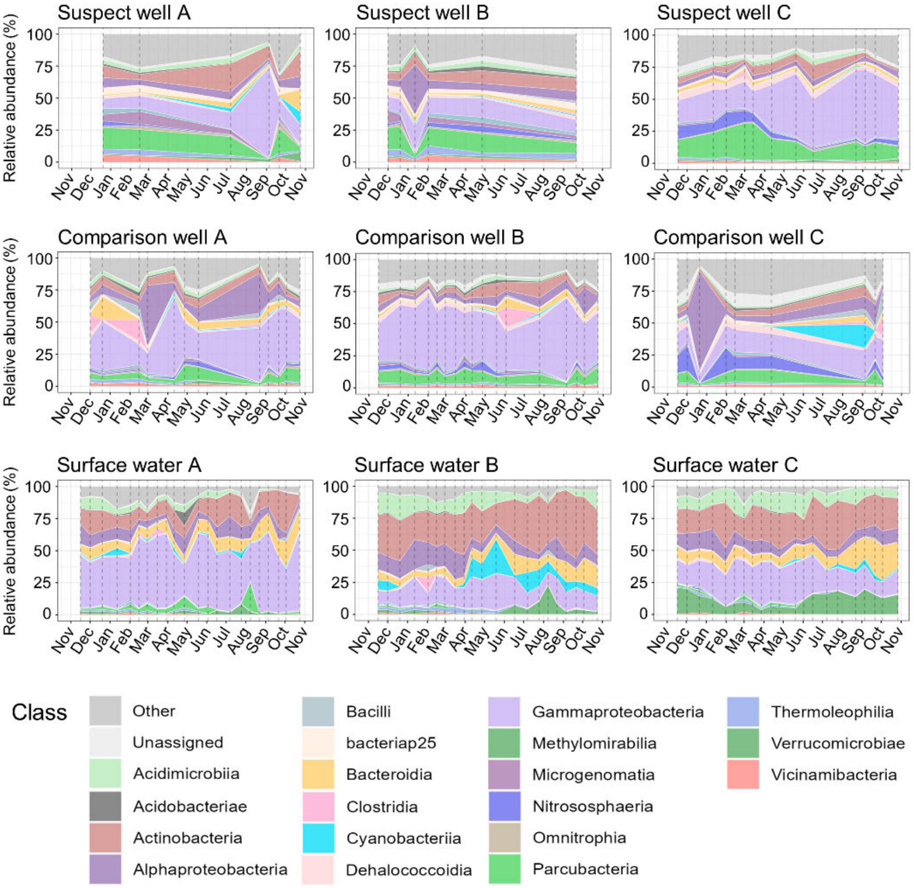

Fig. 21. Stacked area plots of prokaryotic relative abundance data for the nine sampling/measurements sites, showing the top twenty most common classes (based on mean relative abundance across all successfully sequenced samples collected in the study); other classes are combined into the 'Other' category. Successfully sequenced samples are indicated by dashed vertical lines.

3.3.7 Assessment on the use of stable water isotopes for the study of surface water intrusion in shallow wells

Stable water isotopes are a valuable tool for water resource management, as shown by previous work applying stable water isotopes to the study of surface water intrusion in shallow wells in northern Finland based on only one sample per well (Yapiyev et al., 2023). The study described here builds upon that previous work by assessing the seasonal variation of surface water intrusion. Though the application of stable water isotopes for studying surface water intrusion is not new (Criss & Davisson, 1996; Hunt et al., 2005), the method was extended here to provide estimates of the fraction of surface water in the suspect wells based on a simple two-component isotope mixing model using lc-excess* values. Values of lc-excess* were used because normalization of isotopic signal to precipitation can provide a more quantitative assessment of intrusion impact. Application of a mixing model is not always appropriate, as it requires clear labelling of each component (surface water and a groundwater well without intrusion) and can be confounded by other inputs such as snowmelt impact. Here, lc-excess* was used for $R_{surface}$ calculations because using $\delta^{18}O$ overestimated the fraction of surface water ($R_{surface} = 1.27$) in suspect well B in the springtime (April–June) due to depleted isotope composition in the gravel pit pond during that period ($\delta^{18}O = -12.28‰$). The isotope mixing approach could not be meaningfully applied to river water intrusion towards the suspect well in Area A, because the river has stronger direct contributions of precipitation, groundwater, and snowmelt.

3.3.8 Assessment on the use of microbial community analysis for the study of surface water intrusion in shallow wells

Although noticeable similarities were observed between the physicochemical water quality of samples from some wells and surface water bodies – for example, between the comparison well and river in Area A, between suspect well and gravel pit pond in Area B, and between the suspect well and lake in Area C – the same similarities were not seen in microbial community data. In fact, plots of beta diversity showed that, within each study area, there tended to be somewhat more overlap between suspect well and comparison well water samples.

The Area B case seems particularly interesting, because although samples from the suspect well and nearby gravel pit pond had very similar physicochemical water quality, their microbial communities appear quite different (e.g. in the nMDS plot of beta diversity). There are several possible explanations for this. One is that although water from the gravel pit pond travels towards the well (as indicated by isotope data and temperature readings), perhaps the subsurface is quite effective at filtering out or eliminating the surface water microbes such that they do not reach the well in detectable numbers. Another possibility is that while the surface water intrusion may cause some short-term introduction of allochthonous (i.e. foreign) microbes into groundwater (from some nearby surface water body and/or surrounding soils), the overall effect of this disturbance on the groundwater-based communities does not appear to be very large, as shown by the absence of noticeable and/or long-lasting changes in groundwater communities. This is not entirely unsurprising, as the terrestrial subsurface (being low in temperature, oxygen, nutrients and light) is a very different environment to a surface water body, and many microbes that thrive in surface waters cannot therefore be expected to survive for very long in the subsurface environment (Cullimore, 2008).

One limitation of this study is that although water samples were collected twice monthly for a year from the various groundwater wells and surface water bodies, the small volumes collected for DNA sequencing analysis meant that it was not possible to obtain reliable sequencing data from some of the samples (especially those taken from groundwater wells); this problem has also been encountered in other studies (Vargha et al., 2023). Groundwater is a low biomass environment, and hence, larger samples are preferable (e.g. using ultrafiltration methods) (Knappett et al., 2011; Mull & Hill, 2012; Smith & Hill, 2009; Staley et al., 2015). Here, smaller samples were taken for practical and time-related reasons, and the data obtained were still informative, with notable site-specific differences being observed. However, larger samples (at least 4 L or more) appear to produce useful data more reliably (Fiedler et al., 2018; Fillinger et al., 2021; Lin et al., 2012; Ma et al., 2023).

One must admit, too, that if the principal aim of an investigation is to either confirm or refute the presence of surface water intrusion at a given well, then microbial community analysis is unlikely to be the first-preference method, given that there are simpler and more cost-effective methods available.

3.3.9 Summary

Understanding the impact of intruding surface water on shallow wells is important because groundwater provides much of the world’s potable water supply, and reduced quality is a potential public health risk. Here, six wells and three surface water bodies were studied over a period of 12 months, revealing indications of surface water intrusion on the basis of temperature fluctuations, physicochemical and isotopic similarities, and microbial community data.

One of the key findings of this study was the uncovering of a strong groundwater–surface water interaction at suspect well B. This site might be best described as a bank filtration system, whereby much of the water extracted from the well appears to consist not of groundwater but rather of surface water from the nearby pond. 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. Surprisingly, however, no obvious water quality problems were observed at this well, despite the probable intrusion.

Although water from suspect well B showed strong physicochemical similarity with water from the nearby gravel pit pond, and clear indications of surface water intrusion in isotope data, the prokaryotic communities in water from pond and well were noticeably different: prokaryotic alpha diversity was significantly higher in the well than in the pond, and community composition in the well appeared more similar to the comparison well than to the pond. Despite these differences, however, the water from the suspect well did appear to fit the USEPA’s description of ‘groundwater under the direct influence of surface water’ due to marked temperature fluctuations that mirror those of the nearby gravel pit pond. For that reason, risks of microbial contamination originating from the pond cannot be ruled out.

In general, indications of surface water intrusion in suspect wells A and C were less clear than for suspect well B. Suspect well A had, however, the highest cyanobacteria relative abundances of all wells, and suspect well C had slowly declining lc-excess* values and some noticeable fluctuations in the relative abundances of several prokaryotic taxa, which could be further investigated.

Overall, this study revealed seasonal variation in physicochemical water quality, stable water isotope compositions, and microbial community data at the studied sites, as well as some of the strengths and limitations of the methods used, and therefore has relevance to the future study of surface water intrusion in shallow wells.

4 Conclusions and recommendations

This thesis presented findings from three studies of wells extracting shallow groundwater for potable water supply in northern and central Finland. As stated in the introduction, motivations to study groundwater wells can be divided into at least three different categories: (i) the safeguarding of public health, (ii) the improvement of water supply management and groundwater engineering practices, and (iii) the desire to reveal the mysteries of subsurface microbial ecology. This section attempts to summarise the advances made by this work across these three different areas of interest (Fig. 22), and to give recommendations for future research.

