Probabilistic assessment of spatial heterogeneity of natural background concentrations in large-scale groundwater bodies through Functional Geostatistics

L. Guadagnini[a,*], A. Menafoglio[b], X. Sanchez-Vila[a], A. Guadagnini[c]

[a]Department of Civil and Environmental Engineering, Universitat Politècnica de Catalunya, Jordi Girona 1-3, 08034 Barcelona, Spain

[b]Politecnico di Milano, MOX, Department of Mathematics, Piazza L. Da Vinci 32, 20133 Milano, Italy

[c]Politecnico di Milano, Dipartimento di Ingegneria Civile e Ambientale, Piazza L. Da Vinci 32, 20133 Milano, Italy

*Corresponding author: laura.guadagnini@polimi.it, laura.guadagnini@upc.edu

Abstract

We propose and exemplify a framework to assess Natural Background Levels (NBLs) of target chemical species in large-scale groundwater bodies based on the context of Object Oriented Spatial Statistics. The approach enables one to fully exploit the richness of the information content embedded in the probability density function (PDF) of the variables of interest, as estimated from historical records of chemical observations. As such, the population of the entire distribution functions of NBL concentrations monitored across a network of monitoring boreholes across a given aquifer is considered as the object of the spatial analysis. Our approach starkly differs from previous studies which are mainly focused on the estimation of NBLs on the basis of the median or selected quantiles of chemical concentrations, thus resulting in information loss and limitations related to the need to invoke parametric assumptions to obtain further summary statistics in addition to those considered for the spatial analysis. Our work enables one to (*i*) assess spatial dependencies among observed PDFs of natural background concentrations, (*ii*) provide spatially distributed kriging predictions of NBLs, as well as (*iii*) yield a robust quantification of the ensuing uncertainty and probability of exceeding given threshold concentration values via stochastic simulation. We illustrate the approach by considering the (probabilistic) characterization of spatially variable NBLs of ammonium and arsenic detected at a monitoring network across a large scale confined groundwater body in Northern Italy.

Keywords: Natural background level; groundwater quality; chemical status; Kriging; probability density function; uncertainty quantification.

1. Introduction

Robust characterization of the natural chemical signature of a given groundwater system is a key component of modern environmental analysis. Critical aspects associated with this step include the identification of values of sampled concentrations of target chemicals that could be related to geogenic contributions. In this context, it is recognized that markedly high Natural Background Levels (NBLs) of chemical species/compounds of interest can potentially be linked to petrographical (e.g., Hinsby and Condesso de Melo, 2006) or lithological and sedimentological site-specific characteristics (e.g., Redman et al., 2002; Molinari et al., 2013 and references therein) rather than being attributable to anthropogenic actions. Relating high values of sampled concentrations to an anthropogenic rather than a natural contribution may sometimes yield misleading assessments of environmental risks, improper classification of the chemical status (e.g., in terms of a good status, as defined by the European Water Framework Directive, WFD 2000/60/EC GWDD 2006/118/EC Directive 2014/80/EU) of aquifer bodies, as well as setting remediation goals which can be unattainable and/or unsustainable. In this context, modern regulatory frameworks at the European level highlight the need for an appropriate assessment of baseline concentrations, i.e. those that can be ascribed to geogenic effects and are not caused by anthropogenic activities.

Identification and implementation of a complete (generally multicomponent) geochemical model accounting for the complexity of processes driving flow and transport in porous media in the presence of the various sources of uncertainty associated with the ubiquitously heterogeneous subsurface is not always feasible. A series of investigations are then keyed to the development of procedures leading to embedding information within a management framework upon relying on a limited amount of data. The latter typically comprise monitored temporal series of concentration samples (Edmunds et al., 2003, Wendland et al., 2005, Panno et al., 2006, Walter, 2008, Urresti-Estala et al., 2013, Kim et al., 2015; Liang et al., 2017, 2018, 2019).

As an example, one of the main outcomes of the EU funded project BRIDGE (2007), Background cRiteria for the IDentification of Groundwater thrEsholds, is a guideline that allows assessing the natural status of a groundwater body through a Pre-Selection methodology. The latter is based on the identification of pristine groundwater samples within an available collection of observations. This procedure typically yields the estimate of a unique (or bulk) NBL value, which is then assigned to the groundwater body under investigation. According to this approach, concentration values of a chemical species of interest exceeding such a threshold are then ascribed to anthropogenic activities. A notably weak point of such an approach is that it renders a unique NBL value, disregarding spatial variability, this aspect being critical when considering large scale heterogeneous (in terms of petrographic and hydrogeologic characteristics) aquifers. As a further evolution, some authors suggest that the NBL of a natural groundwater system should be expressed in terms of a range of values (e.g., Reimann and Garrett, 2005; Hinsby et al., 2008; Li et al., 2014) rather than being constrained to a single one.

Studies related to characterizing the spatial variability of NBL concentrations include, e.g., the work of Ducci et al. (2016) and Dalla Libera et al. (2017). While the former relies on indicator kriging to demarcate regions associated with given probability of exceeding a target NBL value, the latter authors propose a zonation approach leading to piece-wise uniform NBL concentration maps. The analysis of Molinari et al. (2019) starts from values of the 90^{th} percentile of concentration samples observed at a set of monitoring boreholes. These are then subject to standard variography upon considering alternative variogram models which are then employed in a multimodel context to provide kriging-based spatial distributions of estimates of NBL concentrations. The resulting kriged values are used jointly with the ensuing estimation variance to evaluate spatial distributions of the probability of exceeding predefined threshold values of NBL concentrations, the latter being assumed to be characterized by a log-normal distribution. We emphasize that all of these works rely on the representation of observed temporal series of natural background concentrations by way of through scalar summaries (e.g., the 90^{th} percentile), which are then projected onto a set of locations

of interest where data are not available. Doing so results in a loss of information and requires resorting to additional hypotheses, such as assuming a log-normal distribution for NBL values which is parametrized according to the results of the kriging analysis (as in, e.g., Molinari et al., 2019). The general concept underlying these studies is also consistent with approaches treating the characterization of spatial heterogeneity of aquifer systems within a probabilistic context (e.g., Winter et al., 2003; Short et al., 2010; Perulero Serrano et al., 2014; Bianchi Janetti et al., 2019 and references therein).

Our study rests on the concepts underpinning Object Oriented Data Analysis (Marron and Alonso, 2014). Doing so enables us to consider the information content included in the entire distribution function of NBL concentrations monitored at a given observation borehole as the object of the spatial analysis, instead of being limited to selected moments or quantiles. Such a framework renders (*a*) predictions of the complete distribution of NBL concentrations in a non-parametric setting together with the associated uncertainty, and (*b*) joint assessment of all summary quantities of interest of the distribution (including desired quantiles and probability values). Accordingly, the NBL distributions are embedded in a mathematical space whose elements are probability density functions (Egozcue et al., 2006, Van den Boogaart et al., 2014). Our distinctive objective is to leverage on key elements of Object Oriented Spatial Statistics (O2S2, Menafoglio and Secchi, 2017) to (*i*) quantify spatial dependencies among observations, (*ii*) provide spatially distributed kriging predictions, and (*iii*) yield a robust quantification of the uncertainty associated with NBL spatial distributions through stochastic simulation. As detailed in the following, we first illustrate the theoretical framework, and then demonstrate it to characterize spatial variability of NBL distributions of target chemical species by relying on an extensive set of hydrochemical data collected across a large scale confined groundwater body in Northern Italy.

2. Materials and methods

2.1. Study area and data-set

As a test bed to demonstrate the breadth and potential of our approach, we focus on a groundwater body located in the Emilia-Romagna Region (Northern Italy) and demarcated on the basis of both geological/sedimentological information and anthropogenic impact analyses (Regione Emilia-Romagna, 2010). The area is a portion of the Po Basin fill, a syntectonic sedimentary wedge (Ricci Lucchi, 1984) forming the infill of the Pliocene-Pleistocene fore-deep.

Sedimentological and hydrogeological studies are available in the region (Amorosi et al., 1996; Regione Emilia-Romagna-ENI-AGIP, 1998, Regione Emilia-Romagna, 2010), identifying three main hydrogeological complexes: Apennines alluvial fans, Apennine alluvial plain, and alluvial and deltaic Po plain. The complete aquifer system is characterized by a multilayered confined or semiconfined configuration. The thickness of fine deposits increases towards the northern portion of the plain (Regione Emilia-Romagna, 2010; Farina et al., 2014), where conditions of increased confinement are documented.

Additional information regarding the hydrogeological setting of the region are available in Molinari et al. (2012) and Farina et al. (2014). Our study is keyed to one of the large scale groundwater bodies located in the upper confined portion of the aquifer system, within the hydrogeological complex named as Appenine alluvial plain. Figure 1 depicts limits and planar extent of the groundwater body considered, denoted with the identifier 0610 and characterized by an average depth of 75 m, average thickness of 130 m and area of about 2930 km^2.

