



# Interpretation of hydrogeochemistry of the Upper Freshwater Molasse (*Obere Süßwassermolasse*) in the Munich area (Bavaria, Germany) using multivariate analysis and three-dimensional geological modelling

Aleksandra Kiecak<sup>1</sup> · Jan Huch<sup>1,2</sup> · Alberto Albarrán-Ordás<sup>1</sup> · Lilian Chavez-Kus<sup>1</sup> · Kai Zosseder<sup>1</sup>

Received: 10 February 2023 / Accepted: 6 December 2023 / Published online: 27 December 2023  
© The Author(s) 2023

## Abstract

Intense use of groundwater in urban areas requires appropriate monitoring, which in turn necessitates proper data management with employment of increasingly sophisticated statistical methods and mapping tools. An example of such an urban area with intensive use of groundwater is the study area of GeoPot Project, namely Munich (Germany) and its surroundings. The aim of the presented study was to provide a description of the hydrogeochemical characteristics of the aquifers occurring in the Quaternary and Upper Freshwater Molasse (German: Obere Süßwassermolasse – OSM) sediments and to further improve the understanding of interactions between the aquifers. The focus was put on the identification of hydrochemical facies, the chemical signatures of different water types, an understanding of occurring processes, and spatial relationships between the aquifers. In order to deal with hydrogeochemical data generated for this study, as well as with data coming from existing external databanks (e.g. BIS-BY), a methodology of quality assurance was developed. The analytical methods focused on multivariate statistics. To enhance the interpretation of the obtained clusters, a recently developed three-dimensional geological model was used for better understanding and presentation. It was found that in the study area, deeper aquifer systems represent the most distinct hydrogeochemical signature of the Na–HCO<sub>3</sub> water type. In the remaining clusters, a transition from deeper (alkaline) to shallow (alkaline-earth) groundwater can be observed. The results of the study can be utilized for improved, sustainable groundwater management.

**Keywords** Urban groundwater · Hydrochemistry · Groundwater statistics · Sedimentary rocks · Germany

## Introduction

Urbanization is taking place all over the world and it is an emerging issue with consequences for the economy, environment and society (Kirabo Kacyira 2012; Schirmer et al. 2013). The trend of population shifting from rural to urban areas will increase in the future and, accordingly, there is a severe need for responsible urban management that considers the availability of resources and sustainable handling of environmental impacts.

In recent years the subsurface has increasingly been recognised as a multifunctional resource providing a wide range of services (Volchko et al. 2020), which are also referred to as (geo) potentials (Team 2003; Bayer et al. 2019; Böttcher et al. 2019). Intensive use of urban underground space is seen as a prerequisite for sustainable development of cities, contributing to the achievement of the 17 UN Sustainable Development Goals (Admiraal and Cornaro 2016; Volchko et al. 2020).

One natural resource prone to susceptibility, and a significant factor in urban management, is groundwater (Vázquez-Suñé et al. 2010). Protection of this resource, including a strategy of groundwater monitoring within a dense network and intensive programme, is becoming progressively more critical. Consequently, the number of monitored sample points and data collection exercises will increase, which will lead to the need (and the opportunity) for comprehensive

✉ Aleksandra Kiecak  
aleksandra.kiecak@tum.de

<sup>1</sup> Technical University of Munich, Arcisstraße 21,  
80333 Munich, Germany

<sup>2</sup> Present Address: GAF AG, Arnulfstraße 199, 80634 Munich,  
Germany

data management and more sophisticated data analyses, and form the basis of sustainable groundwater management.

While, most commonly, hydrogeologists must still struggle with the lack of data and an insufficient number of monitoring wells (e.g. Russo and Taddia 2009; Wakode et al. 2018), in urban areas—as in this research case—the biggest challenge is to handle very large quantities of data, which, in turn, requires meticulous descriptions as well as great efforts on quality control. Modern hydrogeological studies of regional scale are becoming more and more a link between geosciences and data science. Such comprehensive regional studies are crucial for integrated sustainable groundwater management (Raiber et al. 2012), whereas data management and data analysis, two main concepts of data science, involve knowledge of statistics, computer science, etc. (Ma 2018).

Large data records require the application of statistical methods and tools to draw conclusions. An appropriate methodology must be developed in order to filter, process and analyse the datasets (Curtis et al. 2018). Multivariate statistical methods can provide powerful and valuable support in the investigation of hydrogeochemical patterns and processes (Menció and Mas-Pla 2008). However, the outcomes of multivariate statistical analyses cannot be assessed without placing them in a hydrogeological three-dimensional (3D) spatial context (Raiber et al. 2012).

Therefore, developments in geological 3D modelling as well as in hydrogeochemical studies (Raiber et al. 2012), enable a new perspective of looking at geological data. An integrated approach combining multiple lines of evidence (3D geological modelling, multivariate statistics, field observations, etc.) benefitted this study in particular, but also more generally benefits the growing understanding of hydrogeochemical processes (Martinez et al. 2017). Such multidimensional approaches are particularly valuable in resource (respectively groundwater) management, where optimal and efficient use of geopotentials has to be guaranteed, assuring the best possible preservation of resources and sustainable spatial planning.

In this context, the Quaternary glaciofluvial gravel plain and the Miocene fluvio-lacustrine sediments of the Molasse Basin in Germany were the subjects of research studies (Böttcher et al. 2019; Zosseder et al. 2019, 2022; Epting et al. 2020; Böttcher and Zosseder 2022). Groundwater in this area is an important and intensively exploited source for drinking water, service water and for geothermal purposes (e.g. heat pumps; Kerl et al. 2012). Moreover, especially in the urban area of Munich, the Quaternary aquifer has become increasingly polluted, where, e.g. considerable quantities of chlorinated hydrocarbons were detected in the 1980s (Rauert et al. 1993). Besides the use of groundwater, the subsurface is exploited in many other ways—for subways, tunnels, infrastructure, stormwater management, etc. The increasing use of the subsurface carries the necessity

for sustainable management of the subsurface (Griffioen et al. 2014), whereas extensive investigation of geological settings, resulting in their representation in 3D models, is a precondition.

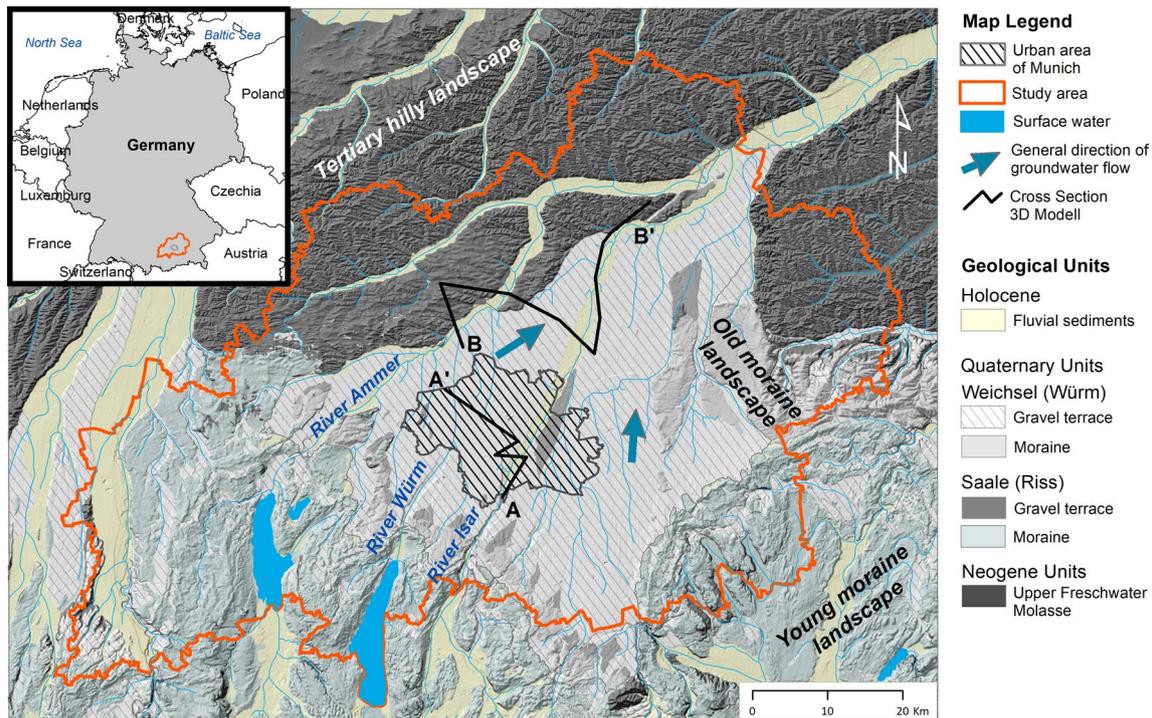
The research projects and in particular the geo-modelling approach by Albarrán-Ordás and Zosseder (2022) led to the development of a geological 3D model for the municipal area of the city of Munich (Albarrán-Ordás and Zosseder 2019, 2020, 2022; Zosseder et al. 2019, 2022). The resulting model predicts the lithological composition in detrital systems consisting of sediment mixtures of different grain sizes. This led to the architectural model representing the interconnectedness of geological bodies with identical lithology in the Quaternary and Molasse deposits (in much greater geological and spatial complexity than previously expected), which is key to identifying and separating geopotentials and optimizing their usage. In particular, it allows an in-depth understanding of the spatial extension of aquifers and their interconnectivity, as well as their lithological configuration, and was a fundamental reference for this study. Hydrogeochemical analysis of the waters within the involved aquifers can be used to validate identified separations or connections of aquifers, which the 3D model is hinting at from a purely spatial side.

The overall aim of this study was to provide a hydrogeochemical characterization of the shallow aquifers in the greater area of Munich with an emphasis on the actual and future utilization of groundwater as a vulnerable resource. In particular, the work involved identification of the hydrochemical facies and the chemical signatures of the different water types, and an understanding of the aquifer processes and spatial relationships between the aquifers. The hydrogeochemical analyses were further used to check the plausibility of the aquifer separation derived from the geological 3D model. The study was performed using different lines of evidence: (1) by applying multivariate statistical methods and (2) relating the outcomes to the geological 3D model by Albarrán-Ordás and Zosseder (2019, 2020) and Zosseder et al. (2022).

## Study area

### Location, geology and hydrogeology

The study area is located in Bavaria (Germany) in the greater area of Munich and covers an area of ~550 km<sup>2</sup> (Fig. 1). The central part of the study area is the Munich Gravel Plain of fluvio-glacial origin, which consists of sandy terraces from Pleistocene glacial periods and modern alluvial and fluvial deposits (Bauer et al. 2006). Coarse-grained carbonate gravels have attained a maximal thickness of 30 m in the city of Munich and up to 60 m in the entire project area, and



**Fig. 1** Location of the study area in the context of geology of southern Bavaria (cross sections A–A' and B–B' are presented in Figs. 12 and 13, respectively)

constitute a very productive porous aquifer (Albarrán-Ordás and Zosseder 2020). The thickness depends on the presence of paleochannel structures carved in the underlying Neogene deposits. Depending on the local geology and the local water table, this aquifer occurs under either confined or (mostly) unconfined conditions. The average hydraulic conductivity is  $5 \times 10^{-3}$  m/s (Böttcher et al. 2019); however, a wide variability in hydraulic conductivity values has been reported (Theel et al. 2020).

In the south, the landscape and geology are dominated by Pleistocene sediments, including pre-Alpine moraines of four glaciations—Würm, Riß, Mindel, and Günz (from young to old)—and corresponding terraces of gravel accumulations (Doppler et al. 2011). The thicknesses vary a lot and reach mostly from a few meters up to 20 or 30 m, reaching in the south even up to 80 m (Zosseder et al. 2022). The sediments are inhomogeneous and comprise boulders, gravels, and sands, as well as clay and silt. Local, mostly isolated and confined porous aquifers are related to lenses or layers of coarse material (Zosseder et al. 2022).

The Upper Freshwater Molasse (German: Obere Süßwassermolasse – OSM) is the upper part of the North Alpine Foreland Basin, also termed the Molasse Basin. The sedimentary rocks range from the middle to the upper Miocene, ca. 17–11 Ma (Kuhleemann and Kempf 2002). The OSM presents tabular geometry and contains sediments of channel deposits (sands, local gravels) and