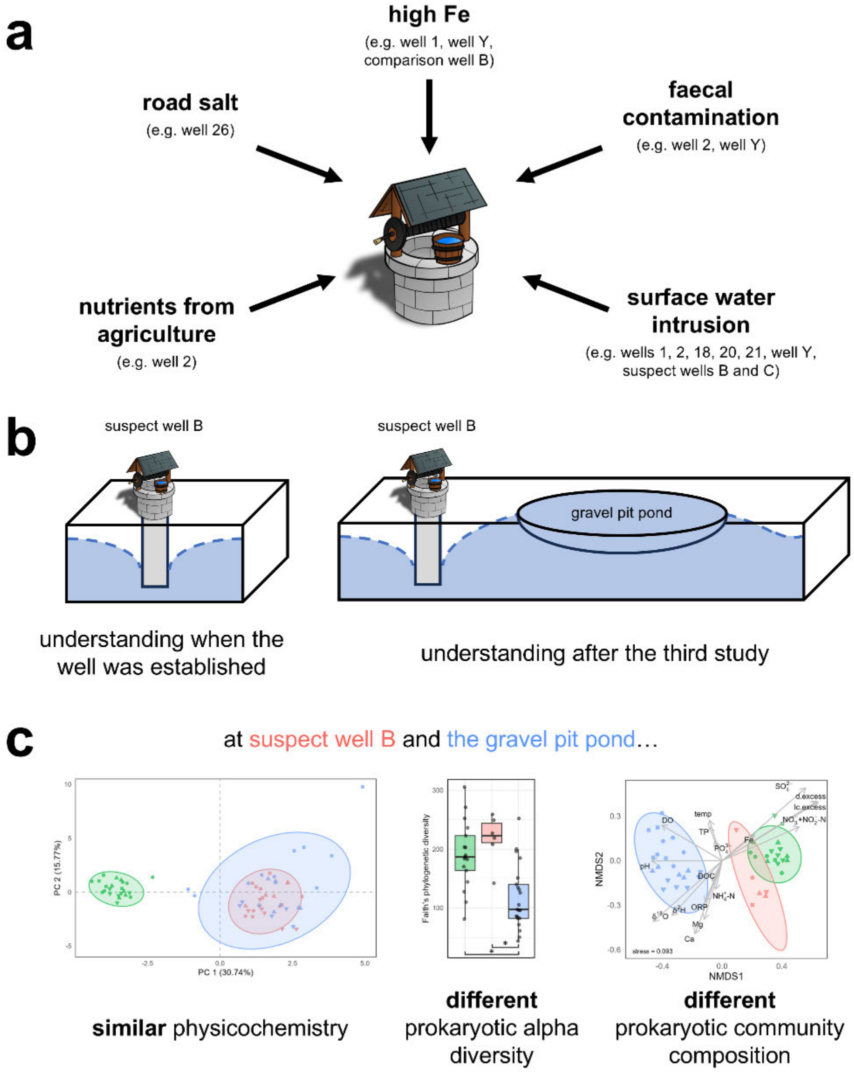

Fig. 22. Overview of some of the main findings from the work described in this thesis. a: Likely factors influencing water quality in the studied wells, based on data collected for this thesis. b: Change in understanding of groundwater–surface water interaction at suspect well B, where data collected for this thesis suggest that most of the well water comes from the nearby gravel pit pond. c: Physicochemical parameters were highly similar in suspect well B and the nearby gravel pit pond; however, prokaryotic alpha diversity and community composition were significantly different.

4.1 Public health

The first study might be said to have advanced the goal of safeguarding public health by highlighting several potentially problematic wells on the basis of physicochemical data, stable water isotopes, microbiological indicators, and microbial community data. Possible influences on well water quality from surface water intrusion, road salt, and agriculture were also discussed with the aid of the collected data and Spearman rank-based correlations (Fig. 22a). This is all potentially useful information for risk management at the studied wells and may be of interest also to water supply professionals and researchers embarking upon similar studies.

In the second study, stable water isotope analyses identified clear indications of surface water intrusion at one of the studied wells, and microbial community data were used to identify potential influences on the wells from surface water intrusion and faecal contamination (Fig. 22a). Again, these findings provide potentially useful information for risk management at the studied wells. One of the other aims of the study was to explore the potential of applying real-time online methods to the monitoring of groundwater physicochemistry. Although there are many examples in the literature of continuous methods being applied to the monitoring of water distribution systems and groundwater level (J. Ikonen et al., 2017; J. M. Ikonen et al., 2017; Luoma et al., 2021; J. Mäkinen et al., 2018; Salo et al., 2019), the second study described in this thesis is one of the few which has implemented real-time online monitoring of groundwater physicochemistry (Oppus, Guico, et al., 2020; Oppus, Guzman, et al., 2020). The real-time online monitoring systems demonstrated potential usefulness for future applications, though more frequent sensor calibration and cleaning may improve the reliability of the measurements.

The third study revealed indications of intrusion in one well consisting of overall physicochemical similarity with a nearby pond, as well as fluctuations in water temperature and isotopic compositions that mirrored the pond (Fig. 22a). Isotope mixing models estimated that 80–95% of the well water comes from the pond. Water from this well appeared to fit the USEPA’s description of ‘groundwater under the direct influence of surface water’, and for that reason the threat of microbial contamination originating from the pond must be considered in risk assessment, as surface water is a known source of microbial contamination in groundwater (P. D. Hynds et al., 2014; Sasakova et al., 2018; Van Driezum et al., 2018).

4.2 Water supply management and groundwater engineering

A major finding described in this thesis relating to water supply management and groundwater engineering was the identification of the strong groundwater–surface water interaction at suspect well B (Fig. 22b). The implications of this have already been described in section 3.3.2. It is worth noting too, however, that the extraction of sand and gravel is considered a common threat to groundwater quality in Finland (Isomäki et al., 2008), in part because sand and gravel deposits are also widely exploited for groundwater extraction, meaning that both uses may often occur in somewhat close proximity. However, it may be that the proximity of wells and gravel pit ponds can also provide unexpected benefits in some cases, by increasing water storage capacity and supplying water to the wells via bank filtration, thus preventing the wells from running dry during periods when groundwater level would otherwise be low. Bank filtration is a long-established and cost-effective method of producing potable water from surface water bodies (Maliva & Missimer, 2012; Ray, 2002; Ray et al., 2002, 2003; Tufenkji et al., 2002), and there may be other sites in Finland and elsewhere in which the reconceptualisation of the groundwater–surface water relationship as a bank filtration system could be exploited to the benefit of potable water supply.

For some of the wells studied here, snapshot sampling of stable water isotopes from well water was sufficient for detecting surface water intrusion. By contrast, however, some wells where intrusion was suspected did not show indications of intrusion in isotope data even over many sampling timepoints. This suggests either (i) that the impact of intrusion at these sites is smaller than initially suspected, or (ii) that the method is less effective at detecting intrusion at these sites. Which of these explanations is true may depend on the site in question. For example, intrusion may be initially suspected at a site where it is not significantly present (based on mistaken site-specific assumptions and deductions), or the nearby surface water body from which the intrusion originates may have similar stable water isotope compositions to the local groundwater – throughout the year, or at certain times of the year (e.g. due to strong precipitation or snowmelt inputs) – making it difficult to differentiate between groundwater and surface water based on isotopes alone. These are worthwhile considerations for any future applications of the stable water isotope method to study surface water intrusion in shallow wells.

4.3 Subsurface microbial ecology

The studies described in this thesis were likely the first major studies to observe the prokaryotic communities of shallow groundwater in Finland (Länsivaara, 2020; Tiwari, 2020; Tiwari et al., 2021), and among the first such studies in the Nordic countries (Knobloch et al., 2021). In all three studies, some logical relationships were observed between well water physicochemistry and dominant prokaryotic taxa (e.g. iron-rich waters and the *Gallionella* genus of iron-oxidising bacteria). As to the issue of surface water intrusion, the first (snapshot) study revealed a correlation between surface water intrusion (as indicated by low d-excess values) and low alpha diversity in the wells, which seemed logical considering that surface waters tends to have lower alpha diversity than groundwaters (Fillinger et al., 2021; Ji, Wang, et al., 2022; Ji, Zhang, et al., 2022). However, the third study, which involved a more in-depth analysis of the temporal variance in wells and surface water bodies, found that even for the well with the strongest indications of surface water intrusion (suspect well B), prokaryotic alpha diversity was significantly higher in the well than in the nearby pond and the prokaryotic communities in the well were found to appear more similar to those of the unaffected comparison well (Fig. 22c). This is in line with findings from a recent study of a riverbank filtration system which found that direct influx of surface water microbes played only a minor role in shaping the studied groundwater communities (Fillinger et al., 2021).

4.4 Future recommendations

From a public health and water supply management perspective, there appears to be great potential for the wider adoption of alternative monitoring methods in shallow groundwater wells, including real-time online monitoring of groundwater physicochemistry, stable water isotope analyses, and analyses of microbial communities. The potential usefulness of these methods should of course 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. Larger water suppliers may have the funds to employ these methods themselves or to pay for outsourcing, whereas smaller suppliers could partner with local authorities, universities, or other research institutes to conduct enhanced monitoring on a temporary basis to investigate specific problems.

As to the question of future research, any further studies of surface water intrusion would do well to investigate both wells and nearby surface water bodies over several sampling/measurement timepoints if a proper understanding of the groundwater–surface water interaction is to be obtained. Special attention should perhaps also be given to time periods where the rate of intrusion is expected to be elevated, and sampling could be 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.

Most aquifer microbes are not freely suspended in groundwater, but are rather attached to surfaces (Griebler et al., 2022). Hence, a complete picture of subsurface microbial ecology cannot be obtained without also investigating solid aquifer samples. These are more challenging to collect compared to groundwater samples but could provide useful information as to the fate of surface water microbes entering the subsurface and to the changes in aquifer community composition brought about by surface water intrusion. Another possible avenue of investigation could be to explore the prokaryotic communities in the soil overlying the groundwater as another potential source of microbes in the well water.