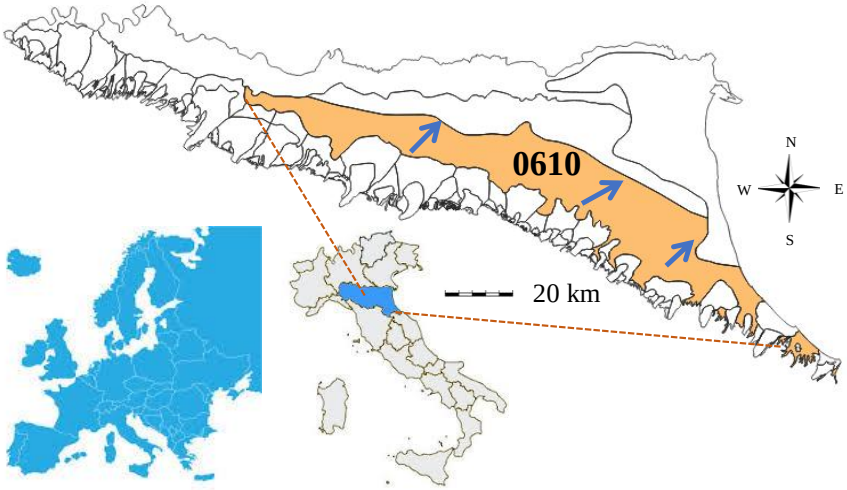

Figure 1. Planar extent of groundwater body 0610 within the Emilia-Romagna Region. Blue arrows correspond to the overall regional-scale groundwater flow direction.

The groundwater body under study is considered to be significantly vulnerable, given its stratigraphic location within the aquifer system and the anthropogenic stresses associated with intensive exploitation for agricultural and civil purposes (Regione Emilia-Romagna, 2010). Being located in the upper confined portion of the complex aquifer system described, its southern limit is in continuity with the recharging areas of alluvial fans. A relevant amount of monitoring boreholes is set within its considerable planar extent, thus yielding a remarkable amount of available chemical data. As evidenced in prior investigations (Molinari et al. 2012, 2019), data about groundwater quality suggest the need to considering regional-scale, spatially heterogeneous distributions of NBL values.

The analyzed data set includes time series of concentrations recorded at several monitoring stations managed by the "Agenzia Regionale per la Prevenzione e l'Ambiente dell'Emilia-Romagna" (ARPAE - Regional Agency for Environmental Protection, Emilia-Romagna). We select monitoring boreholes where 20-year historical records of observations (1987-2008, collected at a six-month interval, albeit not continuously for some wells) are available. We focus on ammonium (NH_4) and arsenic (As), whose documented concentrations locally exceed the limit set by current Italian regulations (also corresponding to the European Drinking Water Standards) set at 0.5 mg/l

and 10 µg/l, respectively, and are seen as critical elements for the achievement of a good chemical status according to Italian Regulation (D. Lgs. 30/09, i.e., Decreto Legislativo n. 30, 16 March 2009) and GWDD 2006/118/EC. A total number of 90 monitoring stations was initially considered by ARPAE to characterize groundwater body 0610. Some of these were associated solely with quantitative measurements of piezometric level. On the basis of a subsequent detailed analysis, monitoring stations that could not be attributed with certainty to the target groundwater body (essentially on the basis of the screen depth) were excluded from the original collection of locations. This has reduced the initial number of monitoring stations attributed to groundwater body 0610 to 62, with a total of 1428 observations. Exclusion of monitoring stations where observations are associated with a temporal window spanning less than 3 years leads to retain 57 monitoring stations with a total of 1354 observations available (among these, ammonium and arsenic have been measured in 1343 and 1193 samples, respectively). Concentrations below detection limit were set equal to half the detection limit. After application of PS, monitoring stations where less than 10 data points are available are further excluded from our analyses. As such, we use 1234 (associated with 44 monitoring stations) for ammonium (see Section 3.1) and 1096 data (related to 43 monitoring stations) for arsenic (see Section 3.2).

2.2 Methodology for Data analysis

2.2.1 NBL estimates

Concentration records are subject to a Pre-Selection (PS) procedure (BRIDGE, 2007) to identify NBL values. This approach enables us to remove samples exceeding certain concentration values, considered indicative of anthropogenic contamination, from the original record of observations. Typically employed criteria for the exclusion of samples deemed as influenced include: (*a*) chloride concentrations > 1000 mg/l, denoting salinity; and (*b*) nitrate (NO_3) concentrations > 10 mg/l, as a signature of anthropogenic influence caused by e.g., fertilizers. Additional criteria (redox conditions, dissolved oxygen, sulfate concentration) can be considered for

sample exclusion (e.g., Hinsby and Condesso de Melo, 2006; Hinsby et al., 2008). For the purpose of our analyses, we follow Molinari et al. (2019) and apply the exclusion criteria listed above.

Data resulting from filtering the raw dataset through PS are considered as observations of naturally occurring NBL concentrations at diverse observation times across the analyzed window. Our analysis rests on monitoring wells which exhibit a time series with more than ten records. We note that the procedure which is then employed for the evaluation of the NBL (e.g., Wendland et al., 2005) relies on (*a*) estimating the median value for the concentrations of the target chemical species identified at each monitoring well via PS, and (*b*) assessing the unique value of NBL associated with the whole water body in terms of a selected percentile (typically the 90^{th}, 95^{th}, or 97.5^{th}).

As illustrated in details in Section 2.2.3, we adopt here a diverse perspective and fully account for the functional nature of the data. The latter are thus analyzed as functional random fields. In this context, the subject of our analysis is the collection of probability functions of NBLs obtained by applying the PS procedure at each monitoring station. By doing so, we go beyond the limitation of relying solely on selected percentiles of such probability functions and take advantage of the complete information content embedded in the entire probability function of NBL reconstructed from the observations at each well.

We structure our study through the following main steps:

1. perform sample selection for historical records at each observation borehole following the adopted exclusion criteria, as indicated in the original BRIDGE (2007) methodology;
2. evaluate the (empirical probability) distribution function of NBLs of (log-transformed) concentrations of the selected chemical species at each observation well;
3. perform spatial prediction and uncertainty quantification of NBL probability density functions (PDFs) at unsampled locations using an object-oriented geostatistical approach (Menafoglio et al., 2014).

We describe the main theoretical elements and the ensuing implementation workflow associated with these steps in the following sections.

2.2.2 Data pre-processing

Data pre-processing aims at extracting an estimate of the NBL PDFs from each temporal series of NBL observations. Each temporal series is considered separately (observations associated with the series being used to build a corresponding histogram) neglecting temporal autocorrelation (additional comments on this choice are given in Section 3). The resulting histogram is then smoothed to yield a continuous estimate of the underlying PDF, as advocated by Machalová et al. (2016) and consistent with the modeling framework employed for the following analysis steps (detailed in Sections 2.2.3-2.2.4). Note that the length of the time-series can have an effect on the accuracy of the PDF estimation, i.e., the longer the time series, the lower the uncertainty in the data-preprocessing. Here, we include all monitoring stations where at least 10 records are available. This is seen as a minimum threshold value to maintain the ability of estimating a density function from the sampled data with a non-parametric approach. In general, the choice of such a threshold should attain a balance between the ability of estimating the PDF with sufficient accuracy, and the need to retain as many measurement sites as possible. This choice is case-specific and depends on the stability of the time-series, the data quality, possible missing data, the density of the measurement locations and their spatial distribution.

2.2.3 Notation and background: geostatistics for PDFs

The smoothed PDF data are considered as the objects of the geostatistical analysis. In the following we denote by $s_1, \dots, s_n$ the n locations in the spatial domain D where the PDFs of NBL are observed, and by $X_{s_1}, \dots, X_{s_n}$ the n smoothed PDFs available at the sampling locations. Here, X_{s_i} denotes the PDF at location s_i, which is a positive function defined on an interval of (log-)concentrations $I = [a, b]$, common to all data. We consider these PDFs as a partial observation of a *functional* random field $\{X_s, s \in D\}$, that is a collection of random functional elements (the PDFs of NBL) indexed by a spatial variable s in D. The goal of the analysis is to provide a kriging prediction of the random field (i.e, the entire PDF, X_{s_0}) at unsampled locations (s_0) in D, based on

the observations available at the monitoring stations. Two key challenges need to be tackled to solve the kriging problem: (*i*) the curse of dimensionality (due to the virtually infinite dimensionality of PDF data, which would need an infinity of point evaluations to be fully characterized), and (*ii*) the data constraints (positivity and unit integral).

To jointly face these challenges, we follow the approach of Menafoglio et al. (2014, 2016a, 2016b), who provide a class of geostatistical methods to analyze datasets of geo-referenced PDFs. These methods are based on the idea of defining an appropriate mathematical space where data are embedded, and use the geometry of the space to perform prediction and stochastic simulation. For instance, if the NBL data were represented through their median (i.e., a scalar summary statistics), the data could be embedded in the space R of real numbers, and analyzed through a typical scalar geostatistics approach. If the NBL data were represented through a set of k summary indices (e.g., mean and standard deviation), a k-dimensional Euclidean space R^k could be used to perform analyses through multivariate geostatistical methods (e.g., Chilès and Delfiner, 1999). Considering functional and constrained data, Menafoglio et al. (2014, 2016a, 2016b) propose to consider a Bayes space (Egozcue et al., 2006; Van den Boogaart et al., 2014), whose elements are PDFs, for embedding and analyzing the data. Bayes spaces provide the generalization to the functional framework of the so-called Aitchison simplex (Aitchison, 1986). In Bayes spaces, appropriate notions of operations between PDFs (e.g., sum $(+)$, or product by a constant $(\cdot)$) as well as of inner product $(\langle\cdot,\cdot\rangle)$ are defined, allowing for the development of a proper theory of kriging and stochastic simulation. For the purpose of this study, we do not present all details of these mathematical constructions and introduce only the key concepts and notation. We refer to Menafoglio et al. (2013, 2014, 2016a, 2016b) for an in-depth introduction to the mathematics underpinning the methods we employ.