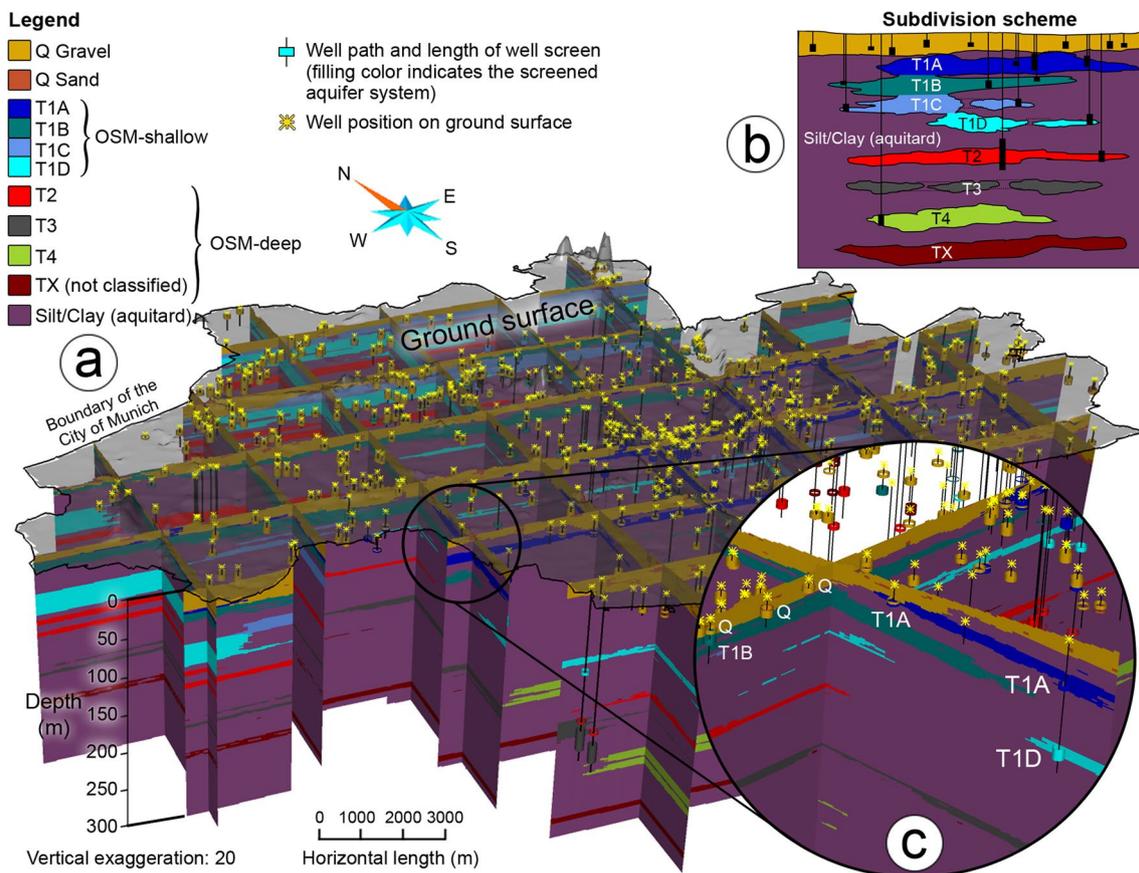
fine-grained material of the alluvial plain (clays, marls; Maurer and Buchner 2007). The mineral composition is comprised of quartz, mica, clay minerals and minor quantities of calcite and dolomite (Rauert et al. 1993). The thickness of the OSM deposits can reach up to 700 m in the project area (southern part of Isar-valley-paleochannel) and generally increases from north to south (Zosseder et al. 2022). The thicknesses of particular groundwater bodies may reach 200 m (Wagner et al. 2003). The OSM sediments are mostly covered by Quaternary deposits. The outcrops of the OSM are generally in the northern part of the study area, which also constitutes the northern margin of the Molasse Basin. Groundwater occurs in sand and gravel layers. Aquifers can be distinguished between “shallow” and “deeper”. This distinction, however, is of practical meaning and not linked to any specific definition of deep groundwater (e.g. as presented by Frape et al. 2003). Shallow aquifers show evidence of anthropogenic impacts and contain relatively young, meteoric water, whereas deeper aquifers are isolated, containing older groundwater. The deeper OSM aquifers are separated and therefore protected from shallow aquifers by silty and clayey sediments, acting as an aquitard (Rauert et al. 1993). The OSM aquifer conditions are mostly confined and the hydraulic conductivity of the OSM layers is two to five orders of magnitude lower than the overlying Quaternary sediment (Jerz 1993; Böttcher and Zosseder 2022).

## Geological 3D models in the greater area of Munich

With the aim of characterizing the existing and potential uses of the subsurface space and the heterogeneity of deposits, the underground space of Munich has been the subject of geological 3D modelling in recent years (Fig. 1, the boundary of the model is almost the same as the urban area of Munich marked in grey). As a result of these works, the “ $D_i$  models” method was conceived as a new methodology for predicting the 3D lithological variability in detrital systems and was proven in a case study in Munich. The details of the model, in particular the input data and sophisticated methodology, can be found in Albarrán-Ordás and Zosseder (2022). The same approach was also implemented in a second model in a broad area located north of Munich (Fig. 1, referred to as Model North; Zosseder et al. 2019). The models were constructed using SKUA-GOCAD™ software (Emerson E&P, St. Louis, USA). The aim of the 3D models was to identify clearly separated aquifers and detect interconnectivity between groundwater bodies.

The focus of the GeoPot project was set as the Munich area, where the need to manage the subsurface resources due to their extensive use is more accentuated, in contrast to its surroundings. Therefore, a detailed lithological and a 3D architectural model was produced for Munich, whereas in the case of the Model North, the priority was given to predicting the lithological composition and no architectural model was implemented.

The first of the models provided the prediction of the lithological composition on a cell-by-cell basis, covering a 310-km<sup>2</sup> area of Munich. This also revealed the complex reservoir configuration presented in the Quaternary and Upper Freshwater Molasse deposits by means of the 3D architectural model (Albarrán-Ordás and Zosseder 2022). Focusing on the OSM sediments, the architectural model showed four extensive coarse-grained geological bodies in Munich that are separated by impermeable beds (Fig. 2). These four geo-bodies are termed T1–T4, from shallower to greater depths, and they host the aquifer systems with the same notation. The aquifer system T1 represents the first OSM groundwater system from the ground surface, having



**Fig. 2** Geological 3D model of the City of Munich with the subdivided aquifer systems Q and T1–T4. **a** 3D fence diagram showing aquifer subdivision through the transparent ground surface; **b** Subdi-

vision scheme; **c** Detailed view of the assignment of aquifer systems to screened wells (modified from Albarrán-Ordás and Zosseder 2019; Zosseder et al. 2019)

various aquifer tiers at different depths (referred to in a previous paper as T1A–T1D; Zosseder et al. 2019; Albarrán-Ordás and Zosseder 2022) and showing multiple hydraulic interaction areas with the Quaternary aquifer throughout the city. The architectural 3D model also indicated that the deeper aquifer systems T2–T4 are considered to be isolated from one another.

The second of the models, Model North, implemented using the  $D_i$  models method, focuses on the area of the Munich Gravel Plain situated immediately to the north of the city. The modelling area extends northwards to the uncovered molasse sediments in the Tertiary Hills region, covering a total surface area of 2,426 km<sup>2</sup> (Fig. 1).

## Methodology

The methodology evolved in this study is presented schematically in a flowchart (Fig. 3) and described in detail in the following sections.

### Data collection

A broad range of hydrogeochemical data derived from groundwater samples taken from boreholes at different depths was collected from the records of the Bavarian Environment Agency (Bayerisches Landesamt für Umwelt, LfU) and several water authorities (e.g. City of Munich). Samples revealing high anthropogenic pollution were excluded

from further assessment. In particular, piezometers located by dump sites, gravel pits or polluted areas were manually selected based on their location and extreme results for certain chemical parameters (mineralisation, pH) compared to local hydrochemical conditions, and consequently expelled. However, diffuse, ubiquitous contamination from agriculture or in the urban areas were not considered as an exclusion criterion.

In total, ca. 98,000 groundwater analyses coming from 7,300 sampling locations were collected. It is important to note that each sampling location represents one well screen section in a borehole. For clarity, hereafter only the term “well” is used, despite the type of borehole (well, piezometer, etc.). Each level of the multilevel groundwater systems was also treated as a separate “well”.

### Additional sampling by the Technical University of Munich (TUM)

In addition to the data collection from reviewed literature and external databases, groundwater sampling was carried out in three sampling campaigns during 2017–2019. Samples were collected at 108 sites, which covered private and municipal drilled wells (32), piezometers (21), and springs (55) of the OSM—Fig. A2 in Section S1 of the electronic supplementary material (ESM). Sampling locations were selected mostly outside the city area, in regions that had previously had a less dense programme of sampling, in order to complement the existing data. In-situ measurements were

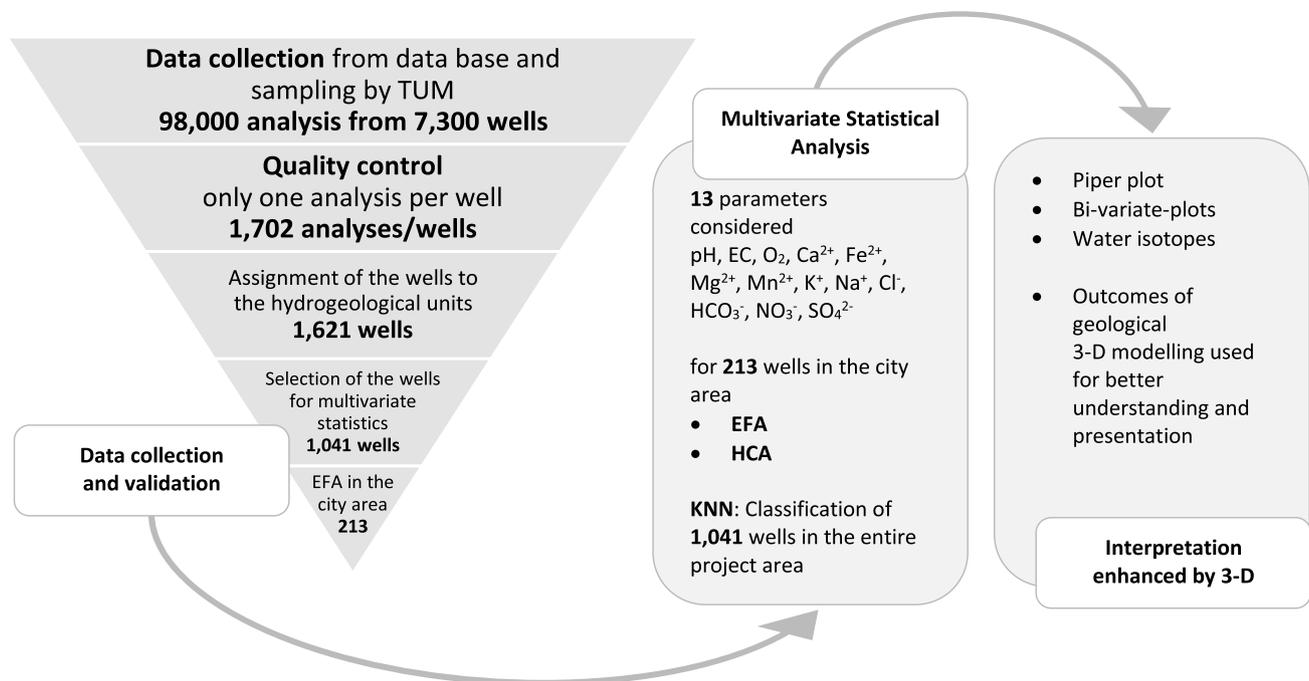


Fig. 3 Methodology flow chart

performed by multiparameter field meters (WTW Multi 3430, Xylem Analytics, Weilheim, Germany) for temperature, pH, electrical conductivity, and dissolved oxygen. Groundwater samples were analysed for 46 parameters, including major and minor ions and trace constituents at the Bavarian Environment Agency. The actual design of the applied analytic methods is described in Chavez-Kus et al. (2016). Additionally, water stable isotopes ( $\delta^{18}\text{O}$  and  $\delta^2\text{H}$ ) were measured at the TUM using a cavity ring-down spectrometer (IWA-45EP; ABB – Los Gatos Research, San Jose, USA).

### Quality assurance and data selection scheme

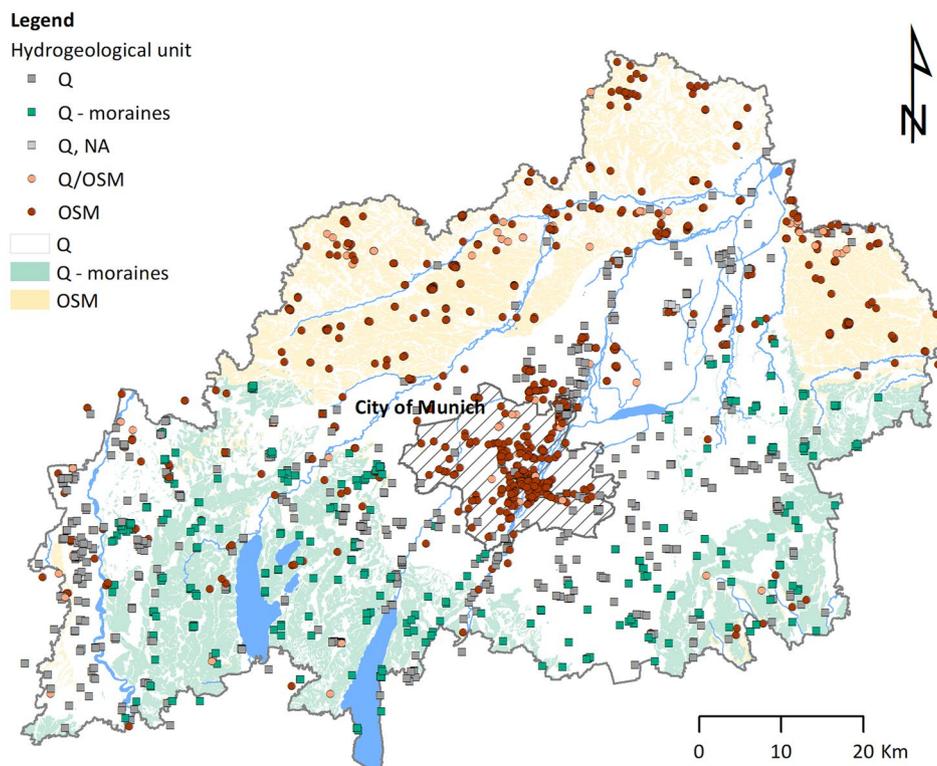
The data set was extremely heterogeneous in many aspects, from the kind of raw data format to the range of analysed parameters. Additionally, the data quality, showing diverse limits of detection, varied considerably over time (data from several decades). Hence, intensive processing was essential prior to further analytical steps.

The analyses were checked for their actuality (more recent than 1978, as proposed in the study of the Bavarian Environment Agency (LfU) by Chavez-Kus et al. 2016), completeness (major ions measured) and quality (charge-balance error  $\leq 10\%$ ). For some wells more than one analysis was available. To ensure the same weighting in spatial statistical assessment and to avoid unwanted clustering effects (e.g. Isaaks and Srivastava 1989), only one analysis for each well

was selected. The selection was performed by considering: (1) novelty (newer analyses were favoured), (2) number of analysed parameters and (3) giving priority to the analyses obtained during a sampling campaign conducted for this study case. The last criterion was introduced for the simple reason that full control over quality was possible for these samples. Finally, 1,702 analyses (coming from the same number of distinct wells) were selected. The data points were well distributed over the whole study area; however, with higher density in the city area of Munich (see Fig. 4).

After careful evaluation of each well (in particular: examination of available geological maps and bore logs, as well as analysing the screen lengths and depths in the geological 3D model), the wells were assigned to the hydrogeological units: (1) Quaternary (612 wells), (2) Quaternary – moraines (320), (3) OSM – shallow (581), and (4) OSM – deep (108). The locations of the wells selected for further assessment are presented in Fig. 4. Additionally, the locations of all available wells along with the number of analyses, as well as the position of sampled points in the aquifers, are shown in Fig. A1 in the ESM. In this step, 81 wells were excluded from statistical analysis, namely the boreholes having more than one well screen section, each of them at different depth intervals, therefore, belonging to more than one hydrogeological unit. This facilitated the description of the characteristics of each hydrogeological unit (results presented in Section S2 in the ESM). In the multivariate analysis, however, the wells capturing more than one hydrogeological unit (or geobody

**Fig. 4** Locations of the groundwater sampling points in the study area



derived from 3D modelling) were saved in order to capture the mixing effects.