Finally, one of the limitations of the third study described here was that prokaryotic community data could not be obtained from all groundwater samples, due most likely to insufficient sample volume. An evaluation of river water sampling methods for amplicon sequencing of bacterial communities has been previously published (Staley et al., 2015), but nothing similar appears to exist for groundwaters. Perhaps the development of a standard groundwater sampling method or sampling guidelines would reduce the likelihood of sequencing failures, thereby aiding future investigations, and perhaps also enabling more ready comparison between different studies. The volume of groundwater collected, the sample transport and storage methods, the pore size of the membrane used for filtering, as well as the methods used for DNA extraction and sequencing tend to vary from study to study, as do the methods used in processing and analysing the sequencing data, presumably leading to considerable variability in the findings. A study to assess the variability arising from each stage of the pipeline would be a welcome addition to the literature, providing useful guidance for future investigations.

List of references

Alfreider, A., Krössbacher, M., & Psenner, R. (1997). Groundwater samples do not reflect bacterial densities and activity in subsurface systems. *Water Research*, *31*(4), 832–840. https://doi.org/10.1016/S0043-1354(96)00311-9

Anderson, M. P. (2005). Heat as a ground water tracer. *Ground Water*, *43*(6), Article 6. https://doi.org/10.1111/j.1745-6584.2005.00052.x

Andrade, L., O'Dwyer, J., O'Neill, E., & Hynds, P. (2018). Surface water flooding, groundwater contamination, and enteric disease in developed countries: A scoping review of connections and consequences. *Environmental Pollution*, *236*, 540–549. https://doi.org/10.1016/j.envpol.2018.01.104

Arvonen, V., Kibocha, S. N., Katko, T. S., & Pietilä, P. (2017). Features of water cooperatives: A comparative study of Finland and Kenya. *Public Works Management & Policy*, *22*(4), 356–377. https://doi.org/10.1177/1087724X17715267

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. *Critical Reviews in Environmental Science and Technology*, *44*(12), Article 12. https://doi.org/10.1080/10643389.2013.781936

Bastviken, D., Cole, J., Pace, M., & Tranvik, L. (2004). Methane emissions from lakes: Dependence of lake characteristics, two regional assessments, and a global estimate. *Global Biogeochemical Cycles*, *18*(4), 2004GB002238. https://doi.org/10.1029/2004GB002238

Bell, E., Lamminmäki, T., Alneberg, J., Andersson, A. F., Qian, C., Xiong, W., Hettich, R. L., Balmer, L., Frutschi, M., Sommer, G., & Bernier-Latmani, R. (2018). Biogeochemical cycling by a low-diversity microbial community in deep groundwater. *Frontiers in Microbiology*, *9*.

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. *Applied and Environmental Microbiology*, *66*(10), Article 10. 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. *Applied and Environmental Microbiology*, *66*(4), Article 4. https://doi.org/10.1128/AEM.66.4.1587-1594.2000

Bettez, N. D., Marino, R., Howarth, R. W., & Davidson, E. A. (2013). Roads as nitrogen deposition hot spots. *Biogeochemistry*, *114*(1–3), Article 1–3. https://doi.org/10.1007/s10533-013-9847-z

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*(1), Article 1. 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., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), Article 8. https://doi.org/10.1038/s41587-019-0209-9

Bomberg, M., Lamminmäki, T., & Itävaara, M. (2016). Microbial communities and their predicted metabolic characteristics in deep fracture groundwaters of the crystalline bedrock at Olkiluoto, Finland. *Biogeosciences*.

Bomberg, M., Nyyssönen, M., Pitkänen, P., Lehtinen, A., & Itävaara, M. (2015). Active microbial communities inhabit sulphate-methane interphase in deep bedrock fracture fluids in Olkiluoto, Finland. *BioMed Research International*, *2015*, 1–17. https://doi.org/10.1155/2015/979530

Bourg, A. C. M., & Bertin, C. (1993). Biogeochemical processes during the infiltration of river water into an alluvial aquifer. *Environmental Science & Technology*, *27*(4), Article 4. https://doi.org/10.1021/es00041a009

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

Brennand, T. A. (2000). Deglacial meltwater drainage and glaciodynamics: Inferences from Laurentide eskers, Canada. *Geomorphology*, *32*(3–4), 263–293. https://doi.org/10.1016/S0169-555X(99)00100-2

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. *Science of the Total Environment*, *717*, 137249. https://doi.org/10.1016/j.scitotenv.2020.137249

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*(7559), Article 7559. https://doi.org/10.1038/nature14486

Brugger, A., Reitner, B., Kolar, I., Quéric, N., & Herndl, G. J. (2001). Seasonal and spatial distribution of dissolved and particulate organic carbon and bacteria in the bank of an impounding reservoir on the Enns River, Austria. *Freshwater Biology*, *46*(8), 997–1016. https://doi.org/10.1046/j.1365-2427.2001.00743.x

Brunke, M., & Gonser, T. (1997). The ecological significance of exchange processes between rivers and groundwater. *Freshwater Biology*, *37*(1), 1–33. https://doi.org/10.1046/j.1365-2427.1997.00143.x

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. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/srep44350

Burkart, M. R., & Stoner, J. D. (2008). Nitrogen in groundwater associated with agricultural systems. In *Nitrogen in the Environment: Sources, Problems, and Management* (J.L. Hatfield and R.F. Follett (eds), pp. 177–202). Elsevier B.V.

Burri, N. M., Weatherl, R., Moeck, C., & Schirmer, M. (2019). A review of threats to groundwater quality in the anthropocene. *Science of The Total Environment, 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. *Scientific Reports*, *9*(1), Article 1. 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*(4), Article 4. 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. *Nature Methods*, *13*(7), Article 7. https://doi.org/10.1038/nmeth.3869

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

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.*

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. *Science of The Total Environment*, *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. *Journal of Environmental Engineering*, *126*(6), 501–508. https://doi.org/10.1061/(ASCE)0733-9372(2000)126:6(501)

Chiy, P. C., & Phillips, C. J. (2002). Effects of sodium and sulphur fertilisers on the concentrations of nitrate and potentially toxic elements (PTEs) in water leached from permanent pasture. *Journal of the Science of Food and Agriculture*, *82*(8), Article 8. https://doi.org/10.1002/jsfa.1113

Christophersen, N., & Hooper, R. P. (1992). Multivariate analysis of stream water chemical data: The use of principal components analysis for the end-member mixing problem. *Water Resources Research*, *28*(1), 99–107. https://doi.org/10.1029/91WR02518

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*(4), Article 4. https://doi.org/10.1002/wat2.1280

Clark, P. U., & Walder, J. S. (1994). Subglacial drainage, eskers, and deforming beds beneath the Laurentide and Eurasian ice sheets. *Geological Society of America Bulletin*, *106*(2), 304–314. https://doi.org/10.1130/0016-7606(1994)106<0304:SDEADB>2.3.CO;2

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 Research*, *43*(19), Article 19. https://doi.org/10.1016/j.watres.2009.06.036

Craun, G. F., Brunkard, J. M., Yoder, J. S., Roberts, V. A., Carpenter, J., Wade, T., Calderon, R. L., Roberts, J. M., Beach, M. J., & Roy, S. L. (2010). Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. *Clinical Microbiology Reviews*, *23*(3), Article 3.

Craun, G. F., Calderon, R. L., & Craun, M. F. (2003). Waterborne Outbreaks in the United States, 1971-2000. In *Drinking Water Regulation and Health*. John Wiley & Sons, Inc. https://doi.org/10.1002/0471721999

Craun, M. F., Craun, G. F., Calderon, R. L., & Beach, M. J. (2006). Waterborne outbreaks reported in the United States. *Journal of Water and Health*, *4*(S2), Article S2. https://doi.org/10.2166/wh.2006.016

Criss, R. E., & Davisson, M. L. (1996). Isotopic imaging of surface water/groundwater interactions, Sacramento Valley, California. *Journal of Hydrology*, *178*(1–4), 205–222. https://doi.org/10.1016/0022-1694(96)83733-4

Cullimore, D. R. (2008). *Practical Manual of Groundwater Microbiology*.

Cushman, J. H., & Tartakovsky, D. M. (2017). *The handbook of groundwater engineering, third edition*.

Dansgaard, W. (1964). Stable isotopes in precipitation. *Tellus*, *16*(4), Article 4. https://doi.org/10.1111/j.2153-3490.1964.tb00181.x

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. *Applied and Environmental Microbiology*, *70*(9), Article 9. https://doi.org/10.1128/AEM.70.9.5695-5697.2004

Dinno, A. (2022). *Package ‘dunn.test’. R package.*

Domagalski, J. L., & Johnson, H. (2012). *Phosphorus and groundwater: Establishing links between agricultural use and transport to streams. U.S. Geological Survey Fact Sheet 2012-3004* (Fact Sheet Fact Sheet 2012-3004; Fact Sheet, Issue Fact Sheet 2012-3004). United States Geological Survey (USGS).

Donner, J. (1996). The Fennoscandian Shield within Fennoscandia. *Bulletin of the Geological Society of Finland*, *68*(1), 99–103. https://doi.org/10.17741/bgsf/68.1.008

Drage, J., & Kennedy, G. (2020). Building a low-cost, internet-of-things, real-time groundwater level monitoring network. *Groundwater Monitoring & Remediation*, *40*(4), Article 4. https://doi.org/10.1111/gwmr.12408

Edmunds, W. M., & Shand, P. (Eds.). (2008). *Natural groundwater quality*. Blackwell Pub.

Emerson, D., & De Vet, W. (2015). The Role of FeOB in Engineered Water Ecosystems: A Review. *Journal AWWA*, *107*(1), Article 1. https://doi.org/10.5942/jawwa.2015.107.0004

Eren, A. M., Sogin, M. L., Morrison, H. G., Vineis, J. H., Fisher, J. C., Newton, R. J., & McLellan, S. L. (2015). A single genus in the gut microbiome reflects host preference and specificity. *The ISME Journal*, *9*(1), 90–100. https://doi.org/10.1038/ismej.2014.97

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 Environment Research*, *13*, 444–456.