2.2.4 Modeling spatial dependence and kriging

As a first step of the geostatistical analysis of the dataset $X_{s_1}, \dots, X_{s_n}$ of PDFs, we model the spatial dependence among data. We assume that (*a*) data are elements of the Bayes space B^2, that is the space of positive functions, whose natural logarithm is square integrable, and (*b*) the field $\{X_s, s \in D\}$ is stationary. This enables us to consider the generalization of the classical variogram to the functional context, which is termed *trace-variogram*. In B^2, the trace-variogram is defined as the function $2\gamma(s_1, s_2)$ that associates with a pair of locations s_1, s_2 (in D) the expected square distance (in B^2) between the NBL PDFs (X_{s_1}, X_{s_2}) at such locations, i.e.,

$$2\gamma(s_1, s_2) = \mathbb{E}\left[d_{B^2}^2\left(X_{s_1}, X_{s_2}\right)\right] = \mathbb{E}\left[\frac{1}{2(b-a)}\int_a^b \int_a^b \ln\left(\frac{X_{s_1}(t)X_{s_2}(s)}{X_{s_1}(s)X_{s_2}(t)}\right) dtds\right]. \quad (1)$$

Interpretation and properties of the trace-variogram for PDF data are very similar to their scalar counterpart. In particular, under stationarity, the trace-variogram depends only on the increment among locations ($s_1 - s_2$), stabilizes at a horizontal asymptote (*sill*), and the distance at which the variogram attains the sill determines the range of association among elements of the field (*range*).

Variogram modeling can be performed in two steps: (*i*) estimating a binned trace-variogram

$$2\gamma(h) = \frac{1}{|N(h)|}\sum_{s_i, s_j \in N(h)} d_{B^2}^2\left(X_{s_i}, X_{s_j}\right), \quad (2)$$

$|N(h)|$ being the number of pairs of sampled sites (approximately) separated by h; and (*ii*) fitting a valid model (e.g., spherical, exponential, matérn) to the empirical estimate (1).

Once the variogram model is estimated, the *functional* kriging prediction for a PDF of NBL at a target location s_0 is based on the best linear unbiased (functional) predictor in the Bayes space B^2. This is defined as the predictor $X_{s_0}^* = \sum_{i=1}^n \lambda_i^* \cdot X_{s_i}$, where symbols denote the linear combination in the Bayes space, and are explicitly written as

$$X_{s_0}^*(t) = \frac{\prod_{i=1}^n X_{s_i}^{\lambda_i^*}(t)}{\int_a^b \prod_{i=1}^n X_{s_i}^{\lambda_i^*}(s)ds}, \quad (3)$$

$\lambda_1^*, \dots, \lambda_n^*$ being scalar weights to be optimized through minimization of the variance of prediction error under unbiasedness. From a practical viewpoint, having estimated the trace-variogram model

2γ, finding the kriging weights reduces to the solution of the very same kriging system associated with scalar geostatistics (see, e.g., Menafoglio and Secchi, 2017, for details).

2.2.5 Stochastic Simulation

Uncertainty quantification for functional kriging can be performed by using conditional stochastic simulation, as originally proposed in Menafoglio et al. (2016b). For this purpose, one necessarily needs to reduce the dimensionality of the data, as it is hardly possible to produce realizations of an infinity of point evaluations of the PDF. Dimensionality reduction can be performed through functional principal component analysis in the Bayes space B^2 (SFPCA, Hron et al., 2016). The SFPCA analysis allows identifying the main directions of variability $(e_1, e_2, ...)$ of the dataset $X_{s_1}, \dots, X_{s_n}$. The elements $e_1, e_2, ...$ are the analogue of the loadings (i.e., the eigenvectors) in multivariate principal component analysis. In particular, loadings $e_1, e_2, ...$ form an orthonormal functional basis of space B^2. Projecting the data along the first K principal components enables one to represent the PDF X_{s_i} through a vector of K coordinates $\boldsymbol{x}_{s_i} = (x_{s_i,1}, \dots, x_{s_i,K})$, thus reducing to K the formerly infinite dimensionality of the PDF. Stochastic simulation of the PDF can be then performed by simulation of the vector of coordinates along the truncated basis $e_1, e_2, \dots, e_K$ at the target location $\boldsymbol{x}_{s_0} = (x_{s_0,1}, \dots, x_{s_0,K})$, based on the coordinate vector available at the sampled sites. Such simulation can be performed through the aid of well-known multivariate methods, such as those based on sequential Gaussian co-simulation (e.g., Chilès and Delfiner, 1999; Kim et al., 2019).

3. Results and discussion

3.1. Ammonium

As stated in Section 2.1, a total of 1234 historical records collected at 44 monitoring stations were available for ammonium concentration after PS, characterized by a number of 12 to 42 observations per monitoring well (with an average of about 28). A preliminary analysis of the data reveals that most locations (41 out of 44) do not display any autocorrelation in the time series of

NBL concentrations (level 1%, as obtained through a Durbin-Watson test on each time series, the p-value of single tests being corrected via Holm's method). Autocorrelation within the temporal series was thus neglected in the data preprocessing. The PDF of NBL log-concentrations (hereafter termed NBL densities or NBL PDFs for ease of illustration) were then estimated at each borehole upon neglecting temporal autocorrelations. The ensuing results are depicted in Figure 2 in terms of smoothed data.

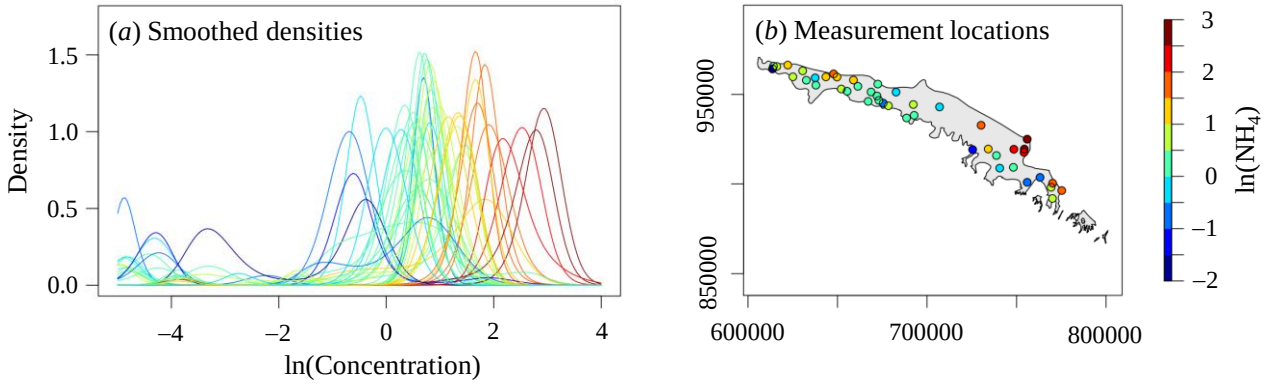

Figure 2. Smoothed data for ammonium log-concentration values and corresponding spatial locations in the investigated aquifer system. Colors are assigned according to the value of the mean related to the corresponding smoothed density. Spatial coordinates are in meters.

Visual inspection of Figure 2 suggests that the highest mean values are associated with the distal portion of the domain, mainly close to the coastal area where water is characterized by high chloride concentrations. A global stationarity assumption of the functional data appears to be supported by the sample trace-semivariogram depicted in Figure 3, which is characterized by a clear asymptote for increasing spatial distances.

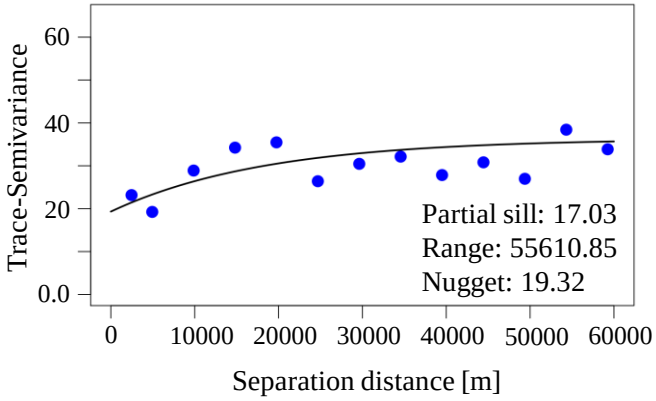

Figure 3. Sample trace-variogram estimated from the smoothed functional data for ammonium log-concentrations and interpreted model with estimated parameters.