## Multivariate statistics

Multivariate statistical methods were employed in order to capture the variability of the great amount of data for selected parameters derived from groundwater found in complex hydrogeological conditions. These multivariate statistical methods included exploratory factor analysis (EFA) and hierarchical cluster analysis (HCA), followed by the application of a machine learning algorithm. The calculations and diagram preparations were performed with the statistical program R (R Core Team 2021) using the packages stats (integrated in R), class (Venables and Ripley 2002), corrplot (Wei and Simko 2021), psych (Revelle 2020), and ggplot2 (Wickham 2009). Besides, for the Piper diagram preparation, the package hydrogeo (English 2017) was used.

EFA is a statistical method that aims to explore the correlation structure among measured variables and identify relationships between them (Goretzko et al. 2019). The dimensionality of variables is reduced by clustering (Hajji et al. 2021), so the data can be summarized in a smaller set of factors for prediction purposes. Because factor analysis uses a correlation matrix, underlying concerns affecting the correlation should be examined prior to the factor analysis: sample size, missing data, normality, linearity, and outliers (Schumacker 2016). Firstly, because it has implications for the selection of the correlation method, the statistical normality of the data was analysed with the Shapiro-Wilk test. The data size was first validated with the Kaiser-Meyer-Olkin test (KMO should be  $>0.5$ ). Next, the Bartlett test of sphericity was performed to identify whether the correlations in the matrix were statistically significant (significance level  $> 0.05$ ; Backhaus et al. 2016). Communality is a measure of how well the variance of each variable is described by a specific set of factors. A measure of sampling adequacy (MSA) shows the proportion of variance in the variables caused by underlying factors (Goretzko et al. 2019). The optimal (lowest) number of factors was determined by generating a scree plot (eigenvalues versus eigenvectors; Conway and Huffcutt 2003). To optimize, variance rotation is applied for better interpretation (e.g. by the oblimin oblique rotation method; Cloutier et al. 2008; Yong and Pearce 2013). As a factor extraction method, maximal likelihood might be preferred (Goretzko et al. 2019).

HCA is a data classification technique, in which the relative positions of all objects in the multidimensional variable space are determined (Odziomek et al. 2017) and distance is used for identification of naturally-occurring groups (clusters). No a priori assumption about the data is made and the classification of the objects into clusters is based on their similarity. In HCA, the squared Euclidean distance metric

and Ward's linkage method were implemented, similar to the methodology presented in Güler et al. (2002) and Cloutier et al. (2008). The Euclidean distance is a measure of similarity performed over all variables included in HCA and is used to identify outlier clusters. Ward's linkage rule creates separate clusters based on analysis of variance and was next applied to link all nonresiduals into distinct clusters (Raiber et al. 2012). The number of clusters was estimated via elbow-, silhouette-, and gap-statistic methods (Kassambara 2015). The HCA results were then visualised in a dendrogram. Cluster analysis allowed grouping of samples in reference to similarities detected by the algorithm.

EFA and HCA were first conducted for the samples in the area Munich, as the most detailed data as well as a detailed 3D architectural model from the geological 3D modelling were available there (Albarrán-Ordás and Zosseder 2020, 2022). A similar methodological approach to the ones presented by Cloutier et al. (2008), Gilabert-Alarcón et al. (2018) and Heine et al. (2021) was implemented. In this part of the study, 13 parameters were considered—pH, electrical conductivity (EC), dissolved oxygen (DO), calcium, iron, magnesium, manganese, potassium, sodium, chloride, hydrogen carbonate, nitrate, and sulphate—observed in 213 wells. The choice of the wells resulted from the prerequisites of the multivariate statistics; in particular, and also to avoid the concerns of missing data, only the wells for which all 13 parameters were available were selected (Cloutier et al. 2008). Moreover, all wells were directly spatially linked to one of the geological bodies identified in the 3D architectural model of the Quaternary and OSM deposits by Albarrán-Ordás and Zosseder (2019, 2020, 2022). This shortened database comprised data from the time period 1992–2019, representing all the geological bodies and aquifer systems as presented in the preceding 3D model study. HCA procedures were applied using the psych package from the R library (Revelle 2020). For the purpose of the multivariate statistical analysis, the parameters with concentrations lower than the detection limit were replaced with a proxy value '0'.

In order to evaluate the characteristics of each cluster more accurately, descriptive statistics were calculated as replenishment for cluster analysis (Ghesquière et al. 2015). The analysis was done on further constituents (e.g. trace elements and water isotopes) not included in the multivariate statistics.

## Classification of data from the greater area of Munich using the *k*-nearest neighbourhood algorithm (KNN)

In order to check if the water types from outside of the city can be grouped in the same water type clusters detected for the city area, the records from Munich's surroundings were fitted to the clusters created for the city. For the classification

of data from the entire project area, a  $k$ -nearest neighbourhood (KNN) algorithm was used. KNN is a supervised machine learning algorithm (Rebala et al. 2019). A total of 828 wells were considered in this step (together with wells from the city area – 1,041 wells, Fig. 3). It is assumed that similar items are closer to each other. The algorithm was first “trained” and tested on the data from the city area and then applied to the data from Munich’s surroundings. The features were scaled to ensure equal contribution while calculating distance. For each row of the test set, the  $k = 1$  nearest training set vectors were found and classified (Venables and Ripley 2002).

### Cluster interpretation enhanced by geological 3D modelling

Firstly, the geological 3D models were used in different forms to enhance the understanding of hydrogeological and hydrogeochemical conditions, by the assignment of hydrogeological units and distinguishing between shallow and deeper wells. A set of 2D cross sections along the wells in both Munich and in the area situated to the north of the city (Model North) was selected and generated from the models. These 2D representations show the predicted prevailing lithologies in the vicinity of the well screens where the samples for this study were collected. Analysis of this picture of the continuity of the lithologies resulted in the assignment of the aquifers “OSM-shallow” and “OSM-deep” to the screened well sections. This was inferred individually for each well screen along the 2D profiles.

For simplification purposes, within the present study, the aquifer systems in the OSM are sorted into one of the following two notations—OSM-shallow or OSM-deep. The first group denotes the aquifer system T1, which presents wide interaction areas with the overlying Quaternary aquifer, whereas the second group refers to the aquifer systems T2, T3, and T4, which have no hydraulic interlinkage to other shallow aquifers. The distinction between the aquifer systems T1–T4, i.e. the 3D architectural model, was implemented in 3D only in the city area. However, in order to facilitate the cluster interpretation with depth also in the northern part of the study area, an interpretation of the OSM-shallow and OSM-deep aquifers along the cross section in Model North is provided.

## Results

### Multivariate statistics for the Munich area

The statistical normality of the data, analysed with the Shapiro-Wilk test, showed that no variable represented the normal distribution. Therefore, a correlation matrix was

calculated with the Spearman method because it is nonparametric and used to perform rank-based correlation analysis (Kassambara 2015).

Relevant relationships among the variables are presented as a correlation matrix (Fig. 5), which allows one to distinguish the following relationships: (1) strong positive correlations among chloride, sulphate, nitrate, EC and calcium (0.61–0.83), (2) strong positive correlations between DO and nitrate (0.65), DO and calcium (0.58), iron and manganese (0.53), carbonates and calcium (0.58), and (3) negative correlation of pH and calcium, EC, nitrate, chloride or carbonates (0.52–0.74).

The KMO test gave a moderate overall measure of sampling adequacy (MSA) of 0.71, indicating that the sample size is adequate for factor analysis. In the Bartlett test, high values of chi-squared ( $\chi^2 = 2,068$ ,  $p$ -value  $\approx 0$ , 78 degrees of freedom) were achieved, indicating that statistically significant correlations exist within the matrix. The determinant of the correlation matrix was positive. Internal consistency reliability was measured by Cronbach’s alpha reliability (Revelle 2020) and demonstrated to be good (0.72).

The ideal number of factors in this case was determined by three methods. The scree-elbow plot of eigenvalues and the empirical Kaiser criterion were in favour of four factors, whereas the parallel analysis favoured five. Finally, by trial and error, a number of three was chosen, giving the most satisfying results: minimum number of variables per factor of four and item to factor ratio of 4.3:1.

Maximum likelihood (ML) was implemented as a factoring method and an oblimin oblique rotation method, allowing the resulting factors to be correlated. The Bartlett method was used for factor scores.

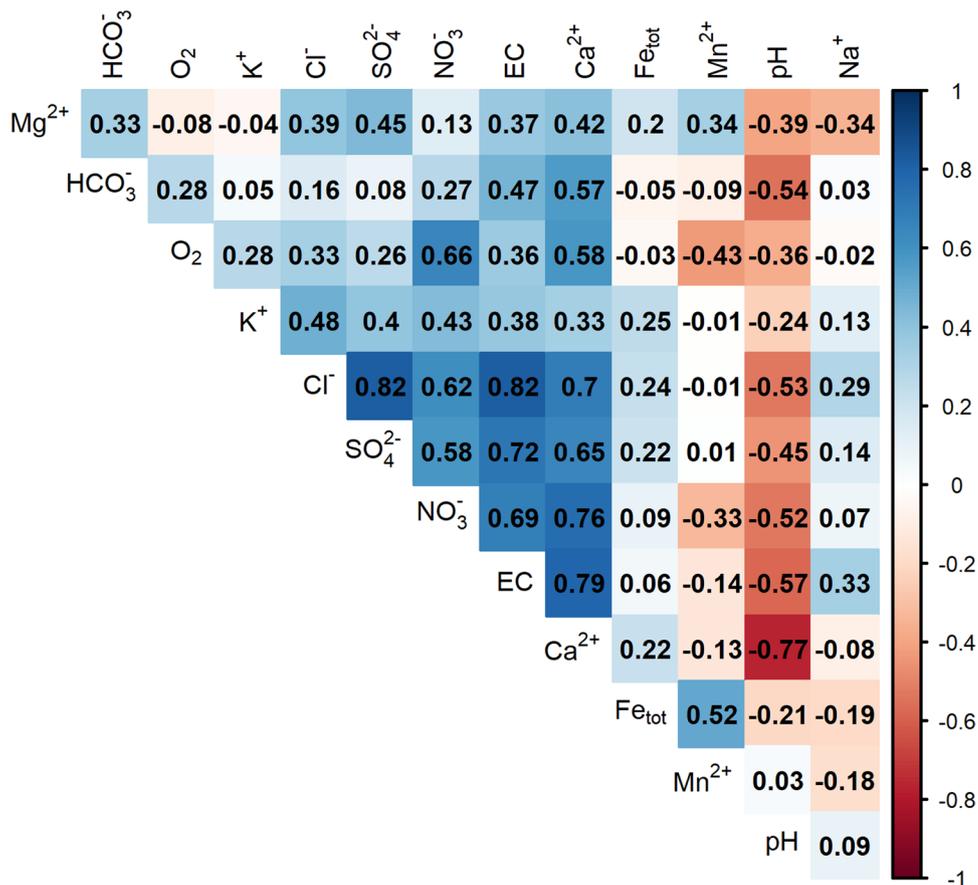
Communalities of individual parameters vary between 0.26 and 1.00 (Table 1). The factor loadings and factor scores for the first two factors obtained in the EFA are presented in Fig. 6 (see also Fig. A3 in the ESM). The three factors explain 60% of the total variance; moreover, factors ML1 and ML2 are correlated at 0.47 (Fig. A4 in the ESM).

Factor ML2 explains 27% of the variance and shows the highest loading ( $>0.7$ ) for positive correlated parameters:  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$  and EC, and also for  $\text{Na}^+$ ,  $\text{K}^+$  and  $\text{NO}_3^-$ . This factor matches the concentrations of anions  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$ , and cations  $\text{Na}^+$ ,  $\text{K}^+$ , to the EC, which is known to be related to the total mineralisation.

DO,  $\text{Ca}^{2+}$  and  $\text{HCO}_3^-$  load high regarding factor ML1. Also,  $\text{NO}_3^-$  and DO load secondarily to this factor, as well as pH (negatively). The opposite loadings for  $\text{HCO}_3^-$  and pH possibly imply that carbonate solubility is controlled by pH (Newman et al. 2016); moreover, the opposite loadings of  $\text{Ca}_2^+$  and  $\text{Na}^+$  reveal the importance of the cation exchange process.