Famiglietti, J. S. (2014). The global groundwater crisis. *Nature Climate Change*, *4*(11), Article 11. https://doi.org/10.1038/nclimate2425

Fiedler, C. J., Schönher, C., Proksch, P., Kerschbaumer, D. J., Mayr, E., Zunabovic-Pichler, M., Domig, K. J., & Perfler, R. (2018). Assessment of microbial community dynamics in river bank filtrate using high-throughput sequencing and flow cytometry. *Frontiers in Microbiology*, *9*, 2887. https://doi.org/10.3389/fmicb.2018.02887

Figueras, M. J., & Borrego, J. J. (2010). New perspectives in monitoring drinking water microbial quality. *International Journal of Environmental Research and Public Health*, *7*(12), Article 12. 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 Research*, *189*, 116631. https://doi.org/10.1016/j.watres.2020.116631

Fitts, C. R. (2013). *Groundwater Science*. Elsevier. https://doi.org/10.1016/B978-0-12-384705-8.00001-7

Fortune, W. B., & Mellon, M. G. (1938). Determination of Iron with o-Phenanthroline: A Spectrophotometric Study. *Industrial & Engineering Chemistry Analytical Edition*, *10*(2), 60–64. https://doi.org/10.1021/ac50118a004

Foster, T., Priadi, C., Kotra, K. K., Odagiri, M., Rand, E. C., & Willetts, J. (2021). Self-supplied drinking water in low- and middle-income countries in the Asia-Pacific. *Npj Clean Water*, *4*(1), 37. https://doi.org/10.1038/s41545-021-00121-6

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*, 15.

Foulquier, A., Mermillod-Blondin, F., Malard, F., & Gibert, J. (2011). Response of sediment biofilm to increased dissolved organic carbon supply in groundwater artificially recharged with stormwater. *Journal of Soils and Sediments*, *11*(2), 382–393. https://doi.org/10.1007/s11368-010-0323-2

Fricker, E. J., Illingworth, K. S., & Fricker, C. R. (1997). Use of two formulations of Colilert and QuantiTray^{TM} for assessment of the bacteriological quality of water. *Water Research*, *31*(10), Article 10. https://doi.org/10.1016/S0043-1354(96)00342-9

Fujii, N., Kuroda, K., Narihiro, T., Aoi, Y., Ozaki, N., Ohashi, A., & Kindaichi, T. (2022). Metabolic potential of the superphylum Patescibacteria reconstructed from activated sludge samples from a municipal wastewater treatment plant. *Microbes and Environments*, *37*(3), n/a. https://doi.org/10.1264/jsme2.ME22012

Gale, J. E. (1993). Defining the hydrogeological framework for environmental studies in Newfoundland and Labrador. *The Scientific Challenge of Our Changing Environment: A Conference Addressing Environmental Change in Newfoundland and Labrador and Similar Regions*, *29*, 155–171.

Gat, J. (2010). *Isotope hydrology: A study of the water cycle*. Imperial College Press. https://doi.org/10.1142/p027

Gat, J. R. (1996). *Oxygen and hydrogen isotopes in the hydrologic cycle*. 38. https://doi.org/10.1146/annurev.earth.24.1.225

Gat, J. R., & Gonfiantini, R. (Eds.). (1981). *Stable isotope hydrology: Deuterium and oxygen-18 in the water cycle*. Internat. Atomic Energy Agency.

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. *Frontiers in Microbiology*, *10*, 1573. https://doi.org/10.3389/fmicb.2019.01573

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017). Microbiome datasets are compositional: And this is not optional. *Frontiers in Microbiology*, *8*, 2224. https://doi.org/10.3389/fmicb.2017.02224

Gollnitz, W. D., Clancy, J. L., Whitteberry, B. L., & Vogt, J. A. (2003). RBF as a Microbial Treatment Process. *Journal - American Water Works Association*, *95*(12), Article 12. https://doi.org/10.1002/j.1551-8833.2003.tb10511.x

Griebler, C., Fillinger, L., Karwautz, C., & Hose, G. C. (2022). Knowledge gaps, obstacles, and research frontiers in groundwater microbial ecology. In *Encyclopedia of Inland Waters* (pp. 611–624). Elsevier. https://doi.org/10.1016/B978-0-12-819166-8.00200-0

Griebler, C., & Lueders, T. (2009). Microbial biodiversity in groundwater ecosystems. *Freshwater Biology*, *54*(4), Article 4.

Griebler, C., Malard, F., & Lefébure, T. (2014). Current developments in groundwater ecology—From biodiversity to ecosystem function and services. *Current Opinion in Biotechnology*, *27*, 159–167. https://doi.org/10.1016/j.copbio.2014.01.018

Gustafsson, J., & Nystén, T. (2000). Trends of chloride concentration in groundwater and results of risk assessment of road salting in Finland. In *Groundwater 2000: Proceedings of the International Conference on Groundwater Research, Copenhagen, Denmark, 6-8 June 2000* (pp. 249–251).

Guzman-Herrador, B., Carlander, A., Ethelberg, S., Freiesleben de Blasio, B., Kuusi, M., Lund, V., Löfdahl, M., MacDonald, E., Nichols, G., Schönning, C., Sudre, B., Trönnberg, L., Vold, L., Semenza, J., & Nygård, K. (2015). Waterborne outbreaks in the Nordic countries, 1998 to 2012. *Eurosurveillance*, *20*(24), Article 24. https://doi.org/10.2807/1560-7917.ES2015.20.24.21160

Halko, N., Martinsson, P.-G., Shkolnisky, Y., & Tygert, M. (2011). An Algorithm for the Principal Component Analysis of Large Data Sets. *SIAM Journal on Scientific Computing*, *33*(5), 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*(2), 264. https://doi.org/10.3390/w11020264

Harrell, F. E. (2020). *Hmisc: Harrell miscellaneous. R package.*

Harwood, V. J., Staley, C., Badgley, B. D., Borges, K., & Korajkic, A. (2014). Microbial source tracking markers for detection of fecal contamination in environmental waters: Relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews, 38*(1), Article 1. https://doi.org/10.1111/1574-6976.12031

Harwood, V., Shanks, O., Korajkic, A., Verbyla, M., Ahmed, W., & Iriarte, M. (2019). General and host-associated bacterial indicators of faecal pollution. In Michigan State University, J. B. Rose, B. Jiménez Cisneros, & UNESCO - International Hydrological Programme (Eds.), *Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management (Global Water Pathogen Project).* Michigan State University. https://doi.org/10.14321/waterpathogens.6

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. *Systematic and Applied Microbiology, 33*(6), Article 6. https://doi.org/10.1016/j.syapm.2010.06.001

He, C., Keren, R., Whittaker, M. L., Farag, I. F., Doudna, J. A., Cate, J. H. D., & Banfield, J. F. (2021). Genome-resolved metagenomics reveals site-specific diversity of episymbiotic CPR bacteria and DPANN archaea in groundwater ecosystems. *Nature Microbiology*, *6*(3), 354–365. https://doi.org/10.1038/s41564-020-00840-5

He, X. (2023). Culture-based approaches to studying “microbial dark matter”. *Proceedings of the National Academy of Sciences, 120*(2), e2219691120. https://doi.org/10.1073/pnas.2219691120

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. *The ISME Journal*, *5*(10), Article 10. https://doi.org/10.1038/ismej.2011.41

Herrmann, M., Geesink, P., Yan, L., Lehmann, R., Totsche, K. U., & Küsel, K. (2020). Complex food webs coincide with high genetic potential for chemolithoautotrophy in fractured bedrock groundwater. *Water Research, 170,* 115306. https://doi.org/10.1016/j.watres.2019.115306

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. *Frontiers in Microbiology*, *10,* 1407. https://doi.org/10.3389/fmicb.2019.01407

Hiscock, K. M., & Grischek, T. (2002). Attenuation of groundwater pollution by bank filtration. *Journal of Hydrology*, *266*(3–4), Article 3–4. https://doi.org/10.1016/S0022-1694(02)00158-0

Hofkes, E. H., Smet, J., & Wijk, C. van. (2002). *Small community water supplies: Technology, people and partnership*. IRC International Water and Sanitation Centre.