An exponential model with nugget was calibrated to the empirical variogram, estimated values of its parameters being included in Figure 3. One may notice the presence of a relevant nugget effect in the structure of spatial dependence, which provides an indication of possible spatial discontinuities in the field of NBL densities. Point Kriging was then performed across a regular grid of 2824 points (of side 983 m and 1048 m along the horizontal and vertical directions, which are taken to correspond to the West-East and South-North directions, respectively). Such a grid encompasses the full aquifer body domain, grid spacing being consistent with the spatial density of the available monitoring network and corresponding to a discretization of the variogram range (Figure 3) with about 50 points. Figure 4a depicts the resulting kriging-based predictions of PDFs of NBL of (log-transformed) ammonium concentrations. We note that, while point Kriging results do not depend on the cell size, the latter can be otherwise influential to the graphical representation associated with the color scale in Figure 4, which can nevertheless capture the overall spatial pattern of the quantities of interest. Cross-validation results (Appendix A) fully support the satisfactory performance of the prediction method.

Figures 4b and 4c illustrate the mean and standard deviation of the predicted NBL densities, respectively. The highest mean values are mostly located in the eastern portion of the domain, close to the coastal groundwater body, with moderate values of standard deviation. These results are consistent with the observation that raw concentration data collected from this area tend to exhibit large NH_4 values that persist over time (see also Figure 2), a finding which is possibly linked to ammonium being more soluble in saline environments as compared to freshwater bodies.

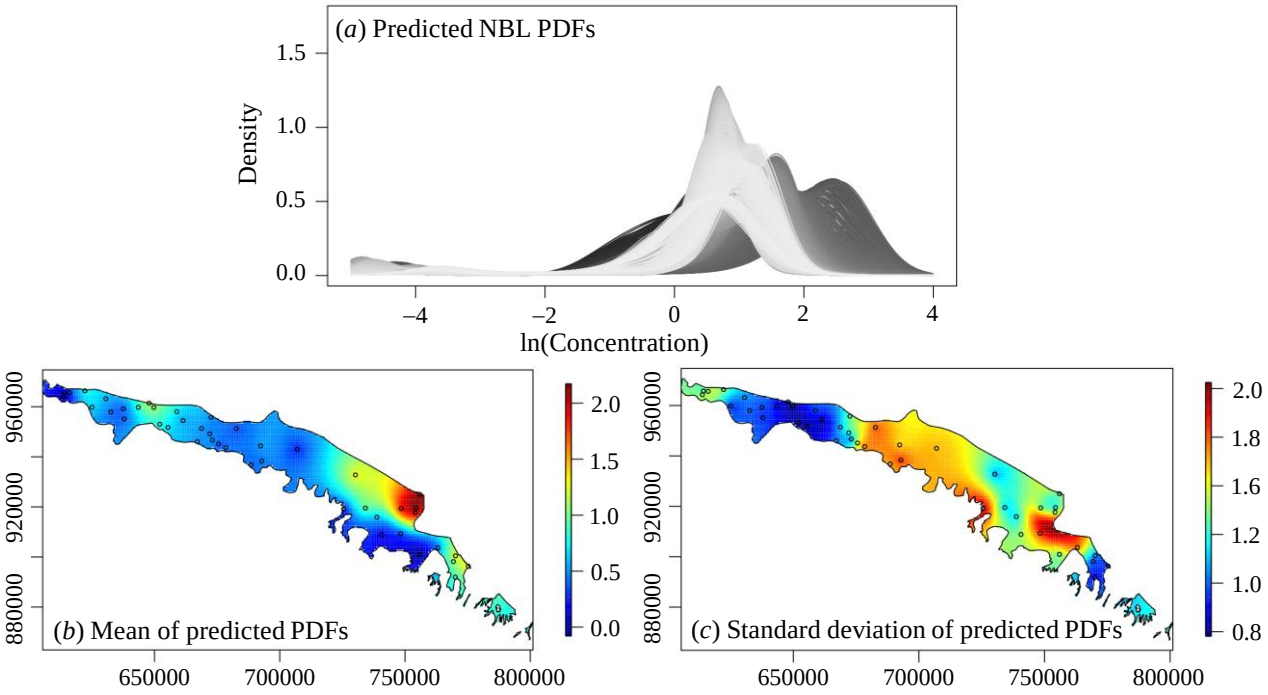

Figure 4. Kriging prediction of NBL densities for ammonium log-concentrations: (a) kriged/predicted densities; (b) mean values, and (c) standard deviation estimated from the kriged densities.

A sector characterized by low mean and high standard deviation values is visible in the south of this area. This result is consistent with the documented pattern associated with experimental data in this region, which are characterized by a temporal evolution displaying high concentration values within a collection of otherwise low values. The central portion of the domain is characterized by modest mean concentration values with high standard deviations. Low mean values and low to moderate values of standard deviation are found within the western area. It is noted that demarcation of zones linked to differing behaviors of the target chemical species is one of the key advantages of the functional analysis approach we employ. Local high values of ammonium are consistent with the documented natural occurrence of paleo-peats (Amorosi et al., 1996; Cremonini et al., 2008) in sample cores collected at other locations across the area of interest, with an overall tendency of ammonium concentrations to increase with depth and with increasing thickness of the fine deposits that confine the aquifer. Further large scale sampling campaigns would be required for a detailed assessment of correspondences with specific local conditions.

Figure 5a depicts the predicted spatial variability of the 90% quantile of the NBL concentration. These results are complemented by Figure 5b, where we depict the spatial distribution of the probability of exceeding the reference NBL value of 4.6 mg/l, which was suggested by Molinari et al. (2012) as representative of the global chemical status of the system upon relying on the classical PS procedure, as proposed by Wendland et al. (2005) and described in Section 2.2.1. The stark variability displayed by the 90^{th} percentile across the domain documents the presence of sectors within which the target chemical species shows differing behavior and suggests the need for considering spatially variable local NBL values. Our results indicate that the probability of exceeding the reference NBL value of 4.6 mg/l is very low across most of the domain, high probability of exceedance being confined within a limited portion of the system.

We note that our results are in general agreement with the findings of Molinari et al. (2019), where areas where such probability was evaluated above 80% are slightly wider than in our findings, while being located in the same sector. We remark that the approach employed by these authors (*i*) is based solely on summary statistics and not on the entire PDFs and (*ii*) relies on a Gaussian assumption to represent (log-transformed) NBL concentrations. Additionally, it is noted that data associated with boreholes with less than 10 records (after PS) were excluded from our analysis to allow for PDF reconstruction and interpretation, while some of these were retained by Molinari et al. (2019). Finally, we highlight that our approach is fully compatible with the possibility of resorting to a multimodel analysis to comprise uncertainty about the choice of the functional format for the variogram model (see e.g., Molinari et al. (2019)). While this element can be of interest, we focus here on the main innovative aspect of our study, which is related to the treatment of the data within the context of a functional geostatistical approach.

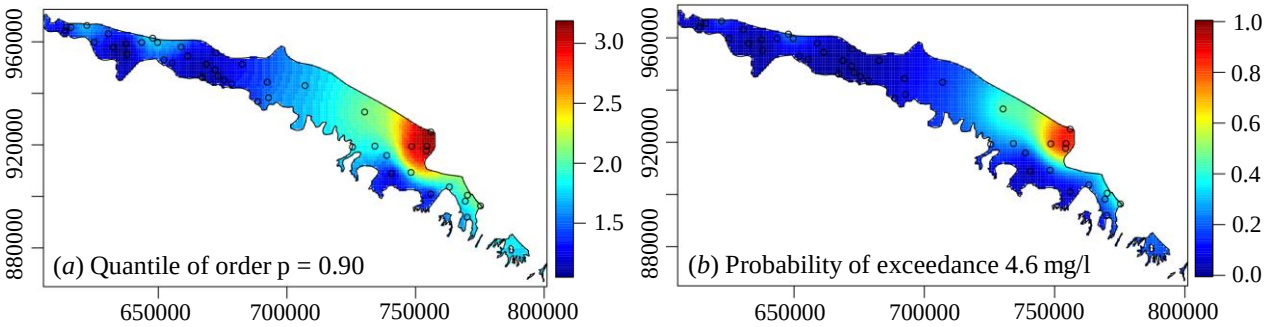

Figure 5. Spatial distributions of predicted (a) quantile of order 90% and (b) probability of exceedance of an ammonium concentration threshold of 4.6 mg/l.

The approach illustrated in Section 2.2.5 was then applied to the smoothed density data projected on the basis generated by the first k = 8 principal components (explaining 99.99% of the total variability) to generate a collection of random realizations of spatial distributions of NBL log-concentration values. The scores x_{s_1,k_1}, x_{s_2,k_2} were modeled as uncorrelated for $k_1 \neq k_2$ and $s_1 \neq s_2$ in the domain, as supported by visual inspection of cross-variograms (not shown). An exponential model was calibrated to the empirical variogram for each spatial field of scores. Conditional Gaussian simulations were performed to yield a Monte Carlo (MC) collection of 100 realizations. The practical implementation relies on the adoption of sequential Gaussian simulation (Abrahamsen and Benth, 2001) as implemented within the R package gstat (Pebesma, 2004), and setting a local neighborhood of 60 km to reduce computational burden. The collection of NBL distributions was then built from the MC ensemble of scores.