$\text{Mn}^{2+}$  has the highest factor loading on the third factor ML3. Similarly,  $\text{Fe}_{\text{tot}}$  and  $\text{Mg}^{2+}$  load to ML3. This factor

**Fig. 5** Correlation matrix of 13 physico-chemical parameters



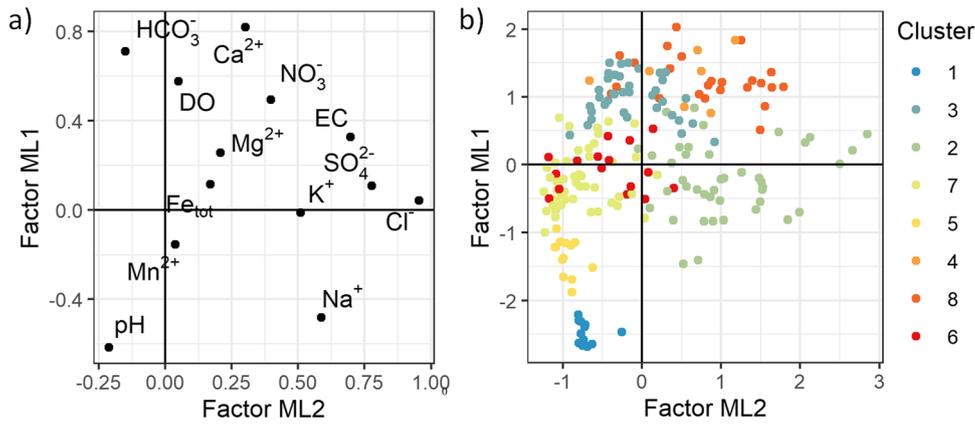
**Table 1** Loadings of the factor matrix (*italic*: loadings > 0.5)

Parameter	Factor ML2	Factor ML1	Factor ML3	Communality	Uniqueness	Complexity
pH	-0.21	<i>-0.62</i>	-0.13	0.57	0.434	1.3
EC	<i>0.7</i>	0.33	-0.06	0.81	0.186	1.4
DO	0.05	<i>0.58</i>	-0.43	0.54	0.463	1.9
Ca <sup>2+</sup>	0.3	<i>0.82</i>	0.01	1	0.005	1.3
Mg <sup>2+</sup>	0.21	0.26	<i>0.56</i>	0.47	0.527	1.7
Na <sup>+</sup>	<i>0.59</i>	-0.48	-0.33	0.44	0.555	2.5
K <sup>+</sup>	<i>0.51</i>	-0.01	-0.07	0.26	0.739	1
Fe <sub>tot</sub>	0.17	0.12	<i>0.52</i>	0.33	0.671	1.3
Mn <sup>2+</sup>	0.04	-0.15	<i>0.75</i>	0.58	0.421	1.1
Cl <sup>-</sup>	<i>0.95</i>	0.04	0.09	0.95	0.051	1
HCO <sub>3</sub> <sup>-</sup>	-0.15	<i>0.71</i>	0	0.43	0.571	1.1
SO <sub>4</sub> <sup>2-</sup>	<i>0.78</i>	0.11	0.14	0.71	0.295	1.1
NO <sub>3</sub> <sup>-</sup>	0.4	<i>0.5</i>	-0.28	0.67	0.331	2.5

may be connected to the redox conditions in the aquifers, especially with respect to the (admittedly low but not negligible) negative loading of DO.

The estimated number of clusters was diversified depending on the method used (6, elbow method, 4, silhouette method, 9, gap statistics). Therefore, finally, the number was determined arbitrarily on eight clusters and these clusters

were the basis for the proceeding interpretation. The cluster analysis allowed grouping of the groundwater samples in reference to similarities in their physicochemical characteristics. In the first attempt at building clusters (not shown), five outliers were identified and excluded from further analysis. The dendrogram in Fig. 7 concerns the remaining 208 data points. This graphical representation of HCA as



**Fig. 6** Plots of **a** factor loadings and **b** factor scores, for the first two factors (ML1 and ML2) obtained in the exploratory factor analysis (EFA) identified by clusters derived from hierarchical cluster analysis (HCA) (see Fig. 7 and cluster description in text)

a dendrogram is presented in combination with pie charts showing the major ions contents.

**Classification of data from the greater area of Munich using the *k*-nearest neighbourhood algorithm (KNN)**

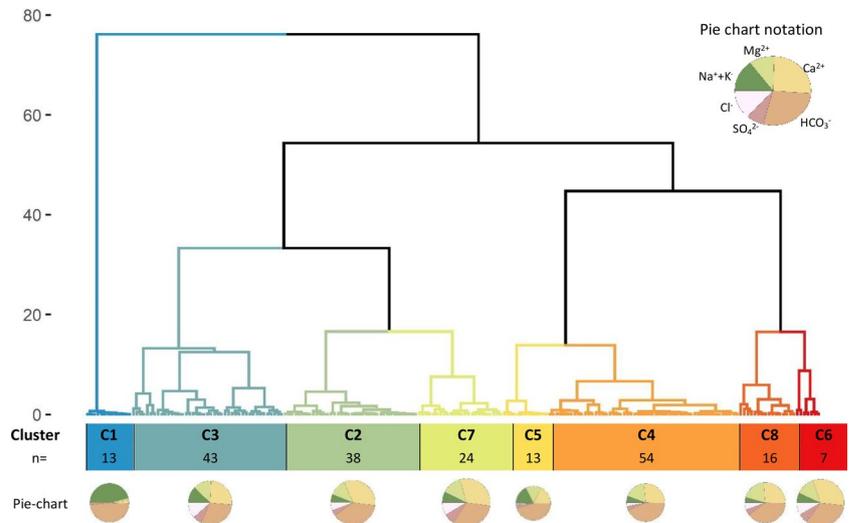
The clusters built for the city area, resulting from the previous step, were used for the classification of the wells from the entire project area. The classification was performed with the KNN-algorithm. The *k*-value, which indicates the count of the nearest neighbours, was found in an iterative selection process. For *k* = 1 the accuracy score of KNN equal to 80 % was found to be satisfactory (95% confidence interval: 69.14, 88.78), with *p*-value < 2.2e-16 and Kappa = 0.7598; hence, KNN was further used to classify the wells.

**Water isotopes**

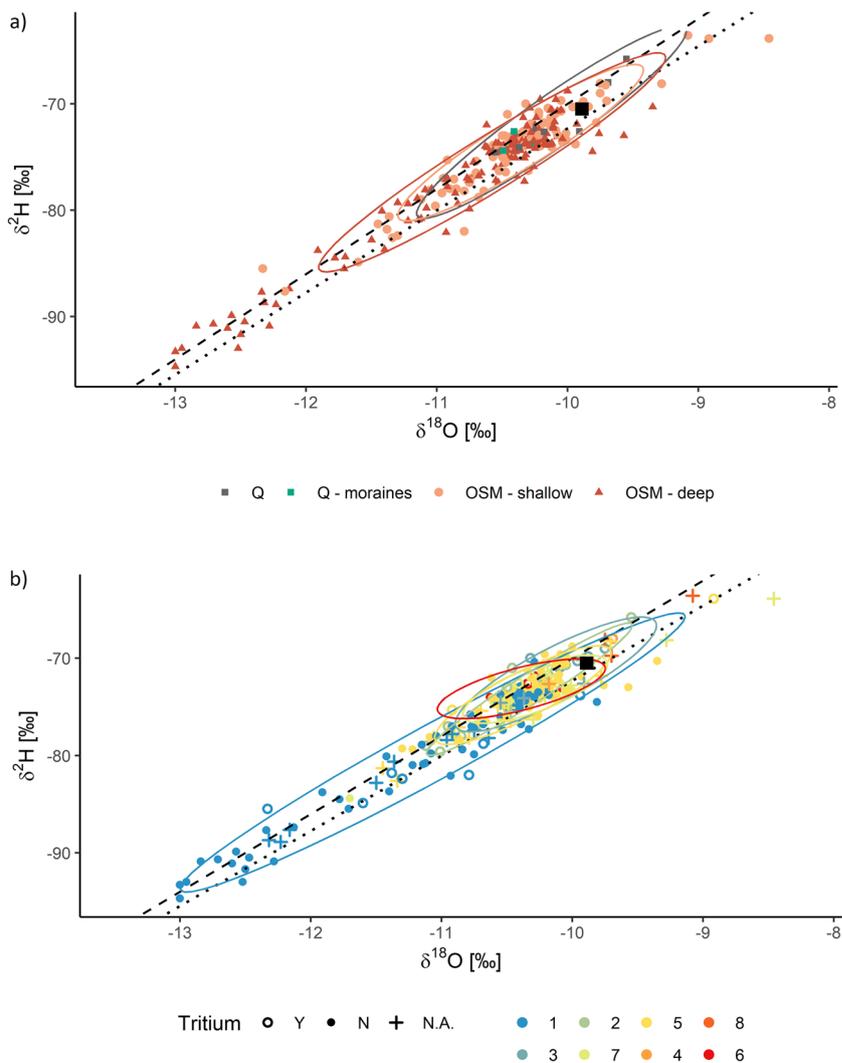
Stable isotopes of oxygen and hydrogen in groundwater (expressed as isotope ratios  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$ , respectively) are used, among others, to analyse groundwater flow pathways, including paleo-groundwater and mixing processes (Tweed et al. 2019). For instance, the values of  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$  for water infiltrating to the subsurface in the Pleistocene period were several per-mill lower than the present (Tweed et al. 2019). Additionally, measurements of the radioactive hydrogen isotope tritium ( $^3\text{H}$ ) provide further insight into the mean age of groundwater (Małozzewski et al. 1983; Zuber et al. 2001).

Figure 8 presents the stable isotope values, as well as the Global Meteoric Water Line (GMWL) and local meteoric water line (LMWL) calculated for Germany (Stumpp et al. 2014), with the formulae:

**Fig. 7** Dendrogram and pie charts for median values in the resulting clusters (size of the pie chart depends on the degree of mineralisation)



**Fig. 8** Plots of  $\delta^2\text{H}$  versus  $\delta^{18}\text{O}$ : **a** for hydrological units, **b** for clusters (colours indicate the clusters, for details see Fig. 7 and cluster description in text; some wells are represented by more than one data point. Notation: dashed line – GMWL, dotted line – LMWL for Germany (Stumpp et al. 2014); the ellipses are normal confidence ellipses calculated for each cluster; the black square represents precipitation after IAEA/WMO (2020) – long-term mean of precipitation measured in Neuherberg near Munich; NA – tritium not analysed



$$\text{GMWL} : \delta^2\text{H} = 8 \cdot \delta^{18}\text{O} + 10 \tag{1}$$

$$\text{LMWL} : \delta^2\text{H} = 7.72 \cdot \delta^{18}\text{O} + 4.9 \tag{2}$$

The observations plot mostly around these lines.

For comparison, an isotopic value for precipitation is presented. Long-term means for stable water isotopes in precipitation were calculated for the measurements from Neuherberg (near Munich) and have the following values:  $\delta^{18}\text{O} -9.89 \pm 0.78\text{‰}$ ,  $\delta^2\text{H} -70.5 \pm 5.8\text{‰}$ , with d-excess  $8.6 \pm 1.3\text{‰}$  (IAEA/WMO 2020).

Groundwater in the study area showed water stable isotope values between  $-13.2$  and  $-8.9\text{‰}$  for  $\delta^{18}\text{O}$  and between  $-94.7$  and  $-63.9\text{‰}$  for  $\delta^2\text{H}$  (Fig. 8a). Water samples from Quaternary aquifers (Q and Q-moraines) plot mostly close to these values, revealing a relatively young age of groundwater in these compartments. The point-cloud of OSM-shallow is spread towards lower values of

$\delta^{18}\text{O}$  and  $\delta^2\text{H}$ , suggesting that water age reaches Pleistocene, since groundwater recharge occurred during cooler conditions from melt water or precipitation (van Geldern et al. 2014). This trend is even more visible for the wells from deeper OSM levels. The highest  $\delta^{18}\text{O}$  and  $\delta^2\text{H}$  values may be explained by evaporation, either natural (subsurface influx of lake water in the south of Munich) or anthropogenic (industrial cooling water) (Rauert et al. 1993). The variability of  $\delta^{18}\text{O}$  and  $\delta^2\text{H}$  values is greatest in cluster 1 and may reflect a spectrum of groundwater ages or mixing processes (Rauert et al. 1993). Data points of the remaining clusters concentrate mostly around the values of modern precipitation.

Regarding the tritium content (Fig. 8b; Fig. A8 in the ESM), no occurrence was found in most of the wells of the OSM-deep. However, even in some samples associated with cluster 1, typical for deeper aquifers, tritium was observed, suggesting mixing processes. Interestingly, in some wells of

**Table 2** The number of objects in each cluster and hydrogeological units in the entire study area

Hydrological unit	Cluster								Total no. of objects
	C1	C3	C2	C7	C5	C4	C8	C6	
Q	-	65	155	28	2	96	4	5	355
Q-moraines	-	23	57	10	3	81	-	1	175
Q/OSM	-	18	13	4	6	17	-	9	67
OSM-shallow	3	92	39	35	41	131	28	-	369
OSM-deep	33	-	-	-	32	4	6	-	75
Total	36	198	264	77	84	329	38	15	1,041

shallow OSM and even in Quaternary moraines, no tritium was measured, revealing water age of more than ca. 70 years (recharged prior to the start of above-ground nuclear bomb testing in 1953), suggesting that these units in some parts of the research area are well isolated from the surface.

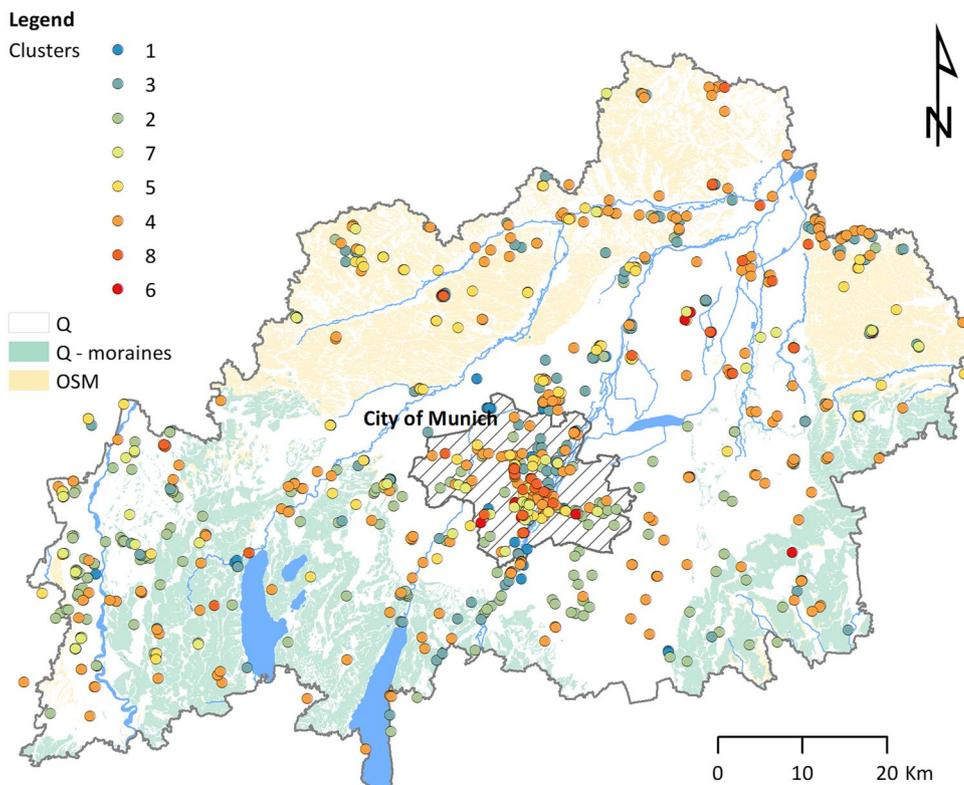
**Cluster description**

As described in section ‘Multivariate statistics for the Munich area’, the primary focus was on the city area, where a detailed 3D architectural model of the Quaternary and OSM deposits was available. A map presenting locations of the wells in the Munich city area (Fig. A5 in the ESM) and a Piper diagram for those sites (Fig. A6 in the ESM), as well