Holmlund, P., Clason, C., & Blomdahl, K. (2016). *The effect of a glaciation on East Central Sweden: Case studies on present glaciers and analyses of landform data*. https://doi.org/10.13140/RG.2.2.26356.14720

Howard, G., Pedley, S., Barrett, M., Nalubega, M., & Johal, K. (2003). Risk factors contributing to microbiological contamination of shallow groundwater in Kampala, Uganda. *Water Research*, *37*(14), Article 14. https://doi.org/10.1016/S0043-1354(03)00235-5

Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., Butterfield, C. N., Hernsdorf, A. W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D. A., Finstad, K. M., Amundson, R., Thomas, B. C., & Banfield, J. F. (2016). A new view of the tree of life. *Nature Microbiology*, *1*(5), Article 5. https://doi.org/10.1038/nmicrobiol.2016.48

Huhtanen, P., Ahvenjärvi, S., & Heikkilä, T. (2000). Effects of sodium sulphate and potassium chloride fertilizers on the nutritive value of timothy grown on different soils. *Agricultural and Food Science*, *9*(2), Article 2. https://doi.org/10.23986/afsci.5653

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. *Journal of Hydrology*, *302*(1–4), Article 1–4. https://doi.org/10.1016/j.jhydrol.2004.07.010

Hunter, P. R., Pond, K., Jagals, P., & Cameron, J. (2009). An assessment of the costs and benefits of interventions aimed at improving rural community water supplies in developed countries. *Science of The Total Environment*, *407*(12), Article 12. https://doi.org/10.1016/j.scitotenv.2009.03.013

Hynds, P. D., Thomas, M. K., & Pintar, K. D. M. (2014). Contamination of groundwater systems in the US and Canada by enteric pathogens, 1990–2013: A review and pooled-analysis. *PLOS One*, *9*(5), Article 5. https://doi.org/10.1371/journal.pone.0093301

Hynds, P., Misstear, B. D., Gill, L. W., & Murphy, H. M. (2014). Groundwater source contamination mechanisms: Physicochemical profile clustering, risk factor analysis and multivariate modelling. *Journal of Contaminant Hydrology*, *159*, 47–56. https://doi.org/10.1016/j.jconhyd.2014.02.001

Ikonen, J. M., Hokajärvi, A.-M., Heikkinen, J., Pitkänen, T., Ciszek, R., Kolehmainen, M., Pursiainen, A., Kauppinen, A., Kusnetsov, J., Torvinen, E., & Miettinen, I. T. (2017). Drinking water quality in distribution systems of surface and ground waterworks in Finland. *Journal of Water Security*, *3*(1), Article 1. https://doi.org/10.15544/jws.2017.004

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. *Journal of Environmental Management*, *198*, 384–392. https://doi.org/10.1016/j.jenvman.2017.04.090

Inkinen, J., Jayaprakash, B., Siponen, S., Hokajärvi, A.-M., Pursiainen, A., Ikonen, J., Ryzhikov, I., Täubel, M., Kauppinen, A., Paananen, J., Miettinen, I. T., Torvinen, E., Kolehmainen, M., & Pitkänen, T. (2019). Active eukaryotes in drinking water distribution systems of ground and surface waterworks. *Microbiome*, *7*(1), Article 1. https://doi.org/10.1186/s40168-019-0715-5

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. *Hydrology and Earth System Sciences*, *19*(3), Article 3. https://doi.org/10.5194/hess-19-1247-2015

Isomäki, E., Valve, M., & Kivimäki, A.-L. (2006). *Small waterworks in Finland*. 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.

Itävaara, M., Nyyssönen, M., Kapanen, A., Nousiainen, A., Ahonen, L., & Kukkonen, I. (2011). Characterization of bacterial diversity to a depth of 1500 m in the Outokumpu deep borehole, Fennoscandian Shield. *FEMS Microbiology Ecology*, *77*(2), 295–309. https://doi.org/10.1111/j.1574-6941.2011.01111.x

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

Ji, L., Zhang, L., Wang, Z., Zhu, X., & Ning, K. (2022). High biodiversity and distinct assembly patterns of microbial communities in groundwater compared with surface water. *Science of The Total Environment*, *834*, 155345. https://doi.org/10.1016/j.scitotenv.2022.155345

Jian, C., Luukkonen, P., Yki-Järvinen, H., Salonen, A., & Korpela, K. (2020). Quantitative PCR provides a simple and accessible method for quantitative microbiota profiling. *PLOS ONE*, *15*(1), Article 1. https://doi.org/10.1371/journal.pone.0227285

Josse, J., & Husson, F. (2016). missMDA: a package for handling missing values in multivariate data analysis. *Journal of Statistical Software*, *70*(1). https://doi.org/10.18637/jss.v070.i01

Kalbus, E., Reinstorf, F., & Schirmer, M. (2006). Measuring methods for groundwater – surface water interactions: A review. *Hydrology and Earth System Sciences*, *10*(6), 873–887. https://doi.org/10.5194/hess-10-873-2006

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. *Journal of Environmental Monitoring*, *12*, 759–768. https://doi.org/10.1039/B917937B

Karro, E., & Lahermo, P. (1999). Occurrence and chemical characteristics of groundwater in Precambrian bedrock in Finland. *Geological Survey of Finland, Special Paper 27*, 85–96.

Karwautz, C., & Griebler, C. (2022). Microbial biodiversity in groundwater ecosystems. In *Encyclopedia of Inland Waters* (pp. 397–411). Elsevier. https://doi.org/10.1016/B978-0-12-819166-8.00119-5

Katko, T. S., Lipponen, M. A., & Rönkä, E. K. T. (2006). Groundwater use and policy in community water supply in Finland. *Hydrogeology Journal*, *14*(1–2), Article 1–2. https://doi.org/10.1007/s10040-004-0351-3

Katsanou, K., & Karapanagioti, H. K. (2017). Surface Water and Groundwater Sources for Drinking Water. In A. Gil, L. A. Galeano, & M. Á. Vicente (Eds.), *Applications of Advanced Oxidation Processes (AOPs) in Drinking Water Treatment* (Vol. 67, pp. 1–19). Springer International Publishing. https://doi.org/10.1007/698_2017_140

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. *International Journal of Environmental Research and Public Health*, *16*(22), Article 22. https://doi.org/10.3390/ijerph16224376

Kers, J. G., & Saccenti, E. (2022). The power of microbiome studies: Some considerations on which alpha and beta metrics to use and how to report results. *Frontiers in Microbiology*, *12*, 796025. https://doi.org/10.3389/fmicb.2021.796025

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. *Science of The Total Environment*, *533*, 566–575. https://doi.org/10.1016/j.scitotenv.2015.06.080

Kitterød, N.-O., Kværner, J., Aagaard, P., Arustienė, J., Bikše, J., Dagestad, A., Gundersen, P., Hansen, B., Hjartarson, Á., Karro, E., Klavins, M., Marandi, A., Radienė, R., Retike, I., Rossi, P. M., & Thorling, L. (2022). Hydrogeology and groundwater quality in the Nordic and Baltic countries. *Hydrology Research*, *53*(7), 958–982. https://doi.org/10.2166/nh.2022.018

Kløve, B., Kvitsand, H. M. L., Pitkänen, T., Gunnarsdottir, M. J., Gaut, S., Gardarsson, S. M., Rossi, P. M., & Miettinen, I. (2017). Overview of groundwater sources and water-supply systems, and associated microbial pollution, in Finland, Norway and Iceland. *Hydrogeology Journal*, *25*(4), Article 4. https://doi.org/10.1007/s10040-017-1552-x

Knappett, P. S. K., Layton, A., McKay, L. D., Williams, D., Mailloux, B. J., Huq, M. R., Alam, M. J., Ahmed, K. M., Akita, Y., Serre, M. L., Sayler, G. S., & van Geen, A. (2011). Efficacy of hollow-fiber ultrafiltration for microbial sampling in groundwater. *Ground Water*, *49*(1), Article 1. https://doi.org/10.1111/j.1745-6584.2010.00712.x

Knobloch, S., Klonowski, A. M., Tómasdóttir, S., Kristjánsson, B. R., Guðmundsson, S., & Marteinsson, V. Þ. (2021). Microbial intrusion and seasonal dynamics in the groundwater microbiome of a porous basaltic rock aquifer used as municipal water reservoir. *FEMS Microbiology Ecology*, *97*(3), fiab014. https://doi.org/10.1093/femsec/fiab014

Knutsson, G. (2008). Hydrogeology in the Nordic Countries. *Episodes*, *31*(1), 148–154. https://doi.org/10.18814/epiiugs/2008/v31i1/020

Kolde, R. (2019). *Pheatmap: Pretty heatmaps. R package.*

Konikow, L. F., & Kendy, E. (2005). Groundwater depletion: A global problem. *Hydrogeology Journal*, *13*(1), Article 1. https://doi.org/10.1007/s10040-004-0411-8

Korbel, K. L., Rutlidge, H., Hose, G. C., Eberhard, S. M., & Andersen, M. S. (2022). Dynamics of microbiotic patterns reveal surface water groundwater interactions in intermittent and perennial streams. *Science of The Total Environment*, *811*, 152380. https://doi.org/10.1016/j.scitotenv.2021.152380

Kortelainen, N. (2007). *Isotopic fingerprints in surficial waters stable isotope methods applied in hydrogeological studies*. Geological Survey of Finland.

Kortelainen, N. (2011). Isotope tracing in groundwater applications. *Geological Survey of Finland Special Paper 49*, 279–284.

Kortelainen, N. M., & Karhu, J. A. (2004). Regional and seasonal trends in the oxygen and hydrogen isotope ratios of Finnish groundwaters: A key for mean annual precipitation. *Journal of Hydrology*, *285*(1–4), 143–157. https://doi.org/10.1016/j.jhydrol.2003.08.014

Kotiaho, M., Fritze, H., Merilä, P., Juottonen, H., Leppälä, M., Laine, J., Laiho, R., Yrjälä, K., & Tuittila, E.-S. (2010). Methanogen activity in relation to water table level in two boreal fens. *Biology and Fertility of Soils*, *46*(6), 567–575. https://doi.org/10.1007/s00374-010-0461-0

Kristensen, P., & Bøgestrand, J. (1996). *Surface water quality monitoring*. European Environment Agency.

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. *Frontiers in Microbiology*, *8*, 1951. https://doi.org/10.3389/fmicb.2017.01951

Kutvonen, H., Rajala, P., Carpén, L., & Bomberg, M. (2015). Nitrate and ammonia as nitrogen sources for deep subsurface microorganisms. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.01079

Lahermo, P. (2002). *Tuhat kaivoa—Suomen kaivovesien fysikaalis-kemiallinen laatu vuonna 1999*. Geologian tutkimuskeskus.