Figure 6 depicts a realization of the spatial field of NBL densities (Figure 6a), the spatial distributions of the 90% quantile (Figure 6b) and the probability of exceeding the threshold value of 4.6 mg/l (Figure 6c). Similar to what we observed in Figure 5, the overall spatial pattern in Figures 6b, c is generally consistent with the results presented by Molinari et al. (2019) (see their Figure 2) and reinforces the concept that assigning a unique NBL value for a given chemical species to a large scale groundwater body can conceal the possibility of identifying regions with high (or low) geogenic contribution. These could in turn be ascribed to low (or high) anthropogenic activity, thus

potentially biasing expectations about results of groundwater protection measures. We recall that Molinari et al. (2019) (*a*) rely on the stringent assumption that the probability density of (log)concentrations can be described through a Gaussian model, and (*b*) parametrize the latter on the basis of kriging results relying solely on summary statistics evaluated from the available data. Rather, we are not limited by any assumption about the specific functional format of probability densities, which are entirely data-driven and are the object of the geostatistical analysis. As such, the tools and implementation workflow we propose is conducive to evaluations of the spatially heterogeneous field of NBL values in a probabilistic context upon maximizing the use of the amount of information embedded in the available data. This is seen as a critical element of a modern decision-making approach grounded on a firm environmental risk assessment practice. Future integration of these findings with other types of (hydro)geological and geochemical information can then yield a complete picture of the natural signature of the system analyzed.

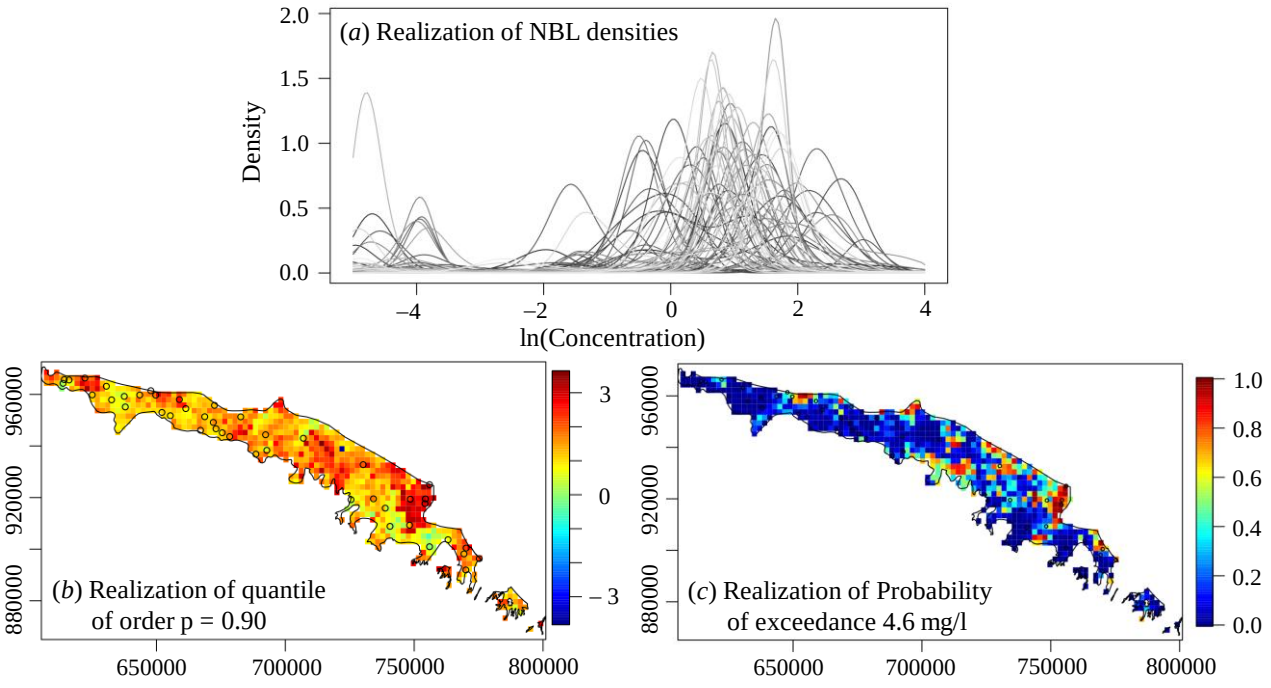

Figure 6. Example of a realization obtained from the (conditional) stochastic simulation of NBL distributions of ammonium log-concentrations. (a) Simulated NBL densities and corresponding spatial distributions, (b) 90% quantiles, (c) probability of exceeding a threshold value of 4.6 mg/l.

3.2. Arsenic

A total of 1096 data collected at 43 monitoring station were available for arsenic after PS (see Section 2.1), with a number of observations per sampling point ranging between 11 to 38 (with an average of about 25). Estimation of the PDF of the NBL concentrations is performed at each borehole location consistently with the approach exemplified in Section 3.1. The resulting smoothed data are depicted in Figure 7.

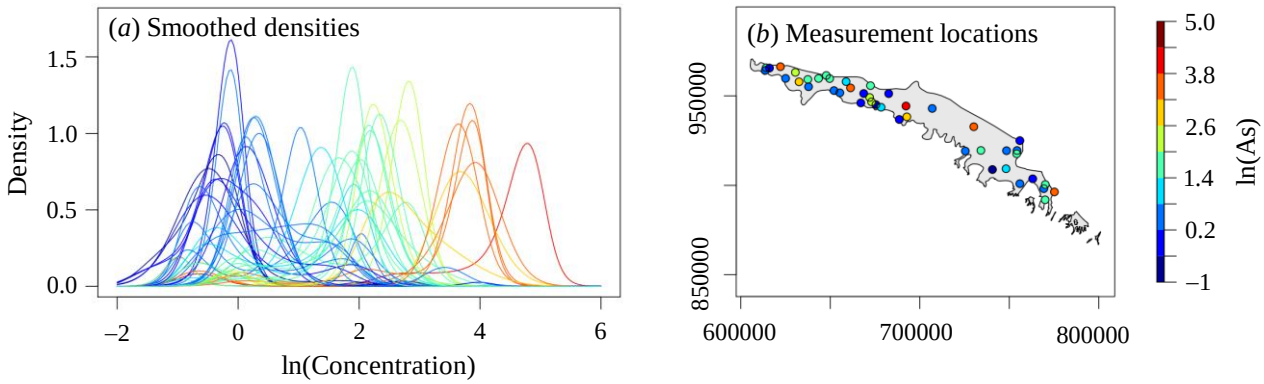

Figure 7. Smoothed data for arsenic log-concentrations (a), and corresponding spatial locations in the investigated aquifer system (b). Colors are assigned according to the value of the mean related to the corresponding smoothed density.

Visual inspection of Figure 7 suggests a significant spatially heterogeneous behavior. The highest mean values are scattered across the whole domain, suggesting that these could be associated with local conditions. These types of results are consistent with the behavior of arsenic, that is typically documented to display a remarkably high degree of spatial variability within a given groundwater body (e.g., Duan et al., 2017; Pi et al., 2018; Smith et al., 2003; Liang et al., 2017, 2018, 2019).

The sample trace-variogram associated with the available densities is depicted in Figure 8, its pattern supporting a global stationarity assumption. An exponential model with nugget was calibrated to the empirical variogram, its estimated parameters being listed in Figure 8. The contribution of the nugget to the total variance is equal to 15%, suggesting the occurrence of variability between sample pairs separated by short distance.

The available smoothed densities were then estimated through Kriging at the same set of unsampled locations considered in the ammonium case, grid spacing corresponding to a discretization of the variogram range (Figure 8) with about 10 points. Cross-validation results (Appendix A) fully support the satisfactory performance of the approach.

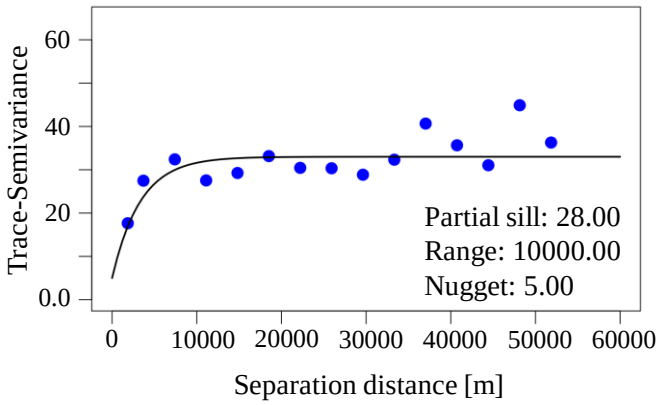

Figure 8. Sample trace-variogram estimated from the smoothed functional data for arsenic log-concentrations and interpreted model with estimated parameters.