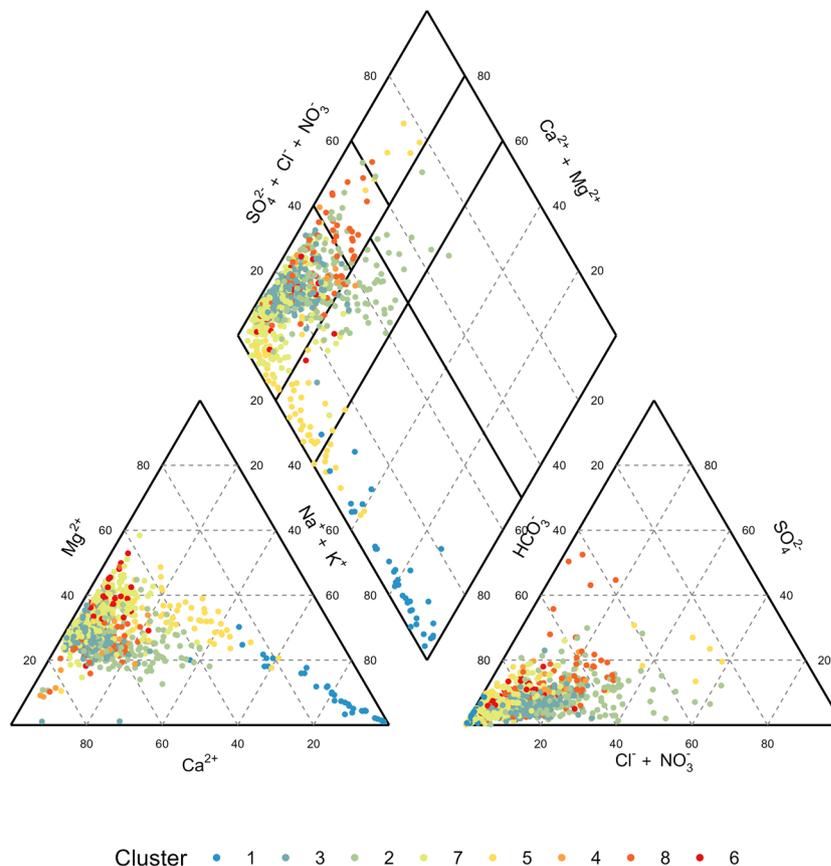
as a table of the number of wells in each aquifer system and cluster (Table A1 in the ESM) are provided.

The clusters described in the following relate to the entire area. A general overview of all observations is given in Table 2 and Figs. 9 and 10, which present by analogy the number of wells in each aquifer and cluster, the locations of the wells within the study area, and the Piper plot. Additionally, the box-plots obtained for the clusters are shown in Fig. A7 in the ESM. In Figs. 9 and 10, the groundwater samples are coloured according to their cluster. It can be seen in Fig. 9 that clusters do not build any spatial, regional pattern in a 2D view; therefore, in the following, they are described in accordance with their occurrence in a 3D geological system. Additional information for cluster interpretation can be concluded from the bi-variate plots (Fig. 11).

**Fig. 9** The locations of wells assigned to clusters in the study area (colours indicate the clusters, for details see Fig. 7 and cluster description in text)



**Fig. 10** Piper plot presenting clusters in the study area (colours indicate the clusters, for details see Fig. 7 and cluster description in text)



The cluster interpretation provided below presents various examples by means of the two geological 3D models implemented in the greater area of Munich. As mentioned in the section ‘[Geological 3D models in the greater area of Munich](#)’, the models refer to (1) the city of Munich (Fig. 12), and (2) the area extending from the north of the city to the Tertiary Hills (Fig. 13). Selected examples are indicated by numbers in Figs. 12 and 13 and described in the text.

### Deeper aquifers

*Cluster C1* is related to well-isolated aquifer systems, with distinct chemistry, representing Na–HCO<sub>3</sub> water type. Due to high concentrations of bicarbonate and very small sulphate and chloride concentrations, as well as high Na content (98 mg/L), the points plot in the lower corner of the diamond on the Piper Plot (Fig. 10). In addition, Ca and Mg content are the lowest. The waters are also characterised by elevated pH (8.3–9.1), and mostly low DO. The data points of C1 (in opposite to all the remaining clusters), shown in Fig. 11b, do not follow the theoretical linear relationship of calcite dissolution (Ghesquière et al. 2015), which indicates that other processes lead to HCO<sub>3</sub><sup>-</sup> enrichment.

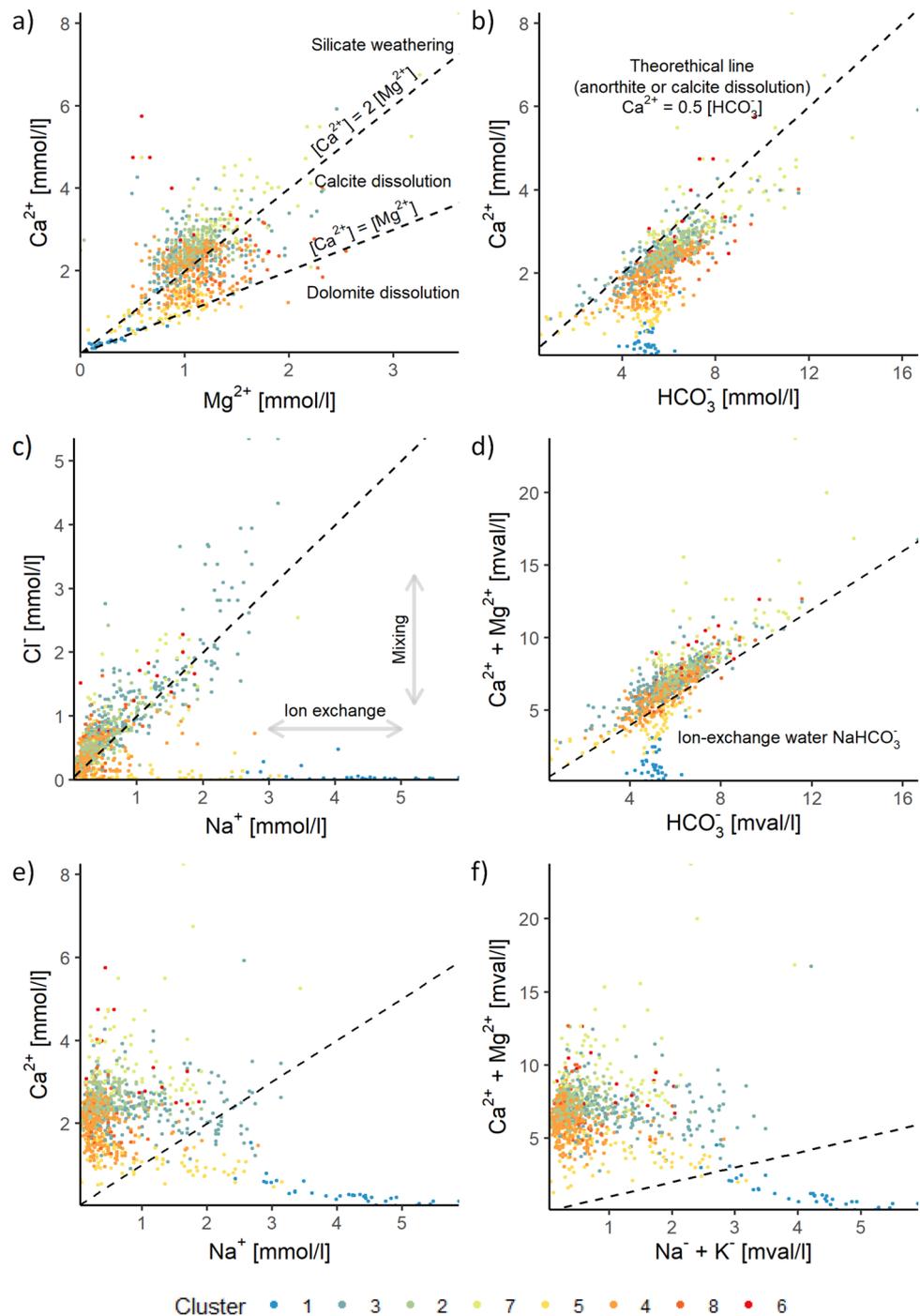
The majority of deeper wells in the city belong to cluster C1, but also some well-isolated wells screened in the deeper

parts of the shallow aquifer system (particularly T1D), in the south of the city (see labels 2, 3, and 4 in Fig. 12). All the aforementioned wells are screened in the deeper aquifer system (marked in green in Fig. 12), well isolated from the overlying shallow groundwater system (marked in grey in Fig. 13). However, at well 1 (Fig. 12) at the depth of between 440 and 500 meters above sea level (masl), it can be observed that the well is characterized by cluster C1, although it is screened in the lower part of the OSM-shallow groundwater system (more specifically, in the aquifer system T1). This part of the shallow Neogene aquifer is very isolated from the remaining overlying aquifer tiers in the southern part of Munich, which would explain the resulting assignment of cluster C1, which typically represents deeper, clearly isolated groundwater systems. In the Munich North model, cluster C1 is observed in wells 14–18 (Fig. 13).

The isolation of C1-waters is also confirmed by the values of stable water isotopes, which present much lower values than modern waters, suggesting that they were recharged during colder than present climatic conditions (Fig. 8). The majority of water samples are also free of tritium.

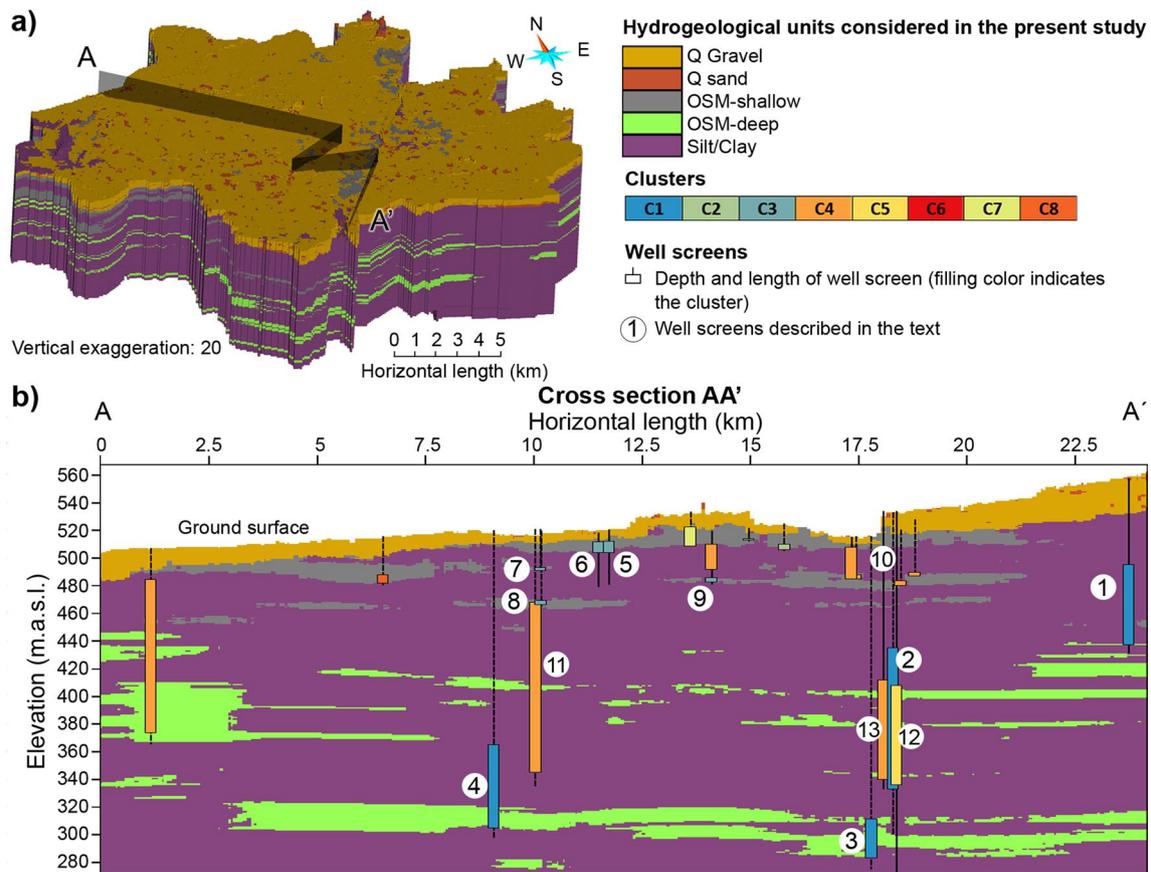
Ortho phosphate and fluoride concentrations are the highest in the deeper aquifers. Some trace elements (B, Ba, Li, Mo, Sr, Ti) show higher concentrations in these deeper-level samples than in shallower-OSM samples or in

**Fig. 11 a–f** Bi-variate plots. Notation: dashed line is 1:1 line, when not described differently; colours indicate the clusters, for details see Fig. 7 and cluster description in text



the Quaternary. The opposite case is observed for uranium, for which concentrations were the lowest among the four hydrogeological units ( $1.3 \mu\text{g/L}$ ). Concerning other trace elements, selenium and strontium analyses also reveal relatively high concentrations, which can be explained by the occurrence of weathered feldspars in the OSM sediments (Kainzmeier et al. 2007). In previous studies also from OSM, but located east of the study area (region 13-Landshut

and 18-Südostbayern), the waters with low EC and DO, and high pH and specific conductivity were described as ion exchange waters (Kainzmeier et al. 2007; Chavez-Kus et al. 2016). The authors related these waters to uprising Malm waters and recalled that higher concentrations of tracer elements (B, Cz, Li, Rb, Th, Ta) present in some OSM wells are typical for Malm-waters that have undergone ion exchange.