Landwehr, J. M., & Coplen, T. B. (2004). Line-conditioned excess: A new method for characterizing stable hydrogen and oxygen isotope ratios in hydrologic systems (4 pages). *Proceedings of an International Conference Held in Monaco, 25–29 October 2004 Organized by the International Atomic Energy Agency and Co-Sponsored by the Abdus Salam International Centre for Theoretical Physics, International Hydrological Programme of UNESCO, Intergovernmental Oceanographic Commission of UNESCO and the Commission Internationale Pour l’Exploration Scientifique de La Mer Méditerranée*. Isotopes in Environmental Studies: Aquatic Forum 2004, Monte-Carlo, Monaco.

Länsivaara, A. (2020). *Characterization of the bacteriome of water treatment processes with next generation sequencing*.

Lavapuro, M., Lipponen, A., Artimo, A., & Katko, T. S. (2008). Groundwater sustainability indicators: Testing with Finnish data. *Boreal Environment Research*, *13*, 381–402.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software*, *25*(1). https://doi.org/10.18637/jss.v025.i01

Lee, J.-H., Lee, B.-J., & Unno, T. (2018). Bacterial communities in ground-and surface water mixing zone induced by seasonal heavy extraction of groundwater. *Geomicrobiology Journal*, *35*(9), Article 9. https://doi.org/10.1080/01490451.2018.1468834

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. *Hydrological Processes*, *21*, 907–919. https://doi.org/10.1002/hyp.6282

Liddle, E. S., Mager, S. M., & Nel, E. L. (2016). The importance of community-based informal water supply systems in the developing world and the need for formal sector support. *The Geographical Journal*, *182*(1), 85–96. https://doi.org/10.1111/geoj.12117

Ligon, G., & Bartram, J. (2016). Literature review of associations among attributes of reported drinking water disease outbreaks. *International Journal of Environmental Research and Public Health*, *13*(6), 527. https://doi.org/10.3390/ijerph13060527

Lin, X., McKinley, J., Resch, C. T., Kaluzny, R., Lauber, C. L., Fredrickson, J., Knight, R., & Konopka, A. (2012). Spatial and temporal dynamics of the microbial community in the Hanford unconfined aquifer. *The ISME Journal*, *6*(9), Article 9. https://doi.org/10.1038/ismej.2012.26

Liu, F., Williams, M. W., & Caine, N. (2004). Source waters and flow paths in an alpine catchment, Colorado Front Range, United States. *Water Resources Research*, *40*(9), 2004WR003076. https://doi.org/10.1029/2004WR003076

Liu, Y., Fang, Y., Hu, H., Tian, F., Dong, Z., & Khan, M. Y. A. (2020). Ecohydrological separation hypothesis: Review and prospect. *Water*, *12*(8), 2077. https://doi.org/10.3390/w12082077

Long, P. E., Williams, K. H., Hubbard, S. S., & Banfield, J. F. (2016). Microbial metagenomics reveals climate-relevant subsurface biogeochemical processes. *Trends in Microbiology*, *24*(8), Article 8. https://doi.org/10.1016/j.tim.2016.04.006

Luef, B., Frischkorn, K. R., Wrighton, K. C., Holman, H.-Y. N., Birarda, G., Thomas, B. C., Singh, A., Williams, K. H., Siegerist, C. E., Tringe, S. G., Downing, K. H., Comolli, L. R., & Banfield, J. F. (2015). Diverse uncultivated ultra-small bacterial cells in groundwater. *Nature Communications*, *6*(1), Article 1. https://doi.org/10.1038/ncomms7372

Luoma, S., & Ikonen, J. (2020). *Vulnerability assessment of the shallow groundwater in Finland* (p. 30). Geological Survey of Finland.

Luoma, S., Majaniemi, J., Pullinen, A., Mursu, J., & Virtasalo, J. J. (2021). Geological and groundwater flow model of a submarine groundwater discharge site at Hanko (Finland), northern Baltic Sea. *Hydrogeology Journal*, *29*(3), 1279–1297. https://doi.org/10.1007/s10040-021-02313-3

Ma, N., Gao, L., Ge, Z., & Li, M. (2023). Hydrochemical characteristics of groundwater in a plain river network region: Establishing linkages between source and water quality variables. *Chemosphere*, *331*, 138809. https://doi.org/10.1016/j.chemosphere.2023.138809

Maa- ja metsätalousministeriö. (2019). *Vesihuollon tilastoja – Maa- ja metsätalousministeriö (Water Supply Statistics – Ministry of Agriculture and Forestry)*. https://mmm.fi/vesi/vesihuolto_tilastot

Macler, B. A., & Merkle, J. C. (2000). Current knowledge on groundwater microbial pathogens and their control. *Hydrogeology Journal*, *8*(1), Article 1. https://doi.org/10.1007/PL00010972

Mahmoud, N., Zayed, O., & Petrusevski, B. (2022). Groundwater quality of drinking water wells in the West Bank, Palestine. *Water*, *14*(3), 377. https://doi.org/10.3390/w14030377

Mäkinen, J., Kallio, E., & Jokela, P. (2018). Managed aquifer recharge and sedimentological characterization within the complex esker deposits in Pälkäne, Finland. *Sustainable Water Resources Management*, *4*(2), 345–359. https://doi.org/10.1007/s40899-017-0195-3

Mäkinen, R., Orvomaa, M., Veijalainen, N., & Huttunen, I. (2008). The climate change and groundwater regimes in Finland. In J. Fehér (Ed.), *Engineering and Industry* (Vol. 2). Trivent Publishing. https://doi.org/10.22618/TP.EI.20162.120014

Maliva, R., & Missimer, T. (2012). Riverbank Filtration. In R. Maliva & T. Missimer, *Arid Lands Water Evaluation and Management* (pp. 631–645). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29104-3_24

Marcheggiani, S., Iaconelli, M., D’angelo, A., Pierdominici, E., La Rosa, G., Muscillo, M., Equestre, M., & Mancini, L. (2008). Microbiological and 16S rRNA analysis of sulphite-reducing clostridia from river sediments in central Italy. *BMC Microbiology*, *8*(1), 171. https://doi.org/10.1186/1471-2180-8-171

Margat, J., & van der Gun, J. (2013). *Groundwater around the world: A geographic synopsis*. Taylor & Francis Group.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, *17*(1), Article 1. https://doi.org/10.14806/ej.17.1.200

Martinez Arbizu, P. (2020). *pairwiseAdonis: Pairwise multilevel comparison using adonis. R package.*

Massmann, G., Pekdeger, A., & Merz, C. (2004). Redox processes in the Oderbruch polder groundwater flow system in Germany. *Applied Geochemistry*, *19*(6), 863–886. https://doi.org/10.1016/j.apgeochem.2003.11.006

McLellan, S. L., & Eren, A. M. (2014). Discovering new indicators of fecal pollution. *Trends in Microbiology*, *22*(12), Article 12. https://doi.org/10.1016/j.tim.2014.08.002

Miettinen, H., Kietäväinen, R., Sohlberg, E., Numminen, M., Ahonen, L., & Itävaara, M. (2015). Microbiome composition and geochemical characteristics of deep subsurface high-pressure environment, Pyhäsalmi mine Finland. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.01203

Misstear, B., Banks, D., & Clark, L. (2017). *Water wells and boreholes*.

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. *The Analyst*, *107*(1275), 659–663. https://doi.org/10.1039/AN9820700659

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. *Journal - American Water Works Association*, *92*(9), Article 9. https://doi.org/10.1002/j.1551-8833.2000.tb09010.x

Mull, B., & Hill, V. R. (2012). Recovery of diverse microbes in high turbidity surface water samples using dead-end ultrafiltration. *Journal of Microbiological Methods*, *91*(3), Article 3. https://doi.org/10.1016/j.mimet.2012.10.001

Murphy, H. M., Prioleau, M. D., Borchardt, M. A., & Hynds, P. D. (2017). Review: Epidemiological evidence of groundwater contribution to global enteric disease, 1948–2015. *Hydrogeology Journal*, *25*(4), Article 4. https://doi.org/10.1007/s10040-017-1543-y

Nnadi, F. N., & Fulkerson, M. (2002). Assessment of groundwater under direct influence of surface water. *Journal of Environmental Science and Health, Part A*, *37*(7), 1209–1222. https://doi.org/10.1081/ESE-120005981

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.* (p. 18).

Nyyssönen, M., Bomberg, M., Kapanen, A., Nousiainen, A., Pitkanen, P., & Itavaara, M. (2012). Methanogenic and sulphate-reducing microbial communities in deep groundwater of crystalline rock fractures in Olkiluoto, Finland. *Geomicrobiology Journal*.

Okkonen, J., Jyrkama, M., & Kløve, B. (2010). A conceptual approach for assessing the impact of climate change on groundwater and related surface waters in cold regions (Finland). *Hydrogeology Journal*, *18*(2), Article 2. https://doi.org/10.1007/s10040-009-0529-9

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)*, 152–155. https://doi.org/10.1109/ECICE50847.2020.9301948

Oppus, C., Guzman, M. A. L., Monje, J. C., Guico, M. L., Ngo, G., Domingo, A., Kwong, J. C., & Retirado, M. G. (2020). Software and data visualization platform for groundwater level and quality monitoring system. *Journal of Physics: Conference Series*, 10.