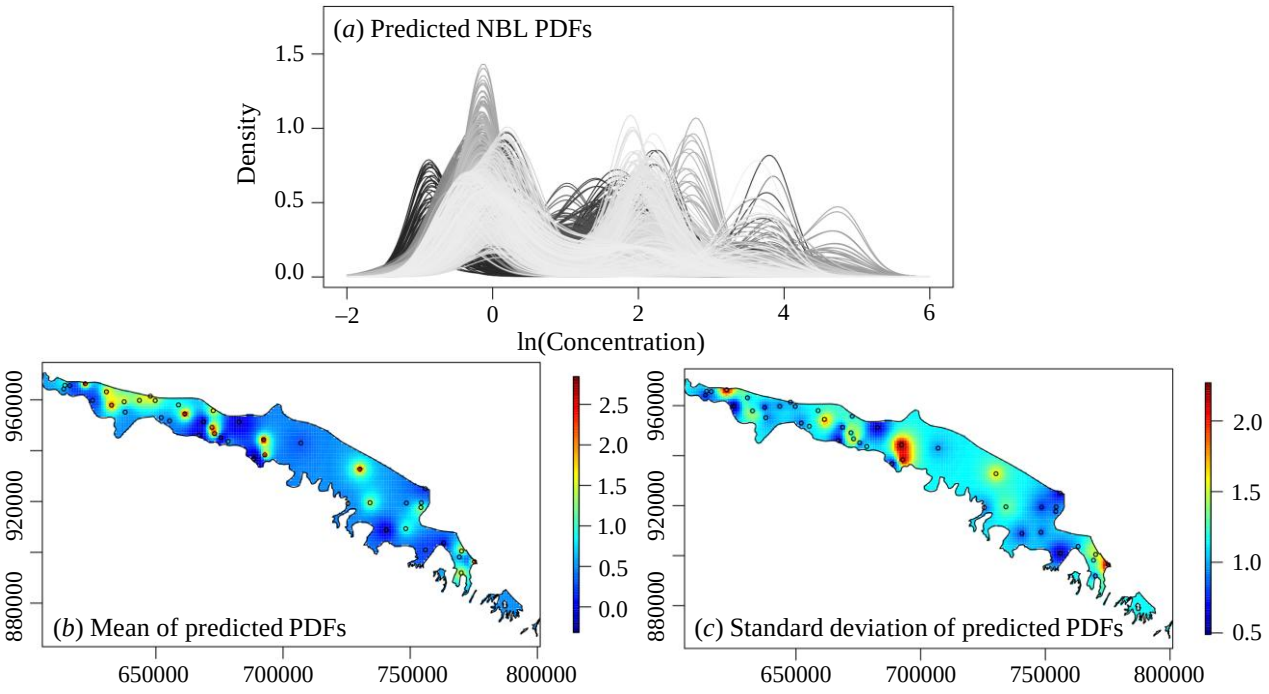

Figure 9. Kriging prediction of NBL densities for arsenic log-concentrations: (a) kriged/predicted density functions; (b) spatial distribution of mean values, and (c) standard deviation estimated from the kriged densities.

Figure 9a depicts the predicted (i.e., based on functional Kriging results) PDFs of NBL log-concentrations of arsenic. Figures 9b and 9c depict the estimates of mean value and standard deviation of NBL densities, respectively. Moderate to high mean values are mostly located in the north-western and central portions of the domain. The associated standard deviation varies from moderate to high values. Areas characterized by high values of the mean value of predicted PDFs appear to be localized in the surrounding of some measurement stations rather than being spread across extended sectors of the domain. This finding is also consistent with possible occurrences of lateral variations of arsenic concentrations, similar to other documented studies across several regions worldwide.

Figure 10a depicts the spatially heterogeneous distribution of the predicted quantile of order 90% of the NBL As log-concentrations. To complement these results, Figure 10b shows the probability of exceeding the reference NBL value of 33 µg/l, which had been evaluated by Molinari et al. (2012) as representative of the global chemical status of the system through the classical PS procedure (Wendland et al., 2005). We found that the probability of exceeding such a threshold value is very modest throughout the system, with the exception of some localized spots where it attains moderate values. This has a clear consequence on the assessment of the chemical status of the system, which would have been (deterministically) classified as requiring attention on the basis of such a performance metric.

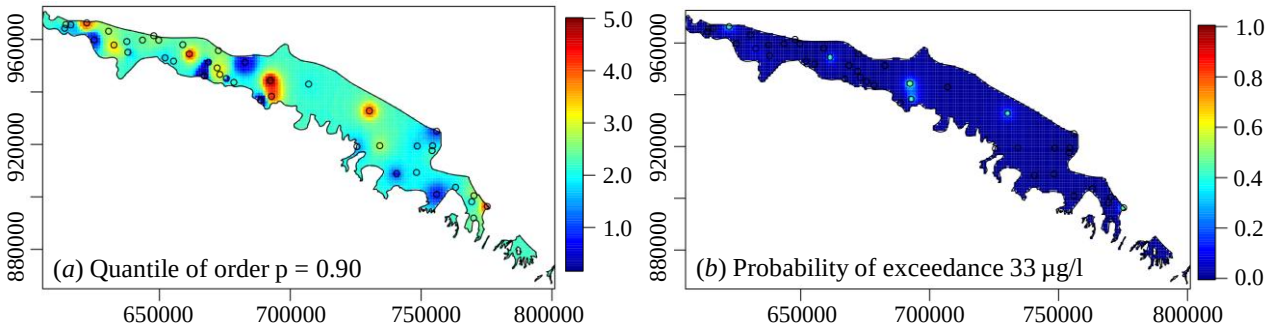

Figure 10. Spatial distribution of predicted (a) 90% quantile and (b) probability of exceedance of an arsenic concentration threshold of 33 µg/l.

We note that Molinari et al. (2019) could not provide spatial maps of exceedance probabilities, because their analysis, grounded solely on summary quantities, resulted in a pure nugget semivariogram. Our results suggest that considering a functional analysis approach might enable one to observe the emergence of some degree of spatial correlation when the complete density associated with observations is embedded in the methodology.

Similar to the case of ammonium, we applied the stochastic simulation approach described in Section 2.2.5 to the smoothed density data projected on the basis generated by the first $k = 8$ principal components (explaning 99.99% of the total variability). Modeling of the scores and of the ensuing empirical variogram is performed as described in Section 3.1 for ammonium, conditional Gaussian simulations being then performed to yield a collection of 100 MC realizations. The same stochastic generation methodology described in Section 3.1 has been employed, a local neighborhood of 60 km being set to alleviate computational time. Figure 11 depicts a selected realization of the spatial field of NBL densities (Figure 11a), the corresponding spatial distributions of quantiles of order 90% (Figure 11b), and the probability of exceeding the threshold value of 33 µg/l (Figure 11c). The occurrence of localized spots associated with significant probability of high natural arsenic concentrations are consistent with the documented presence at some depths in the aquifer system of sediments whose composition includes a vegetal-rich fraction (see, e.g., Molinari et al., 2013, 2014). These types of solid matrices are prone to potentially adsorb significant arsenic amounts that can then be mobilized by variations of redox conditions (see, e.g., Molinari et al., 2013, 2014, 2015). A detailed analysis to evaluate possible relationships and consistency with local conditions would require additional large scale sampling campaigns which can be subject of future studies.

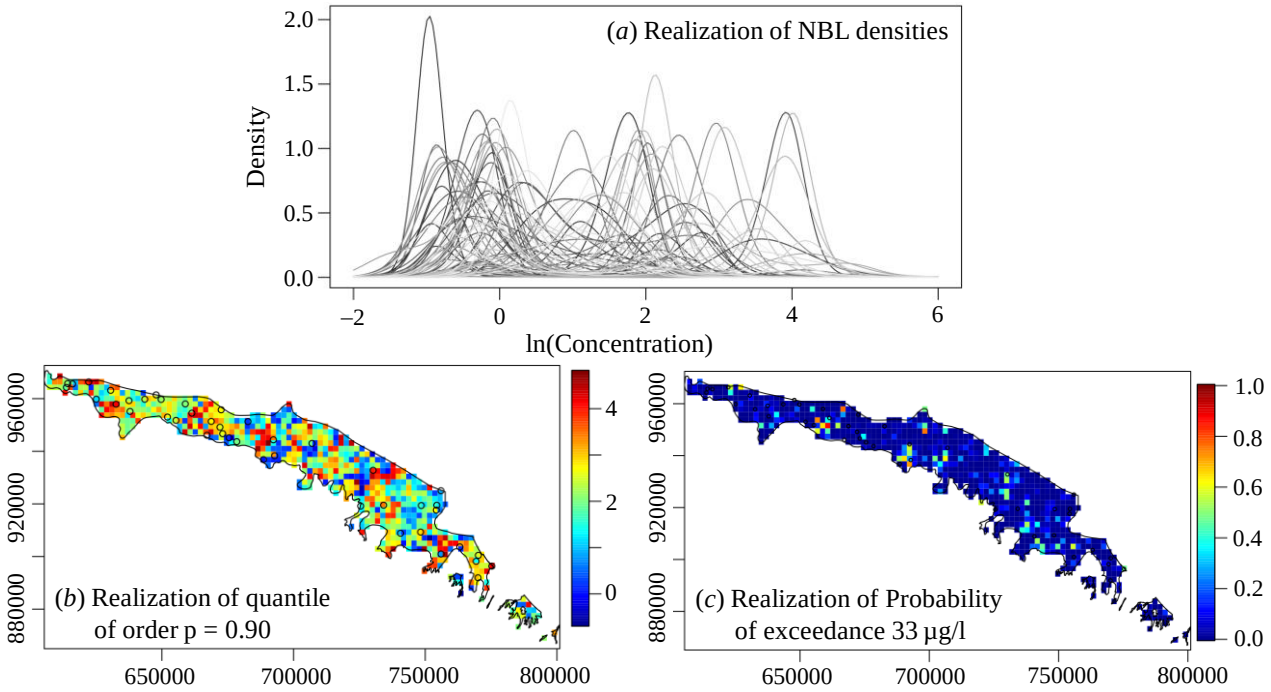

Figure 11. Example of a realization obtained from the (conditional) stochastic simulation of NBL distributions of arsenic. (a) Simulated NBL densities and corresponding spatial distributions of (b) 90% quantiles, and (c) probability of exceeding the threshold value of 33 µg/l.