**Fig. 12** Cluster interpretation for the city of Munich enhanced by geological 3D modelling: **a** 3D view of the hydrogeological units considered in the present study adapted from the 3D architectural

model of Albarrán-Ordás and Zosseder (2022); **b** Cross section A–A' showing the hydrogeological units and the cluster assignment by various examples of well screens

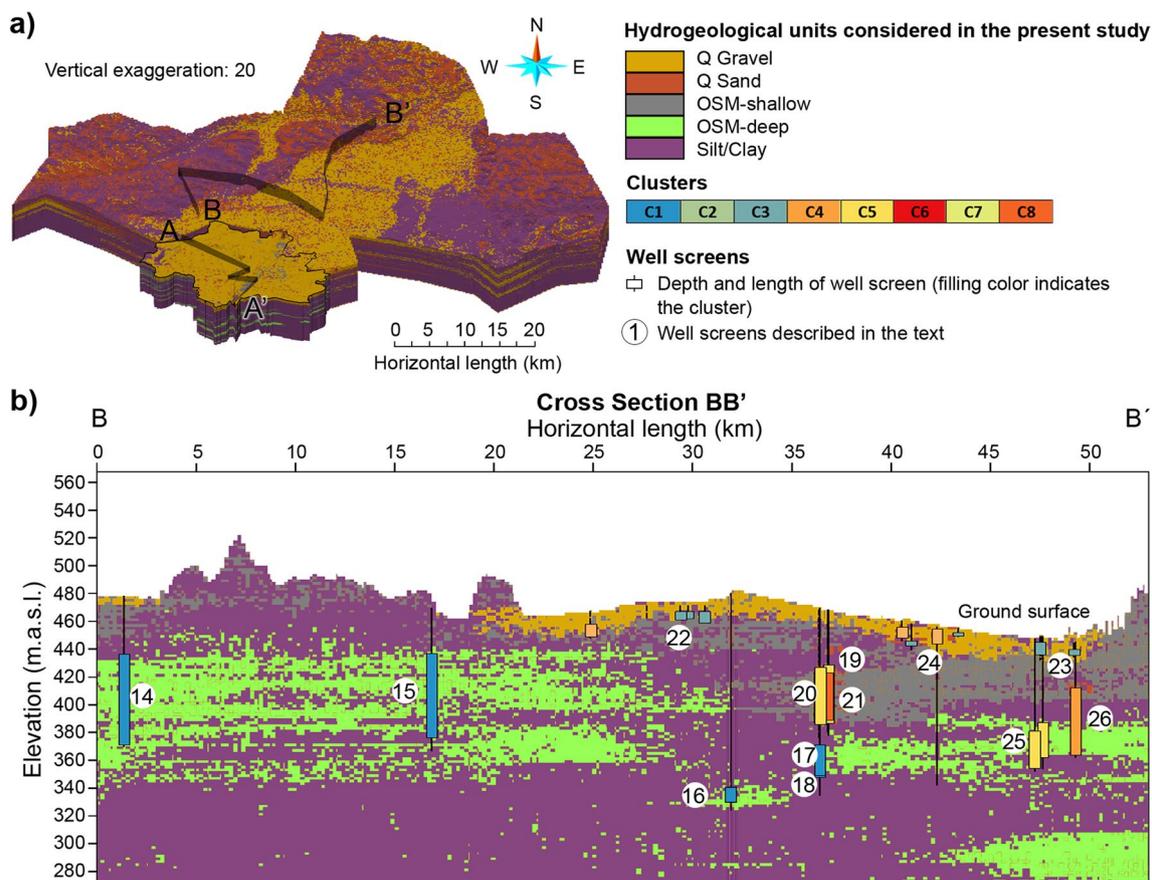
**Transition zone**

A transition from deeper to shallow groundwater may be observed in the remaining clusters. With increasing depth, the content of Ca and Mg decreases and the content of Na decreases due to ion-exchange processes (Rauert et al. 1993). Cluster C5 represents mixed conditions, which is also apparent from the Piper diagram (Fig. 10), where the points representing this cluster plot between C1 and all the other points. The transition then goes further, but not as obvious, from C1 and C5 towards C7, followed by C3 and finally C4. The shift is also visible on the stable water isotopes plot (Fig. 8b).

Cluster C5 is similar to C1, except in Na content, and is characterised by low chloride, low DO and low EC. The pH values are relatively elevated, but not as distinctly as in C1. Ion exchange is the main process leading to increase of Na concentration by constant Cl content in both clusters (Fig. 11c, in contrast to the remaining clusters, where mixing and ion exchange are balanced, with small domination of mixing). Mg<sup>2+</sup> and Ca<sup>2+</sup> concentrations are governed predominantly by calcite weathering in cluster C5 (Fig. 11a).

The ion-exchange water of Na–HCO<sub>3</sub>-type in clusters C1 and C5 is also apparent in Fig. 11d, as data points indicate that  $c(\text{Ca}^{2+} + \text{Mg}^{2+}) < c(\text{HCO}_3^-)$ . In Germany, this type of water is exclusively associated with ion-exchange processes, because there are no deposits of this kind of salt in the subsurface (Hölting and Coldewey 2013). Ion exchange is also an explanation for the deficit of Ca<sup>2+</sup> in comparison to Na<sup>+</sup> in clusters C1 and C5 as presented in Fig. 11e (Appelo and Postma 2004; Ghesquière et al. 2015). The tendency is noticeably even clearer in Fig. 11f, which shows  $c(\text{Ca}^{2+} + \text{Mg}^{2+})$  in comparison to  $c(\text{Na}^- + \text{K}^-)$ .

The wells from this cluster are mostly screened in isolated aquifer systems; however, not as deep as C1, and the influence of younger water is noticeable. Cluster C5 lies in transition zones between deeper (C1) and shallow aquifer systems; therefore, groundwater samples of this cluster show a broader typification from Ca–HCO<sub>3</sub> to Na,K–HCO<sub>3</sub> waters. In this sense, it was observed that wells 19 and 20 (Fig. 13), which are screened in the gravel-dominant aquifer system at depths between 380 and 420 masl, belong to cluster C5, whereas the underlying close wells 17 and 18 were



**Fig. 13** Cluster interpretation in the northern part of the study area enhanced by geological 3D modelling: **a** 3D view of the prevailing lithological classes (gravel, sand and clay/silt); **b** Cross section B–B'

showing the interpreted hydrogeological units and the cluster assignment by various examples of well screens

assigned to cluster C1. The mixed conditions at wells 19 and 20 can be explained by the proximity of an interaction zone between the Quaternary aquifer and the first shallow groundwater system in the OSM (from ca. 40–50 km on cross-section B–B'). It can also be observed that the aforementioned well groups 19–20 and 17–18 are separated by a fine-grained succession, thus causing an isolation of the deeper groundwater assigned to cluster C1, which is typical for isolated groundwater. Cluster C5 is also visible in four wells situated at location 25. In this case, the mixed conditions can be explained by the lateral extension of the gravel aquifer at depths of 400 masl. It can also be observed that well 26, with a long well screen reaching the uppermost OSM groundwater system, belongs to cluster C4 and represents an exception to this interpretation.

### Shallow aquifers

Clusters C3, C2 and C7 comprise wells screened in the Quaternary and shallow OSM. The dendrogram reveals

some indications of similarity between these three clusters (Fig. 7). These are characterised by similar pH and EC values, as well as concentrations of  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{HCO}_3^-$ , and  $\text{SO}_4^{2-}$ . Cluster C2 is characterised by the highest DO. C3 has a higher  $\text{Cl}^-$  median concentration than C2 or C7, also with outliers.

Possibly, the anthropogenic impacts (mostly diffuse contamination sources, i.e. use of fertilizers or road-salt in winter) are visible through elevated contents of nitrate, chloride, potassium, and sodium (Wagner et al. 2003, 2011)—for example, German limit-values for nitrate in drinking water (>50 mg/L) were exceeded in 30 analysed samples for C2, 18 for C3, and 15 for C7. Elevated  $\text{NO}_3^-$  concentration (including several outliers) indicates use of fertilizers in the vicinity of the wells. Elevated contents of  $\text{SO}_4^{2-}$  and  $\text{Mg}^{2+}$  are also indicators of anthropogenic contamination due to agricultural activity and wastewater release (Senbayram et al. 2015; Torres-Martínez et al. 2020). Previous studies conducted in the immediate vicinity of study area have shown that elevated concentrations of nitrate are caused by release of high amounts of manure and, to a lesser extent,

synthetic fertilizers due to intensified agriculture (Wild et al. 2018, 2020).

Cluster C6 is similar to C7, but with lower DO and  $\text{NO}_3^-$ . These clusters almost overlap each other on the Piper plot (Fig. 10) and factor scores plot (Fig. 6). Cluster C7 belongs to shallow wells, screened in areas with hydraulic interaction between the Quaternary aquifer and the underlying aquifer parts of the OSM. Cluster C6 is a relatively small group of 15 wells screened at shallow depths. EC medians are the highest among all the clusters (870 and 878  $\mu\text{S}/\text{cm}$ , respectively). In cluster C6, elevated Fe and Mn concentrations are noticed; moreover, the DO is relatively low, which suggests reductive conditions, as confirmed via EFA (see section ‘Classification of data from the greater area of Munich using *k*-nearest neighbourhood algorithm (KNN)’). Iron and manganese exceedances are generally associated with each other, as they have similar chemical behaviour, being subject to pH and redox conditions (Montcoudiol et al. 2015).  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  concentrations are governed predominantly by silicate weathering in clusters C6 and C7, as well as in C2 and C3 (Fig. 11a).

Clusters C4 and C8 expose some similarities and overlay each other on both Piper and factor scores plots (Figs. 6 and 10). Cluster C4 represents a high variety of wells screened in different aquifer systems. These clusters are influenced by both processes of calcite dissolution and silicate weathering (Fig. 11a). Relatively high concentrations of silica are observed in the shallow OSM, which may be related to the occurrence of feldspar in the OSM sediments (Wagner et al. 2003).

As apparent on the dendrogram (Fig. 7), clusters C6 and C8 are closely related, both being assigned to the OSM-shallow; however, C8 represents slightly greater depths. C6 has more Ca, but less Mg than C8 and shows higher mineralisation. The cluster interpretation enhanced by geological 3D modelling is more complex for the clusters C2, C3, C4, C6, C7 and C7, as the differences between the hydrogeochemical patterns are less distinct than for the clusters C1 and C5. On the one hand, in the model of Munich, shallow wells 5 and 6, as well as those lying a bit deeper (7, 8 and 9), belong to cluster C3. This can also be observed in the Munich North model at wells 22 and 23. On the other hand, the four wells in Munich represented by well number 10 are marked with cluster C4 and are located very close to the uppermost tier of the shallow aquifer. Similarly, one can also observe in the Munich North model (Fig. 13) that clusters C3 and C4, which are typical for shallow OSM groundwater, are found interspersed in well 24. However, an exception for cluster C4 is shown in well 11, also located in Munich, reaching great depths from 460 to 340 masl. The latter scenario, as well as that for well 26 (Fig. 13), clearly underlines the difficulty of interpreting the results in the case of long well screens that mix groundwater from various aquifers at different depths.

This can be also seen, for instance, in wells 2, 12 and 13 (Fig. 13), which are partially screened very close together at similar depths and show three different clusters (C1, C4 and C5).

Among the trace elements, the following are present: Sr with a median above 100  $\mu\text{g}/\text{L}$ ; B, Ba – above 10  $\mu\text{g}/\text{L}$ ; Cu, Cr, Li, U, Zn – above 1  $\mu\text{g}/\text{L}$ . The presence of uranium above 1  $\mu\text{g}/\text{L}$  in more than 50% of samples is noteworthy, not only in the Quaternary deposits, but also in the moraines and shallow OSM. As concluded by Banning et al. (2013), uranium primarily originates from lignitic inclusions in the OSM sediment; uranium was transported to and accumulated in lowland moor peats, and nowadays get mobilised, for example, by application of agricultural fertilizer containing nitrate.

## Discussion

The study and utilisation of geopotentials, in particular groundwater abstraction, and the use of shallow (as defined by Halilovic et al. 2022) and deep geothermal energy generates a great amount of data over time and space. Thus, the quality of the data and their spatial distribution may be extremely heterogeneous, so a large amount of data is not automatically a guarantee for establishing a good understanding of the local conditions. Quality assurance processes, as presented in this study, must be developed, implemented, and accordingly documented prior to analysing the data.

Further, understanding the local groundwater chemistry means developing a conceptual model and constitutes a complex problem, which may be resolved by using different lines of evidence. Privett (2019) concluded that using appropriate lines of evidence allows one to increase the understanding and solve challenging geological problems. Also, Freedman et al. (2019) stated that evolving a conceptual model should be associated with applying adequate lines of evidence and a suitable level of complexity to represent hydrogeological systems. Two main aspects resulting from the study are discussed in the following sections.

### Coupling the 3D geological models with hydrogeochemical data

In this study, the three main lines of evidence in characterising hydrogeochemistry were: (1) multivariate statistics, complemented by (2) descriptive statistics, and finally coupled with (3) geological 3D modelling. These lines of evidence did not follow one another as listed here, but rather followed each other in a few iterative steps. Coupling the 3D geological models with hydrogeochemical data can be seen from different perspectives: first of all, the model serves as

a starting point for grouping hydrochemical data, but at the same time hydrogeochemistry poses an additional validation of the model. Moreover, geological 3D modelling provides a better understanding of the geometry of the aquifers. In older studies, the main criterion for the grouping of wells was their depth. Thanks to the 3D model, other criteria can be examined such as the affiliation of wells to certain aquifers, interconnections between groundwater bodies, etc. On the other hand, the presented utilisation of the 3D geological modelling outcomes reveals a possible qualitative verification of the model, meaning that greater confidence in the model with subdivided horizons was established (after the definition of verification by Diaz-Maurin and Saaltink 2021).

A first understanding of the hydrogeochemical characteristics of the study area was obtained by defining four main hydrogeological units and assigning objects to these units (often taking advantage of the outcomes of the geological 3D modelling), followed by calculating descriptive statistics for 46 parameters (Zosseder et al. 2022; results described briefly in Section S2 of the ESM). The results of this approach were satisfying but did not allow for a deepened assessment of hydrogeochemical patterns; therefore, more sophisticated multivariate statistical methods were implemented.