Oswald, K., Graf, J. S., Littmann, S., Tienken, D., Brand, A., Wehrli, B., Albertsen, M., Daims, H., Wagner, M., Kuypers, M. M., Schubert, C. J., & Milucka, J. (2017). Crenothrix are major methane consumers in stratified lakes. *The ISME Journal*, *11*(9), 2124–2140. https://doi.org/10.1038/ismej.2017.77

Pedersen, K., Bengtsson, A. F., Edlund, J. S., & Eriksson, L. C. (2014). Sulphate-controlled diversity of subterranean microbial communities over depth in deep groundwater with opposing gradients of sulphate and methane. *Geomicrobiology Journal, 31*(7), 617–631. https://doi.org/10.1080/01490451.2013.879508

Phillips, D. L., Newsome, S. D., & Gregg, J. W. (2005). Combining sources in stable isotope mixing models: Alternative methods. *Oecologia, 144*(4), 520–527. https://doi.org/10.1007/s00442-004-1816-8

Pietilä, P. E., & Vihanta, J. (2018). The role of cooperatives in the provision of water services in Finland. In *Challenges and opportunities facing water-service co-operatives: Cases from Argentina and Finland. WATERLAT-GOBACIT Network Working papers.* (Vol. 5).

Piispanen, R., & Nykyri, T. (1997). Acidification of groundwater in water-filled gravel pits—A new environmental and geomedical threat. *Environmental Geochemistry and Health, 19,* 111–126.

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*(3), Article 3. https://doi.org/10.3390/resources4030637

Pitkänen, T., Karinen, P., Miettinen, I. T., Lettojärvi, H., Heikkilä, A., Maunula, R., Aula, V., Kuronen, H., Vepsäläinen, A., Nousiainen, L.-L., Pelkonen, S., & Heinonen-Tanski, H. (2011). Microbial contamination of groundwater at small community water supplies in Finland. *AMBIO, 40*(4), Article 4. https://doi.org/10.1007/s13280-010-0102-8

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. *Environmental Science & Technology, 47*(23), Article 23. https://doi.org/10.1021/es403489b

Porowski, A., Porowska, D., & Halas, S. (2019). Identification of sulfate sources and biogeochemical processes in an aquifer affected by peatland: Insights from monitoring the isotopic composition of groundwater sulfate in Kampinos National Park, Poland. *Water, 11*(7), Article 7. https://doi.org/10.3390/w11071388

Pritchard, M., Mkandawire, T., & O'Neill, J. G. (2008). Assessment of groundwater quality in shallow wells within the southern districts of Malawi. *Physics and Chemistry of the Earth, Parts A/B/C, 33*(8–13), 812–823. https://doi.org/10.1016/j.pce.2008.06.036

Props, R., Kerckhof, F.-M., Rubbens, P., De Vrieze, J., Hernandez Sanabria, E., Waegeman, W., Monsieurs, P., Hammes, F., & Boon, N. (2017). Absolute quantification of microbial taxon abundances. *The ISME Journal, 11*(2), 584–587. https://doi.org/10.1038/ismej.2016.117

Purkamo, L., Bomberg, M., Nyyssönen, M., Kukkonen, I., Ahonen, L., Kietäväinen, R., & Itävaara, M. (2013). Dissecting the deep biosphere: Retrieving authentic microbial communities from packer-isolated deep crystalline bedrock fracture zones. *FEMS Microbiology Ecology, 85*(2), 324–337. https://doi.org/10.1111/1574-6941.12126

Purkamo, L., Kietäväinen, R., Miettinen, H., Sohlberg, E., Kukkonen, I., Itävaara, M., & Bomberg, M. (2018). Diversity and functionality of archaeal, bacterial and fungal communities in deep Archaean bedrock groundwater. *FEMS Microbiology Ecology*, *94*(8). https://doi.org/10.1093/femsec/fiy116

Purkamo, L., Kietäväinen, R., Nuppunen-Puputti, M., Bomberg, M., & Cousins, C. (2020). Ultradeep microbial communities at 4.4 km within crystalline bedrock: Implications for habitability in a planetary context. *Life*, *10*(1), 2. https://doi.org/10.3390/life10010002

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 Research*, *41*(D1), Article D1. https://doi.org/10.1093/nar/gks1219

Ratnayaka, D. D., Brandt, M. J., & Johnson, K. M. (2009). Groundwater Supplies. In *Water Supply* (pp. 109–147). Elsevier. https://doi.org/10.1016/B978-0-7506-6843-9.00012-3

Ray, C. (Ed.). (2002). *Riverbank Filtration: Understanding Contaminant Biogeochemistry and Pathogen Removal*. Springer Netherlands. https://doi.org/10.1007/978-94-010-0479-4

Ray, C., Grischek, T., Schubert, J., Wang, J. Z., & Speth, T. F. (2002). A Perspective of Riverbank Filtration. *Journal AWWA*, *94*(4), 149–160. https://doi.org/10.1002/j.1551-8833.2002.tb09459.x

Ray, C., Melin, G., & Linsky, R. B. (Eds.). (2003). *Riverbank filtration: Improving source-water quality* (Vol. 43). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-48154-5

Reasoner, D. J., & Geldreich, E. E. (1985). A new medium for the enumeration and subculture of bacteria from potable water. *Applied and Environmental Microbiology*, *49*(1), Article 1. https://doi.org/10.1128/AEM.49.1.1-7.1985

Rivera, A. (Ed.). (2014). *Canada’s groundwater resources*. Fitzhenry & Whiteside.

Robertson, J. B., & Edberg, S. C. (1997). Natural protection of spring and well drinking water against surface microbial contamination. I. Hydrogeological parameters. *Critical Reviews in Microbiology*, *23*(2), Article 2. 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. (2021). RESCRIPt: Reproducible sequence taxonomy reference database management. *PLOS Computational Biology*, *17*(11), 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. *Hydrological Processes*, *29*(13), 2895–2921. https://doi.org/10.1002/hyp.10403

Roslev, P., Bjergbaek, L. A., & Hesselsoe, M. (2004). Effect of oxygen on survival of faecal pollution indicators in drinking water. *Journal of Applied Microbiology*, *96*(5), Article 5. https://doi.org/10.1111/j.1365-2672.2004.02209.x

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. *Journal of Hydrology, 531*, 992–1002. https://doi.org/10.1016/j.jhydrol.2015.11.009

Sallanko, J., Hietala, J., Lindström, R., Kytövaara, A., & Väisänen, T. (2013). More efficient alkalisation with ground limestone in groundwater treatment. *Journal of Water Supply: Research and Technology-Aqua*, *62*(3), Article 3. https://doi.org/10.2166/aqua.2013.025

Salminen, J. M., Nystén, T. H., & Tuominen, S. M. (2011). Review of approaches to reducing adverse impacts of road deicing on groundwater in Finland. *Water Quality Research Journal*, *46*(2), Article 2. https://doi.org/10.2166/wqrjc.2011.002

Salo, H., Mellin, I., Sikkilä, M., Nurminen, J., Äijö, H., Paasonen-Kivekäs, M., Virtanen, S., & Koivusalo, H. (2019). Performance of subsurface drainage implemented with trencher and trenchless machineries. *Agricultural Water Management*, *213*, 957–967. https://doi.org/10.1016/j.agwat.2018.12.010

Sasakova, N., Gregova, G., Takacova, D., Mojzisova, J., Papajova, I., Venglovsky, J., Szaboova, T., & Kovacova, S. (2018). Pollution of surface and ground water by sources related to agricultural activities. *Frontiers in Sustainable Food Systems*, *2*, 42. https://doi.org/10.3389/fsufs.2018.00042

Saxena, G., Bharagava, R. N., Kaithwas, G., & Raj, A. (2015). Microbial indicators, pathogens and methods for their monitoring in water environment. *Journal of Water and Health*, *13*(2), Article 2. https://doi.org/10.2166/wh.2014.275

Scherer, H. W. (2005). Fertilizers and fertilization. In *Encyclopedia of Soils in the Environment* (pp. 20–26). Elsevier Ltd.

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*(10), Article 10. 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. *Journal of Water and Health*, *6*(2), Article 2. https://doi.org/10.2166/wh.2008.022

Singh, S., Pitchers, R., & Hassard, F. (2022). Coliphages as viral indicators of sanitary significance for drinking water. *Frontiers in Microbiology*, *13*, 941532. https://doi.org/10.3389/fmicb.2022.941532

Smith, C. M., & Hill, V. R. (2009). Dead-end hollow-fiber ultrafiltration for recovery of diverse microbes from water. *Applied and Environmental Microbiology*, *75*(16), Article 16. https://doi.org/10.1128/AEM.00456-09

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.

Soveri, J. (1981). On the concentrations of certain substances in the melt water and the groundwater in Finland. *Studies in Environmental Science*, *17*, 777–784.

Soveri, J. (1985). *Influence of meltwater on the amount and composition of groundwater in quaternary deposits in Finland* (Vesientutkimuslaitoksen julkaisuja 63; p. 92). Vesihallitus. National Board of Waters.