4. Conclusions

We propose and apply a theoretical framework and the ensuing operational workflow to obtain a rigorous (probabilistic) assessment of Natural Background Levels (NBL) of concentrations of target chemical species in large-scale groundwater bodies, which are usually characterized by a high degree of heterogeneity of sedimentological and hydrogeochemical conditions. Our approach enables one to fully consider the richness of information embedded in the available historical records of routinely monitored concentrations which are then typically employed (e.g., by Environmental Agencies) to assess the chemical status of a groundwater body. On these bases, we suggest a change of perspective in the way one should consider evaluating NBL concentrations in a modern probabilistic risk assessment context. Rather than focusing on selected (statistical) moments or percentiles (i.e., summary statistics) evaluated on the basis of sample probability distributions of concentrations at individual boreholes, we associate with each monitoring station the entire distribution of NBL concentrations. The latter is represented through its (estimated) density

518 function, which we model as a random point in a Bayesian Hilbert space and then analyze in the
519 context of Object Oriented Data Analysis. The merits of the approach are exemplified through an
520 application targeting the evaluation of the main characteristics of the spatial variability of the NBLs
521 of two selected chemical species (ammonium and arsenic) within a large scale groundwater body in
522 Northern Italy.

523 Our study leads to the following major conclusions.

524 1. The approach enables one to identify local trends within a given groundwater body, as
525 quantified in terms of spatial heterogeneity of NBL concentrations, in a probabilistic
526 context, without being limited to relying solely on selected quantiles of the distribution of
527 concentrations extracted from historical records. As such, it is possible to demarcate sectors
528 where distinct NBL spatial patterns emerge from an average system behavior, to be then
529 integrated within a decision-making activity.
530 2. The approach is fully consistent with modern requirements of tailoring the objective of
531 environmental actions to spatially varying conditions. This forms the platform to set
532 appropriate and cost-effective remediation goals and actions for deteriorated groundwater
533 bodies which account for the complete set of information embedded in the historical records.
534 Relying on rigorously assessed spatial distributions of probabilities of exceeding given NBL
535 concentration thresholds hampers the risk of assigning exceedingly high values of natural
536 background concentrations to areas subject to anthropogenic activities or otherwise setting
537 very low background levels within regions where the geogenic contribution can be
538 significant. Lack of consideration of these elements could lead to setting unrealistic
539 remediation goals.
540 3. Having the ability to generate multiple conditional spatial realizations of NBL densities
541 enables a complete uncertainty quantification (see our exemplary results in Section 3) which
542 would be otherwise impossible with standard methods of analysis currently adopted in
543 practical applications targeting large scale groundwater bodies. These elements are markedly

relevant in such systems, whose hydrogeologic, lithologic, and geochemical characteristics can be associated with large spatial heterogeneity.

Key values of the study are methodological as well as operational. From a methodological standpoint, the workflow we propose includes elements of innovation which go beyond limitations of other typically used approaches, including the possibility of effectively using the full information content embedded in data which are routinely monitored by local authorities and public environmental agencies. From an operational standpoint, it provides an appraisal of the probability that a given threshold value of concentration of geogenic origin can be exceeded in the exemplary areas considered. The ability to provide a robust and data-driven quantification of probability of exceedance provides an important element of flexibility in decision-making under uncertainty. The nature of the approach allows accounting for specific local needs, as viewed in the broad regional context, as well as the possibility of updating the results of the analysis as data become available. As such, it enables one to structure corrective actions according to levels of priorities related to target concentration thresholds and associated probability distributions linked to specific areas, which might be characterized by distinct local requirements. In this sense, our results can provide a support to identify localized areas where detailed hydrogeological studies can be promoted with the aim, e.g., to constrain uncertainty associated with predicted NBL values and associated probability of exceedance.

Appendix A

The performance of the proposed approach is assessed through a leave-one-out cross-validation (LOO CV) analysis. Here, for each site s_i in D, the PDF of the NBL PDF X_{s_i} is left out of the sample and a training set built upon all of the other NBL PDFs, $\left\{X_{s_j}, j \neq i\right\}$, is considered for calibration of the geostatistical model, following the same steps and parameter settings as in Section 3. Kriging is then used to predict the left-out NBL PDF X_{s_i}, yielding a prediction $X_{s_i}^{*(-i)}$. The prediction error for each site is evaluated through the sum of squared errors (SSE) as

$$\mathrm{SSE}(X_{s_i}) = d_{B^2}^2\left(X_{s_i}, X_{s_i}^{*(-i)}\right) = \frac{1}{2(b-a)}\int_a^b\int_a^b \ln\left(\frac{X_{s_i}(t)X_{s_i}^{*(-i)}(s)}{X_{s_i}(s)X_{s_i}^{*(-i)}(t)}\right)\mathrm{d}t\mathrm{d}s. \quad \text{(A1)}$$

Table A1 lists the summary statistics of SSE, as assessed via LOO CV for ammonium (first row) and arsenic (second row). It is noted that the LOO CV analyses for these chemical species are performed separately. Overall, the order of magnitude of the errors is fully consistent with the estimated sills of the trace-variograms (estimated sills: 36.35 and 33.00 for ammonium and arsenic, respectively).

Chemical Species	Min	Q1	Median	Mean	Q3	Max
Ammonium	8.81	18.16	26.58	29.97	40.01	109.08
Arsenic	9.84	19.91	26.96	35.43	36.32	134.37

Table A1: Summary statistics for SSE (A1).

The LOO CV analysis is additionally used to evaluate the ability of our conditional simulation theoretical approach and operational workflow to represent prediction uncertainty. As an example, Figure A1 depicts the results obtained at two locations (denoted as FC17−01 and RE17−03) for ammonium (top panels) and arsenic (bottom panels). Predicted NBL PDFs at these locations are depicted with dashed black curves, whereas grey curves correspond to the $B = 100$ conditional simulations at the site. The observed PDFs are represented as thick black curves. Visual inspection of Figure A1 suggests that conditional simulations well represent the uncertainty associated with the

predictions for both chemical species. For instance, one can observe that, even as the kriging error for arsenic at location RE17-03 appears to be quite high, the conditional simulations at the site suggest that a high uncertainty is associated with the prediction. Note that here, the test NBL PDF is well captured by the simulated collection of realizations.

To quantitatively assess the performance in terms of uncertainty quantification, we then compute the distance between the test curve and the ensemble of Monte Carlo simulations as

$$\mathrm{SSE}_{\mathrm{sim}}\left(\mathrm{X}_{\mathrm{s}_\mathrm{i}}\right) = \min\left\{\mathrm{d}^2_{\mathrm{B}^2}\left(\mathrm{X}_{\mathrm{s}_\mathrm{i}}, \mathrm{X}^{\mathrm{b}(-\mathrm{i})}_{\mathrm{s}_\mathrm{i}}\right), \mathrm{b} = 1, \ldots \mathrm{B}\right\} \quad \text{(A2)}$$

where $\mathrm{X}^{\mathrm{b}(-\mathrm{i})}_{\mathrm{s}_\mathrm{i}}$ denotes the b-th conditional simulation when the i-th observation is left out of the sample. In practice, the smaller $\mathrm{SSE}_{\mathrm{sim}}\left(\mathrm{X}_{\mathrm{s}_\mathrm{i}}\right)$, the closer the ensemble is to the test observation $\mathrm{X}_{\mathrm{s}_\mathrm{i}}$. For instance, values of $\mathrm{SSE}_{\mathrm{sim}}\left(\mathrm{X}_{\mathrm{s}_\mathrm{i}}\right)$ for ammonium at locations FC17-01 and RE17-03 are 5.96 and 7.56, respectively, their counterparts corresponding to arsenic being 16.46 and 10.28, respectively. Table 2 lists the summary statistics associated with $\mathrm{SSE}_{\mathrm{sim}}$ for both chemical species. One can note that the ensemble is typically quite close to the test observation, with an average $\mathrm{SSE}_{\mathrm{sim}}$ of 10.67 and 9.02 for ammonium and arsenic, respectively.