Multivariate statistics have successfully been applied in hydrogeochemistry studies, especially when dealing with large-scale data (e.g. Heine et al. 2021), when presenting information on human impacts on hydrogeochemistry (Menció and Mas-Pla 2008) and distinguishing between distinct hydrogeochemical zones, as well as increasing the general understanding of hydrogeological processes (Cloutier et al. 2008). All these aspects were addressed in the presented study.

The hydrogeochemical groups (clusters) obtained by multivariate statistical methods and their dependencies could be explained using plots and diagrams, and, finally, being presented in 3D view has thus enhanced their understanding. However, the clusters were based on selected parameters. After creating the clusters by means of multiple statistical methods, it was necessary to fall back on descriptive statistics in order to characterise the hydrogeochemistry more precisely. In this way, descriptive statistics, which constitute a more traditional approach, played an important role in this study, as it enabled further insight into the hydrogeochemistry—for instance, the patterns of trace element occurrence could not be described otherwise.

The clusters obtained in this study, based on the hydrogeochemistry, were not as unambiguous as expected, because the local geology offers multiple pathways for water exchange and mixing at almost all levels and over the whole area, as has already been hinted at in the 3D model (Albarrán-Ordás and Zosseder 2022). Therefore, the authors support the idea that the clusters actually not only present an

effective way to highlight distinct water groups where possible, but they also effectively highlight relationships between mixed waters. However, some wells located close to each other and screened at almost the same depths represented different clusters, probably due to some local conditions not captured in the 3D model, or due to the long well screen lengths reaching across different groundwater systems. The latter circumstance is economically explicable, but complicates capturing the hydrogeochemistry characteristics of distinct levels—for example, Rauert et al. (1993) suspected that samples analysed in their study represented mixed waters originating from great depth intervals and, therefore, of different ages.

The two main studies highlighted so far that combined hydrogeochemistry with 3D geological modelling were reported by Martinez et al. (2017) and Raiber et al. (2012). Use of 3D geological models in these studies, similar to the one reported here, has enhanced understanding of the aquifer systems, and their spatial relationships and interconnections. Taking into account the current development of 3D modelling and its numerous applications (Albarrán-Ordás and Zosseder 2022), and a successful implementation as presented in this study, new applications in this thematic area can be expected to follow.

## Remarks on the conceptual regional hydrogeochemical model

A conceptual regional hydrogeochemical model was developed based on the results of multivariate statistics supported by geological 3D modelling. Based on that, the following can be stated:

- The deeper aquifers, as seen in the geological 3D models, represent the most distinct hydrogeochemical signature of the Na–HCO<sub>3</sub> water type. In the remaining clusters, a transition from deeper (alkaline) to shallow (alkaline-earth) groundwater can be observed.
- The dissolution of carbonate gravels present in the Quaternary layers, by recharge waters, contributes Ca<sup>2+</sup> and HCO<sub>3</sub><sup>-</sup> ions. Whereas the Mg<sup>2+</sup> and Ca<sup>2+</sup> concentrations are governed predominantly by calcite weathering, in some of the other zones, silicate weathering is the governing process (Fig. 11a).
- Due to cation exchange processes, Ca<sup>2+</sup> is exchanged by Na<sup>+</sup> in the deeper levels (Borzi et al. 2021).
- As the occurrence of nitrate is commonly related to agriculture and livestock, its presence in OSM groundwater (cluster 8) points towards the vulnerability of shallow OSM, particularly where these layers are not covered by impermeable clays or silts.

- According to Bertleff et al. 1993 and Heine et al. 2021, ion exchange waters of the Na–HCO<sub>3</sub> type were recharged during colder-than-present climatic conditions. This is confirmed by the stable water isotopes results, which present much lower values than modern waters.
- In previous studies of the OSM (Andres and Egger 1985; Wagner et al. 2003) a distinction between deeper and shallow groundwater was made based on the tritium content; waters occurring deeper than 30 m were usually free of tritium. Tritium content data from the data bank of LfU and previous studies were used as a supporting argument in classifying wells. An interface of tritium-free old water and young tritium-containing water was estimated at a depth of ~30 m (Andres and Egger 1985). This estimation of depth is still in use by local authorities for distinguishing between the shallow and deep OSM (e.g. Wagner et al. 2003; Chavez-Kus et al. 2016). However, ~30 years ago, Rauert et al. (1993) observed that, in the Munich area, under hydraulic stress (by intensive exploitation of deeper aquifers), young tritium-containing waters have reached the depths of 60 m. Tritium-free water was observed in deeper aquifers as well as in the well-isolated shallow OSM aquifer, mostly in the southern part of the project area. The former establishment of the tritium-interface depth is therefore unjustified.
- The data set was inhomogeneous in many ways, e.g. it was collected over a few decades and, therefore, characterized by multiple limits of detection for the same parameters. It was challenging to work with censored data ('censored' meaning data that have been reported above a predetermined value). Simple methods, such as replacing censored data with the detection limit or half of the detection limit, cause loss of valuable data. For that reason, descriptive statistics can be calculated using the NADA package in the statistic programme R (R Core Team 2021) with "regression on order statistics" (ROS; Helsel 2012). ROS is a semi-parametric method used to estimate summary statistics and plot model distributions with censored data; it assumes that data can be fit to a known distribution by a least-squares regression on a probability plot (ITRC 2013).

## Conclusions

The presented investigation of the hydrogeochemical characteristics of groundwater in the shallow aquifers of the greater area of Munich was a continuation of extensive research on the shallow deposits of the Quaternary and Neogene subsurface in the wider area of Munich, conducted at the Chair of Hydrogeology, TUM, Germany (Kerl et al. 2012; Böttcher et al. 2019; Albarrán-Ordás and Zosseder 2020,

2022; Theel et al. 2020; Zosseder et al. 2022; Böttcher and Zosseder 2022).

In this study, different methods were applied in order to characterise the groundwater chemistry in the most suitable way. It was found that the integration of descriptive and multivariate statistical techniques with 3D geological modelling gave the most satisfying outcome. The methods used posed different lines of evidence and were complementary to one another. The presented workflow of dealing with inhomogeneous data sets and the methodology of quality assurance demonstrate the importance of adequate data management. Presented procedures can be (after adaptation) implemented elsewhere.

Three water types were classified according to their hydrogeochemical composition and location in the 3D geological system. It was recognized that deeper aquifer systems are characterised by the most separate hydrogeochemical signature of the Na–HCO<sub>3</sub> water type. Also, a transition from deeper (alkaline) to shallow (alkaline-earth) groundwater was observed, as well as an anthropogenic impact on water quality in the shallow aquifers.

The outcomes of this study improve understanding of the regional hydrogeology. The results can be utilised as a stimulus to other studies (in particular the geological 3D model and the derived architectural model) for improved groundwater management and sustainable usage of geopotentials.

This study has, as usual, some limitations resulting, for instance, from selected methodologies. First of all, the goal of this work was to capture the spatial variability of hydrogeochemical facies. Therefore, from the available analyses, only one fulfilling the defined criteria (actuality and number of analysed parameters) was selected for each well. Investigation of the temporal variability of physiochemical parameters can be an interesting possibility for future research.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10040-023-02761-z>.

**Acknowledgements** Grateful acknowledgment is given to the Bavarian Environment Agency (LfU) for their support and supervision, as well as for access to the borehole database of the Soil Information System of Bavaria. The authors thank Manuel Gossler (TUM) and Florian Heine (TUM) for their help in R programming, Marco Kerl (TUM) for productive cooperation within the project, Susanne Thiemann (TUM) for analysing water stable isotopes, Günter Kus (LfU) for instructions on the standardized sampling methods of the Bavarian Environment Agency and for information to hydrogeochemical aspects of the study area, Thomas Schaller (TUM) for help in the sampling campaigns, and finally all the institutions that provided us with data. Lastly, Emerson E&P is kindly acknowledged for providing licenses for the Software product SKUA™ Engineering Modelling in the scope of the Emerson Academic Program, which supported the results of this publication.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This study was funded by the Bavarian State Ministry of the Environment and Consumer Protection (StMUV) and was part of the GeoPot Project.

## Declarations

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Admiraal H, Cornaro A (2016) Why underground space should be included in urban planning policy: and how this will enhance an urban underground future. *Tunn Undergr Sp Technol* 55:214–220. <https://doi.org/10.1016/j.tust.2015.11.013>
- Albarrán-Ordás A, Zosseder K (2019) 3D facies modelling based on particle size distributions of borehole data: a case study of geore-source management in the city of Munich. In: 5th European Meeting on 3D Geological Modelling, Bern, Switzerland, May 2019
- Albarrán-Ordás A, Zosseder K (2020) Geostatistical relief modelling of the Quaternary aquifer basis in the Munich Gravel Plain with the use of big datasets. *J Appl Region Geol* 171:1–19. <https://doi.org/10.1127/zdgg/2020/0206>
- Albarrán-Ordás A, Zosseder K (2022) The Di models method: geological 3-D modeling of detrital systems consisting of varying grain fractions to predict the relative lithological variability for a multipurpose usability. *Bull Eng Geol Environ* 81:34. <https://doi.org/10.1007/s10064-021-02538-2>
- Andres G, Egger R (1985) A new tritium interface method for determining the recharge rate of deep groundwater in the Bavarian Molasse basin. *J Hydrol* 82:27–38. [https://doi.org/10.1016/0022-1694\(85\)90044-7](https://doi.org/10.1016/0022-1694(85)90044-7)
- Appelo CAJ, Postma D (2004) *Geochemistry, groundwater and pollution*. CRC, London
- Backhaus K, Erichson B, Plinke W, Weiber R (2016) *Multivariate Analysemethoden [Multivariate analysis]*. Springer, Heidelberg, Germany
- Banning A, Demmel T, Rude TR, Wrobel M (2013) Groundwater uranium origin and fate control in a river valley aquifer. *Environ Sci Technol* 47:13941–13948. <https://doi.org/10.1021/es304609e>
- Bauer M, Neumann P, Scholz M, Thuro K (2006) The geology of the city of Munich and its significance for ground modelling in urban areas. In: Proc. IAEG Int. Congr., Geological Society London, Nottingham, England, pp 83–92
- Bayer P, Attard G, Blum P, Menberg K (2019) The geothermal potential of cities. *Renew Sustain Energy Rev* 106:17–30. <https://doi.org/10.1016/j.rser.2019.02.019>
- Bertleff B, Ellwanger D, Szenkler C, Eichinger L, Trimborn P, Wolfendale N (1993) Interpretation of hydrochemical and hydroisotopic measurements on paleogroundwaters in Oberschwaben, South German Alpine Foreland, with focus on quaternary geology. In: *Isotope techniques in the study of past and current environmental changes in the hydrosphere and the atmosphere*. Proceedings of a symposium, Vienna, 19–23 April 1993, International Atomic Energy Agency, Vienna, pp 337–357
- Borzi G, Idaszkin YL, Tanjal C, Galliari J, Carol E (2021) Assessment of anthropogenic pollution using multiple hydrogeochemical tools and statistical analysis in rural plain basins of the Argentinian Pampean Plain. *River Res Appl* 37:826–842. <https://doi.org/10.1002/rra.3800>
- Böttcher F, Zosseder K (2022) Thermal influences on groundwater in urban environments: a multivariate statistical analysis of the subsurface heat island effect in Munich. *Sci Total Environ* 810:152193. <https://doi.org/10.1016/j.scitotenv.2021.152193>
- Böttcher F, Casasso A, Götzl G, Zosseder K (2019) TAP - Thermal aquifer Potential: a quantitative method to assess the spatial potential for the thermal use of groundwater. *Renew Energy* 142:85–95. <https://doi.org/10.1016/j.renene.2019.04.086>
- Chavez-Kus L, Kainzmeier B, Muhr C, Paul R, Riße I, Scholz M, Wilferth T, Pokowitz C, Blomenhofer A (2016) *Geowissenschaftliche Landesaufnahme in der Planungsregion 18 Südoberbayern*. Bayerisches Landesamt für Umwelt, Augsburg, Germany
- Cloutier V, Lefebvre R, Therrien R, Savard MM (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J Hydrol* 353:294–313. <https://doi.org/10.1016/j.jhydrol.2008.02.015>
- Conway JM, Huffcutt AI (2003) A review and evaluation of exploratory factor analysis practices in organizational research. *Organ Res Methods* 6:147–168. <https://doi.org/10.1177/1094428103251541>
- Curtis ZK, Li SG, Liao HS, Lusch D (2018) Data-driven approach for analyzing hydrogeology and groundwater quality across multiple scales. *Groundwater* 56:377–398. <https://doi.org/10.1111/gwat.12584>
- Diaz-Maurin F, Saaltink MW (eds) (2021) *Model validation for subsurface dynamics*. GEOTHERMICA, Reykjavík, Iceland
- Doppler G, Kroemer E, Rögner K, Wallner J, Jerz H, Grotenthaler W (2011) Quaternary stratigraphy of southern Bavaria. *E&G Quat Sci J* 60:329–365. <https://doi.org/10.3285/eg.60.2-3.08>
- English M (2017) Package ‘hydrogeo’: groundwater data presentation and interpretation (Version 0.6-1). <https://cran.r-project.org/web/packages/hydrogeo/hydrogeo.pdf>. Accessed Dec 2023
- Epting J, Böttcher F, Mueller MH, García-Gil A, Zosseder K, Huggenberger P (2020) City-scale solutions for the energy use of shallow urban subsurface resources: bridging the gap between theoretical and technical potentials. *Renew Energy* 147:751–763. <https://doi.org/10.1016/j.renene.2019.09.021>
- Frape SK, Blyth A, Blomqvist R, McNutt RH, Gascoyne M (2003) Deep fluids in the continents: II. crystalline rocks. In: *Treatise on geochemistry*. Elsevier, Amsterdam, pp 541–580
- Freedman V, Connelly M, Rockhold M, Hasan N, Mehta S, McMahon WJ, Kozak M, Hou ZJ, Bergeron M (2019) A multiple lines of evidence approach for identifying geologic heterogeneities in conceptual site models for performance assessments. *Sci Total Environ* 692:450–464. <https://doi.org/10.1016/j.scitotenv.2019.07.213>
- Ghesquière O, Walter J, Chesnaux R, Rouleau A (2015) Scenarios of groundwater chemical evolution in a region of the Canadian Shield based on multivariate statistical analysis. *J Hydrol Reg Stud* 4:246–266. <https://doi.org/10.1016/j.ejrh.2015.06.004>
- Gilbert-Alarcón C, Daesslé LW, Salgado-Méndez SO, Pérez-Flores MA, Knöller K, Kretschmar TG, Stump C (2018) Effects of reclaimed water discharge in the Maneadero coastal aquifer, Baja California, Mexico. *Appl Geochem* 92:121–139. <https://doi.org/10.1016/j.apgeochem.2018.03.006>