Staley, C., Gould, T. J., Wang, P., Phillips, J., Cotner, J. B., & Sadowsky, M. J. (2015). Evaluation of water sampling methodologies for amplicon-based characterization of bacterial community structure. *Journal of Microbiological Methods*, *114*, 43–50. https://doi.org/10.1016/j.mimet.2015.05.003

Stegen, J. C., Fredrickson, J. K., Wilkins, M. J., Konopka, A. E., Nelson, W. C., Arntzen, E. V., Chrisler, W. B., Chu, R. K., Danczak, R. E., Fansler, S. J., Kennedy, D. W., Resch, C. T., & Tfaily, M. (2016). Groundwater–surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. *Nature Communications*, *7*(1), Article 1. https://doi.org/10.1038/ncomms11237

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. *Proceedings of the National Academy of Sciences*, *103*(7), 2363–2367. https://doi.org/10.1073/pnas.0506361103

Storey, M. V., van der Gaag, B., & Burns, B. P. (2011). Advances in on-line drinking water quality monitoring and early warning systems. *Water Research*, *45*, 741–747. https://doi.org/10.1016/j.watres.2010.08.049

Stumpp, C., Brüggemann, N., & Wingate, L. (2018). Stable Isotope Approaches in Vadose Zone Research. *Vadose Zone Journal*, *17*(1), Article 1. https://doi.org/10.2136/vzj2018.05.0096

Tabatabai, M. A. (1974). A rapid method for determination of sulfate in water samples. *Environmental Letters*, *7*(3), Article 3. https://doi.org/10.1080/00139307409437403

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., … Treidel, H. (2013). Ground water and climate change. *Nature Climate Change*, *3*(4), Article 4. https://doi.org/10.1038/nclimate1744

THL. (2022). *Waterborne outbreaks*. https://thl.fi/en/web/environmental-health/water/waterborne-outbreaks

Tian, R., Ning, D., He, Z., Zhang, P., Spencer, S. J., Gao, S., Shi, W., Wu, L., Zhang, Y., Yang, Y., Adams, B. G., Rocha, A. M., Detienne, B. L., Lowe, K. A., Joyner, D. C., Klingeman, D. M., Arkin, A. P., Fields, M. W., Hazen, T. C., … Zhou, J. (2020). Small and mighty: Adaptation of superphylum Patescibacteria to groundwater environment drives their genome simplicity. *Microbiome*, *8*(1), Article 1. https://doi.org/10.1186/s40168-020-00825-w

Tiwari, A. (2020). *Monitoring microbial quality, contaminant sources and bacterial communities of bathing water in Finland* [Unpublished]. http://rgdoi.net/10.13140/RG.2.2.28464.07681

Tiwari, A., Hokajärvi, A.-M., Domingo, J. S., Elk, M., Jayaprakash, B., Ryu, H., Siponen, S., Vepsäläinen, A., Kauppinen, A., Puurunen, O., Artimo, A., Perkola, N., Huttula, T., Miettinen, I. T., & Pitkänen, T. (2021). Bacterial diversity and predicted enzymatic function in a multipurpose surface water system – from wastewater effluent discharges to drinking water production. *Environmental Microbiome*, *16*(1), Article 1. https://doi.org/10.1186/s40793-021-00379-w

Tiwari, A., Kauppinen, A., & Pitkänen, T. (2019). Decay of Enterococcus faecalis, Vibrio cholerae and MS2 coliphage in a laboratory mesocosm under brackish beach conditions. *Frontiers in Public Health*, *7*, 269. https://doi.org/10.3389/fpubh.2019.00269

Treidel, H., Martin-Bordes, J. L., & Gurdak, J. J. (Eds.). (2011). *Climate change effects on groundwater resources: A global synthesis of findings and recommendations* (First issued in paperback). CRC Press.

Tufenkji, N., Ryan, J. N., & Elimelech, M. (2002). The promise of bank filtration. *Environmental Science & Technology*, *36*(21), 422A-428A. https://doi.org/10.1021/es022441j

Väisänen, U. (2004). *Groundwater and factors affecting its quality: Examples from the Rovaniemi district of Northern Finland and western Nicaragua*.

Valentin-Alvarado, L. E., Fakra, S. C., Probst, A. J., Giska, J. R., Jaffe, A. L., Oltrogge, L. M., West-Roberts, J., Rowland, J., Manga, M., Savage, D. F., Greening, C., Baker, B. J., & Banfield, J. F. (2024). Autotrophic biofilms sustained by deeply sourced groundwater host diverse bacteria implicated in sulfur and hydrogen metabolism. *Microbiome*, *12*(1), 15. https://doi.org/10.1186/s40168-023-01704-w

Van Driezum, I. H., Chik, A. H. S., Jakwerth, S., Lindner, G., Farnleitner, A. H., Sommer, R., Blaschke, A. P., & Kirschner, A. K. T. (2018). Spatiotemporal analysis of bacterial biomass and activity to understand surface and groundwater interactions in a highly dynamic riverbank filtration system. *Science of The Total Environment*, *627*, 450–461. https://doi.org/10.1016/j.scitotenv.2018.01.226

Vargha, M., Róka, E., Erdélyi, N., Németh, K., Nagy-Kovács, Z., Kós, P. B., & Engloner, A. I. (2023). From source to tap: Tracking microbial diversity in a riverbank filtration-based drinking water supply system under changing hydrological regimes. *Diversity*, *15*(5), 621. https://doi.org/10.3390/d15050621

Vasconcelos, J., & Harris, S. (1992). *Consensus Method for Determining Groundwaters Under the Direct Influence of Surface Water Using Microscopic Particulate Analysis (MPA)* (EPA 910/9-92-029; Issue EPA 910/9-92-029). USEPA.

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 Research*, 9. https://doi.org/10.1016/j.watres.2017.03.040

Vigneron, A., Cruaud, P., Langlois, V., Lovejoy, C., Culley, A. I., & Vincent, W. F. (2020). Ultra-small and abundant: Candidate phyla radiation bacteria are potential catalysts of carbon transformation in a thermokarst lake ecosystem. *Limnology and Oceanography Letters*, *5*(2), 212–220. https://doi.org/10.1002/lol2.10132

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. *Geophysical Research Letters, 37*(20), Article 20. https://doi.org/10.1029/2010GL044571

Wallin, A. (2016). *Pohjaveden otto talousvedeksi – teknisten, sosiaalisten ja ympäristöön liittyvien riskien ja ongelmien vertailu.*

Wang, H., Knabe, D., Engelhardt, I., Droste, B., Rohns, H.-P., Stumpp, C., Ho, J., & Griebler, C. (2022). Dynamics of pathogens and fecal indicators during riverbank filtration in times of high and low river levels. *Water Research*, *209*, 117961. https://doi.org/10.1016/j.watres.2021.117961

Wei, T. (2017). *Corrplot: Visualization of a correlation matrix. R package.*

WHO. (2011). *Guidelines for Drinking-Water Quality* (4th edn). World Health Organization.

WHO. (2022). *Guidelines for drinking-water quality: Fourth edition incorporating the first and second addenda* (Fourth edition incorporating the first and second addenda). World Health Organization.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (Second edition). Springer. https://doi.org/10.1007/978-3-319-24277-4

Winter, T. C., Harvey, J. W., Franke, O. L., & Alley, W. M. (1998). *Ground water and surface water: A single resource* (Circular) [Circular].

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*(6102), 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. *Journal of Hydrology*, *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 Resources Research, 59*(2). https://doi.org/10.1029/2022WR033056

Zaryab, A., Noori, A. R., Wegerich, K., & Kløve, B. (2017). *Assessment of water quality and quantity trends in Kabul aquifers with an outline for future drinking water supplies*.

Zhou, Y., Kellermann, C., & Griebler, C. (2012). Spatio-temporal patterns of microbial communities in a hydrologically dynamic pristine aquifer. *FEMS Microbiology Ecology*, *81*(1), Article 1. https://doi.org/10.1111/j.1574-6941.2012.01371.x

Original publications

I 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. *Microbiology Spectrum*, *9*(3), e00179-21. https://doi.org/10.1128/Spectrum.00179-21

II 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*, 864, 161199. https://doi.org/10.1016/j.scitotenv.2022.161199

III Lyons, K. J., Yapiyev, V., Lehosmaa, K., Ronkanen, A-K., Rossi, P. M. & Kujala, K. (Manuscript). Assessing the influence of surface water intrusion on shallow groundwater wells using water quality measurements, stable water isotopes, and microbial community analysis.

Original publications are not included in the electronic version of the dissertation.

ACTA UNIVERSITATIS OULUENSIS

SERIES C TECHNICA

948. Järvinen, Topias (2024) Sensor applications of MoS_2 and WS_2 thin films
949. Autio, Anna (2024) Integrated modelling of groundwater-surface water interactions in the boreal landscape
950. Büngener, Lina (2024) The fate of nitrogen and microplastics as wastewater contaminants in treatment wetlands under cold climate conditions
951. Sun, Zhaodong (2024) Facial video-based non-contact physiological signal measurement and applications
952. Tahir, Muhammad Naeem (2024) Advanced vehicular communications through cellular and short-range networks exploiting road weather and traffic observation data
953. Ren, Zhongfei (2024) Adsorption and advanced oxidation/reduction process for elimination of per- and polyfluoroalkyl substances and pharmaceutical pollutants in water
954. Beddiar, Djamila Romaissa (2024) Deep learning-based automatic captioning for medical imaging
955. Zhao, He (2024) Bismuth halide perovskites as photocatalysts for hydrogen production
956. Kaksonen, Rauli (2024) Transparent and tool-driven security assessment for sustainable IoT cybersecurity
957. Yastrebova-Castillo, Anastasia (2024) LEO satellite constellations for autonomous system operation in the Arctic region : situational awareness and communication aspects
958. Airaksinen, Susanna (2024) Oxide scale formation of stainless steel in transition of annealing and reheating towards fossil-free methods
959. Bahador, Nooshin (2024) Assessment of neurological function with multimodal and multichannel physiological signal analysis using machine and deep learning techniques
960. Akbar, Rehman (2024) CMOS integrated wideband, flexible and scalable beamforming architecture
961. Asad Ullah, Muhammad (2024) Machine-type direct-to-satellite communications : modeling and performance analysis
962. Yue, Xin (2024) Lignin processing by deep eutectic solvents : from structural chemistry to its advanced valorization

Book orders:
Virtual book store
https://verkkokauppa.omapumu.com/fi/

UNIVERSITY OF OULU P.O. Box 8000 FI-90014 UNIVERSITY OF OULU FINLAND

ACTA UNIVERSITATIS OULUENSIS

UNIVERSITY OF OULU

ISBN 978-952-62-4210-1 (Paperback)
ISBN 978-952-62-4211-8 (PDF)
ISSN 0355-3213 (Print)
ISSN 1796-2226 (Online)