- Goretzko D, Pham TTH, Bühner M (2019) Exploratory factor analysis: current use, methodological developments and recommendations for good practice. *Curr Psychol* 2. <https://doi.org/10.1007/s12144-019-00300-2>
- Griffioen J, van Wensem J, Oomes JLM, Barends F, Breunese J, Bruining H, Olsthoorn T, Stams AJM, van der Stoep AEC (2014) A technical investigation on tools and concepts for sustainable management of the subsurface in The Netherlands. *Sci Total Environ* 485–486:810–819. <https://doi.org/10.1016/j.scitotenv.2014.02.114>
- Güler C, Thyne GD, McCray JE, Turner AK (2002) Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol J* 10:455–474. <https://doi.org/10.1007/s10040-002-0196-6>
- Hajji S, Karoui S, Nasri G, Allouche N, Bouri S (2021) EFA-CFA integrated approach for groundwater resources sustainability in agricultural areas under data scarcity challenge: case study of the Souassi aquifer, central-eastern Tunisia. *Environ Dev Sustain*. <https://doi.org/10.1007/s10668-020-01155-5>
- Halilovic S, Böttcher F, Kramer SC, Piggott MD, Zosseder K, Hamacher T (2022) Well layout optimization for groundwater heat pump systems using the adjoint approach. *Energy Convers Manag* 268:116033. <https://doi.org/10.1016/j.enconman.2022.116033>
- Heine F, Zosseder K, Einsiedl F (2021) Hydrochemical zoning and chemical evolution of the deep upper Jurassic thermal groundwater reservoir using water chemical and environmental isotope data. *Water* 13:1162. <https://doi.org/10.3390/w13091162>
- Helsel DR (2012) Statistics for censored environmental data using Minitab and R, 2nd edn. Wiley, Hoboken, NJ
- Hörling B, Coldewey WG (2013) Hydrogeologie [Hydrogeology]. Spektrum, Heidelberg, Germany
- IAEA/WMO (2020) Global Network of Isotopes in Precipitation. The GNIP Database. <https://nucleus.iaea.org/wiser>. Accessed Dec 2023
- Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, New York
- ITRC (2013) Groundwater statistics and monitoring compliance: statistical tools for the Project Life Cycle. The Interstate Technology & Regulatory Council, Washington, DC
- Jerz H (1993) Geologie von Bayern II - Das Eiszeitalter in Bayern: Erdgeschichte, Gesteine, Wasser, Boden [Geology of Bavaria part 2: stratigraphy, rocks, water, soil]. E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, Germany
- Kainzmeier B, Thom P, Wrobel M, Pukowitz C, Lischeid G, Pamer R (2007) Geowissenschaftliche Landesaufnahme in der Planungsregion 13 Landshut [Geological study of planning region 13 Landshut]. Bayerisches Landesamt für Umwelt, Augsburg, Germany
- Kassambara A (2015) Multivariate analysis 1: practical guide to cluster analysis in R. STHDA. <http://www.sthda.com/english/>. Accessed Dec 2023
- Kerl M, Runge N, Tauchmann H, Goldscheider N (2012) Hydrogeologisches Konzeptmodell von München: Grundlage für die thermische Grundwassernutzung [Conceptual hydrogeological model of the City of Munich, Germany, as a basis for geothermal groundwater utilisation]. *Grundwasser* 17:127–135. <https://doi.org/10.1007/s00767-012-0199-8>
- Kirabo Kacyira A (2012) Addressing the sustainable urbanization challenge. *UN Chron* 49:58–60. <https://doi.org/10.18356/f813137d-en>
- Kuhlemann J, Kempf O (2002) Post-Eocene evolution of the North Alpine Foreland Basin and its response to Alpine tectonics. *Sediment Geol* 152:45–78. [https://doi.org/10.1016/S0037-0738\(01\)00285-8](https://doi.org/10.1016/S0037-0738(01)00285-8)
- Ma X (2018) Data science for geoscience: leveraging mathematical geosciences with semantics and open data. In: Handbook of mathematical geosciences. Springer, Cham, Switzerland, pp 687–702
- Małoszewski P, Rauert W, Stichler W, Herrmann A (1983) Application of flow models in an alpine catchment area using tritium and deuterium data. *J Hydrol* 66:319–330. [https://doi.org/10.1016/0022-1694\(83\)90193-2](https://doi.org/10.1016/0022-1694(83)90193-2)
- Martinez JL, Raiber M, Cendón DI (2017) Using 3D geological modelling and geochemical mixing models to characterise alluvial aquifer recharge sources in the upper Condamine River catchment, Queensland, Australia. *Sci Total Environ* 574:1–18. <https://doi.org/10.1016/j.scitotenv.2016.09.029>
- Maurer H, Buchner E (2007) Fluvial systems of the Upper Freshwater Molasse (North Alpine Foreland Basin, SW-Germany). *Zeitschr Dtsch Gesellsch Geowissensch* 158:249–270. <https://doi.org/10.1127/1860-1804/2007/0158-0249>
- Menció A, Mas-Pla J (2008) Assessment by multivariate analysis of groundwater-surface water interactions in urbanized Mediterranean streams. *J Hydrol* 352:355–366. <https://doi.org/10.1016/j.jhydrol.2008.01.014>
- Montcoudiol N, Molson J, Lemieux JM (2015) Groundwater geochemistry of the Outaouais Region (Québec, Canada): a regional-scale study. *Hydrogeol J* 23:377–396. <https://doi.org/10.1007/s10040-014-1190-5>
- Newman BD, Havenor KC, Longmire P (2016) Identification of hydrochemical facies in the Roswell Artesian Basin, New Mexico (USA), using graphical and statistical methods. *Hydrogeol J* 24:819–839. <https://doi.org/10.1007/s10040-016-1401-3>
- Odziomek K, Rybinska A, Puzyn T (2017) Unsupervised learning methods and similarity analysis in chemoinformatics. In: Handbook of computational chemistry. Springer, Cham, Switzerland, pp 2095–2132
- Privett KD (2019) The lines of evidence approach to challenges faced in engineering geological practice. *Q J Eng Geol Hydrogeol* 52:141–172. <https://doi.org/10.1144/qjgeh2018-131>
- R Core Team (2021) R: a language and environment for statistical computing. <https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing>. Accessed Dec 2023
- Raiber M, White PA, Daughney CJ, Tschirner C, Davidson P, Bainbridge SE (2012) Three-dimensional geological modelling and multivariate statistical analysis of water chemistry data to analyse and visualise aquifer structure and groundwater composition in the Wairau Plain, Marlborough District, New Zealand. *J Hydrol* 436–437:13–34. <https://doi.org/10.1016/j.jhydrol.2012.01.045>
- Rauert W, Wolf M, Weise SM, Andres G, Egger R (1993) Isotope-hydrogeological case study on the penetration of pollution into the deep tertiary aquifer in the area of Munich, Germany. *J Contam Hydrol* 14:15–38. [https://doi.org/10.1016/0169-7722\(93\)90039-U](https://doi.org/10.1016/0169-7722(93)90039-U)
- Rebala G, Ravi A, Churiwala S (2019) Handbook of computational chemistry. Springer, Cham, Switzerland
- Revelle W (2020) psych: procedures for psychological, psychometric, and personality research. R package version 2.0.12.. <http://personality-project.org/r/psych-manual.pdf>. Accessed Dec 2023
- Russo L, Taddia G (2009) Groundwater in the urban environment: management needs and planning strategies. *Am J Environ Sci* 5:493–499. <https://doi.org/10.3844/ajessp.2009.493.499>
- Schirmer M, Leschik S, Musolff A (2013) Current research in urban hydrogeology: a review. *Adv Water Resour* 51:280–291. <https://doi.org/10.1016/j.advwatres.2012.06.015>
- Schumacker RE (2016) Using R with multivariate statistics. Sage, Thousand Oaks, CA
- Senbayram M, Gransee A, Wahle V, Thiel H (2015) Role of magnesium fertilisers in agriculture: plant-soil continuum. *Crop Pasture Sci* 66:1219–1229. <https://doi.org/10.1071/CP15104>
- Stumpp C, Klaus J, Stichler W (2014) Analysis of long-term stable isotopic composition in German precipitation. *J Hydrol* 517:351–361. <https://doi.org/10.1016/j.jhydrol.2014.05.034>
- Team GM (2003) GeoMol - assessing subsurface potentials of the Alpine Foreland Basins for sustainable planning and use of natural

- resources. Project Report. Bayerisches Landesamt für Umwelt, Augsburg, Germany
- Theel M, Huggenberger P, Zosseder K (2020) Assessment of the heterogeneity of hydraulic properties in gravelly outwash plains: a regionally scaled sedimentological analysis in the Munich gravel plain, Germany. *Hydrogeol J* 28:2657–2674. <https://doi.org/10.1007/s10040-020-02205-y>
- Torres-Martínez JA, Mora A, Knappett PSK, Ornelas-Soto N, Mahl-knecht J (2020) Tracking nitrate and sulfate sources in groundwater of an urbanized valley using a multi-tracer approach combined with a Bayesian isotope mixing model. *Water Res* 182:115962. <https://doi.org/10.1016/j.watres.2020.115962>
- Tweed S, Leblanc M, Cartwright I, Bass A, Travi Y, Marc V, Bach TN, Duc ND, Massuel S, Kumar US (2019) Stable isotopes of water in hydrogeology. In: *Encyclopedia of water*. Wiley, Hoboken, NJ, pp 1–10
- van Geldern R, Baier A, Subert HL, Kowol S, Balk L, Barth JAC (2014) Pleistocene paleo-groundwater as a pristine fresh water resource in southern Germany: evidence from stable and radiogenic isotopes. *Sci Total Environ* 496:107–115. <https://doi.org/10.1016/j.scitotenv.2014.07.011>
- Vázquez-Suñé E, Carrera J, Tubau I, Sánchez-Vila X, Soler A (2010) An approach to identify urban groundwater recharge. *Hydrol Earth Syst Sci* 14:2085–2097. <https://doi.org/10.5194/hess-14-2085-2010>
- Venables W, Ripley B (2002) *Modern applied statistics with S*, 4th edn. Springer, New York
- Volchko Y, Norrman J, Ericsson LO, Nilsson KL, Markstedt A, Öberg M, Mossmark F, Bobylev N, Tengborg P (2020) Subsurface planning: towards a common understanding of the subsurface as a multifunctional resource. *Land Use Policy* 90:104316. <https://doi.org/10.1016/j.landusepol.2019.104316>
- Wagner B, Töpfner C, Lischeid G, Scholz M, Klinger R, Klaas P (2003) Hydrochemische Hintergrundwerte der Grundwässer Bayerns [Hydrogeochemical background values of groundwater in Bavaria]. Bayerisches Geologisches Landesamt, Munich, Germany
- Wagner B, Walter T, Himmelsbach T, Clos P, Beer A, Budziak D, Dreher T, Fritsche HG, Hübschmann M, Marcziniek S, Peters A, Poeser H, Schuster H, Steinel A, Wagner F, Wirsing G (2011) A web map service for background groundwater chemistry in Germany [Hydrogeochemische Hintergrundwerte der Grundwässer Deutschlands als Web Map Service]. *Grundwasser* 16:155–162. <https://doi.org/10.1007/s00767-011-0161-1>
- Wakode HB, Baier K, Jha R, Azzam R (2018) Impact of urbanization on groundwater recharge and urban water balance for the city of Hyderabad, India. *Int Soil Water Conserv Res* 6:51–62. <https://doi.org/10.1016/j.iswcr.2017.10.003>
- Wei T, Simko T (2021) R package “corrplot”. Visualization of a correlation matrix. (Version 0.92), <https://github.com/taiyun/corrplot>. Accessed Dec 2023
- Wickham H (2009) *ggplot2. Elegant graphics for data analysis*. Springer, New York
- Wild LM, Mayer B, Einsiedl F (2018) Decadal delays in groundwater recovery from nitrate contamination caused by low O<sub>2</sub> reduction rates. *Water Resour Res* 54(9996–10):012. <https://doi.org/10.1029/2018WR023396>
- Wild LM, Rein A, Einsiedl F (2020) Monte Carlo simulations as a decision support to interpret δ<sup>15</sup>N values of nitrate in groundwater. *Groundwater* 58:571–582. <https://doi.org/10.1111/gwat.12936>
- Yong AG, Pearce S (2013) A beginner’s guide to factor analysis: focusing on exploratory factor analysis. *Tutor Quant Methods Psychol* 9:79–94. <https://doi.org/10.20982/tqmp.09.2.p079>
- Zosseder K, Chavez-Kus L, Kerl M, Albarrán-Ordás A, Gossler MA, Theel M, Kiecak A, Huch J, Schaller T (2019) Abschlussbericht zum Forschungsvorhaben “Parameterbestimmung für die Abschätzung der geologischen Nutzungspotenziale in der Planungsregion 14 und im tertiären Untergrund des Großraums München” [Final report on the research project “Determination of parameters for the assessment of the geological utilization potential in planning region 14 and in the tertiary subsoil of the greater Munich area”]. Technical University of Munich, Chair of Hydrogeology, Munich. <https://www.lfu.bayern.de/geologie/geopot/index.htm>. Accessed Dec 2023
- Zosseder K, Kerl M, Albarrán-Ordás A, Gossler MA, Kiecak A, Chavez-Kus L (2022) Die hydraulischen Grundwasserverhältnisse des quartären und des oberflächennahen tertiären Grundwasserleiters im Großraum München [Hydraulic groundwater conditions of the quaternary and tertiary aquifer in the wider area of Munich]. *Geologica Bavarica* 122. Bayerisches Landesamt für Umwelt, Augsburg, Germany
- Zuber A, Michalczyk Z, Maloszewski P (2001) Great tritium ages explain the occurrence of good-quality groundwater in a phreatic aquifer of an urban area, Lublin, Poland. *Hydrogeol J* 9:451–460. <https://doi.org/10.1007/s100400100149>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.