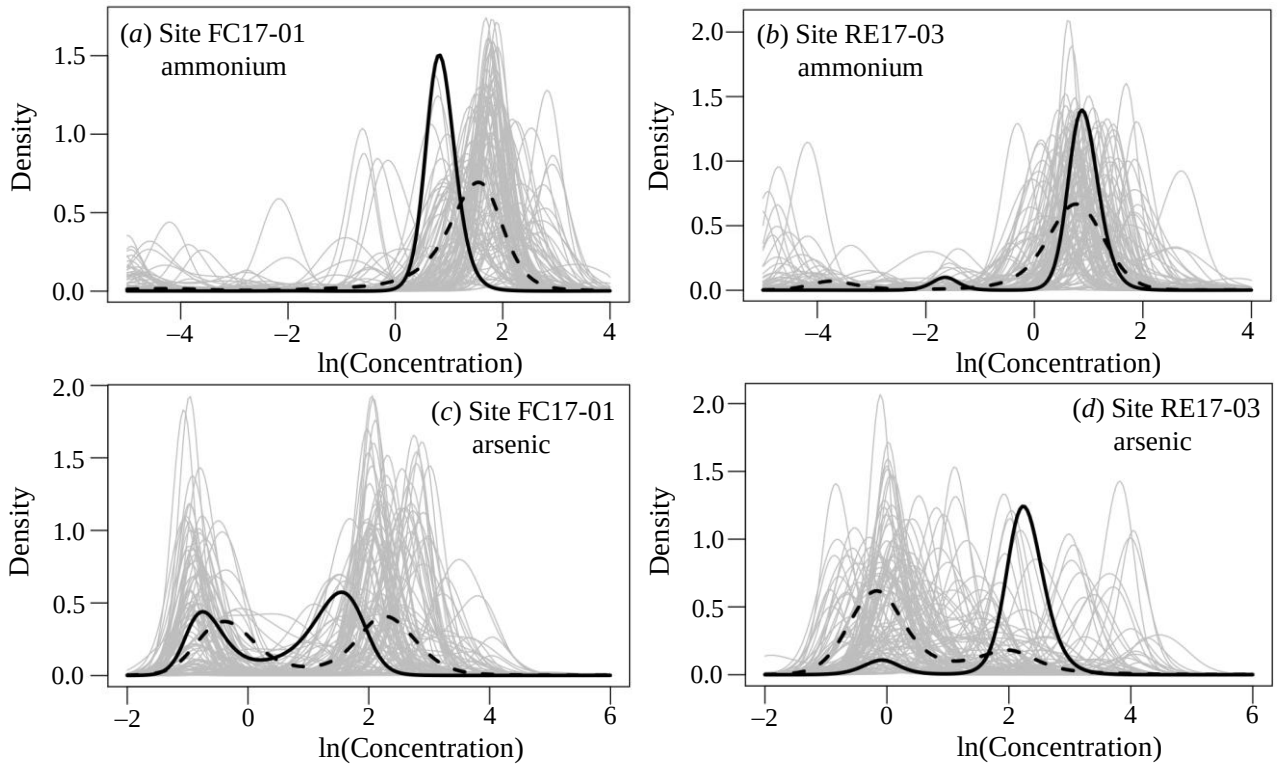

Figure A1. Leave-one-out cross-validation results at sites FC17-01 and RE17-03. Dashed black curves correspond to predicted NBL PDFs, whereas grey curves correspond to the $B = 100$ conditional simulations at the site; observed PDFs are represented as thick black curves.

Chemical Species	Min	Q1	Median	Mean	Q3	Max
Ammonium	2.30	5.70	7.88	10.67	13.15	41.92
Arsenic	1.05	5.31	7.14	9.023	10.56	30.38

Table A2: Summary statistics for $\mathrm{SSE}_{\mathrm{sim}}$ *(A2).*

References

Aitchison, J., 1986. The statistical analysis of compositional data. Chapmann and Hall, New York.

Abrahamsen, P., Benth, F., 2001. Kriging with inequality constraints. Math. Geol. 33(6), 719-744. https://doi.org/10.1023/A:1011078716252.

Amorosi, A., Farina, M., Severi, P., Preti, D., Caporale, L., Di Dio, G., 1996. Genetically related alluvial deposits across active fault zones: an example of alluvial fan-terrace correlation from the upper Quaternary of the southern Po Basin, Italy. Sediment. Geol. 102, 275-295. https://doi.org/10.1016/0037-0738(95)00074-7.

Bianchi Janetti, E., Guadagnini, L., Riva, M., Guadagnini, A., 2019. Global sensitivity analysis of multiple conceptual models with uncertain parameters driving groundwater flow in a regional-scale sedimentary aquifer. J. Hydrol. 574, 544-556. https://doi.org/10.1016/j.jhydrol.2019.04.035

BRIDGE - Background cRiteria for the IDentification of Groundwater thrEsholds 2007. http://nfp-at.eionet.europa.eu/irc/eionet-circle/bridge/info/data/en/index.htm (accessed 21 February 2020); or https://cordis.europa.eu/result/rcn/51965_en.html (accessed 21 February 2020).

Chilès, J.P., Delfiner, P., 1999. Geostatistics: Modeling spatial uncertainty, John Wiley & Sons, New York.

621 Cremonini, S., Etiope, G., Italiano, F., Martinelli, G., 2008. Evidence of possible enhanced peat
622 burning by deep-origin methane in the Po River delta Plain (Italy). J. Geol. 116, 401-413.
623 https://doi.org/10.1086/588835.
624 Dalla Libera, N., Fabbri, P., Mason, L., Piccinini, L., Pola, M., 2017. Geostatistics as a tool to
625 improve the natural background level definition: an application in groundwater. Sci. Total
626 Environ. 598, 330-340. https://doi.org/10.1016/j.scitotenv.2017.04.018.
627 Decreto Legislativo n. 30 del 16 marzo 2009 (Legislation Decree n. 30, 16 March, 2009).
628 Application of the Directive 2006/118/CE, related to the protection of groundwater
629 resources from pollution and deterioration (Attuazione della direttiva 2006/118/CE, relativa
630 alla protezione delle acque sotterranee dall'inquinamento e dal deterioramento). Gazzetta
631 Ufficiale n. 79 of 4 April 2009 (in Italian).
632 Directive 2000/60/EC - Water Framework Directive (WFD). Directive of the European Parliament
633 and of the Council of 23 October 2000 establishing a framework for Community action in
634 the field of water policy, OJ L327, 22 Dec 2000, 1-73.
635 Directive 2006/118/EC, GroundWater Daughter Directive (GWDD). Directive of the European
636 Parliament and of the Council of 12 December 2006 on the protection of groundwater
637 against pollution and deterioration, OJ L372, 27 Dec 2006, 19-31.
638 Directive 2014/80/EU amending Annex II to Directive 2006/118/EC of the European Parliament
639 and of the Council on the Protection of Groundwater Against Pollution and Deterioration,
640 OJ L182, 21 June 2014, 52-55.
641 Duan, Y., Gan, Y., Wang, Y., Liu, C., Yu, K., Deng, Y., Zhao, K., Dong, C., 2017. Arsenic
642 speciation in aquifer sediment under varying groundwater regime and redox conditions at
643 Jianghan Plain of Central China. Sci. Total Environ. 607-608, 992-1000.
644 https://doi.org/10.1016/j.scitotenv.2017.07.011.
645 Ducci, D., Condesso de Melo, M.T., Preziosi, E., Sellerino, M., Parrone, D., Ribeiro, L., 2016.
646 Combining natural background levels (NBLs) assessment with indicator kriging analysis to

improve groundwater quality data interpretation and management. Sci. Total Environ. 569-570, 569-584. https://doi.org/10.1016/j.scitotenv.2016.06.184.

Edmunds, W.M., Shand, P., Hart, P., Ward, R.S., 2003. The natural (baseline) quality of groundwater: a UK pilot study. Sci. Total Environ. 310 (1-3), 25-35. https://doi.org/10.1016/S0048-9697(02)00620-4.

Egozcue, J. J., Díaz-Barrero, J.L., Pawlowsky-Glahn, V., 2006. Hilbert space of probability density functions based on Aitchison geometry. Acta Math. Sin. Engl. Ser. 22(4), 1175- 1182. https://doi.org/10.1007/s10114-005-0678-2.

Farina, M., Marcaccio, M., Zavatti, A., 2014. Experiences and perspectives to monitoring of groundwater resources: the contribution of Emilia Romagna (Esperienze e prospettive nel monitoraggio delle acque sotterranee: il contributo dell'Emilia-Romagna), Pitagora Editrice, Bologna (in Italian).

Hinsby, K., Condesso de Melo, M.T., 2006. Application and evaluation of a proposed methodology for derivation of groundwater threshold values-a case study summary report. In Deliverable D22 of the EU project ‘‘BRIDGE’’ 2006. http://nfp-at.eionet.europa.eu/Public/irc/eionet-circle/bridge/library?l=/deliverables/d22_final_reppdf/_EN_1.0_&a=d. (accessed 21 February 2020).

Hinsby, K., Condesso de Melo, M.T., Dahl, M., 2008. European case studies supporting the derivation of natural background levels and groundwater threshold values for the protection of dependent ecosystems and human health. Sci. Total Environ. 401(1-3), 1-20. https://doi.org/10.1016/j.scitotenv.2008.03.018.

Hron, K., Menafoglio, A., Templ, M., Hruzova, K., Filzmoser, P., 2016. Simplicial principal component analysis for density functions in Bayes spaces. Comput. Stat. Data An. 94, 330–350. https://doi.org/10.1016/j.csda.2015.07.007.

671 Kim, K.H., Yun, S.T., Kim, H.K., Kim, J.W., 2015. Determination of natural backgrounds and
672 thresholds of nitrate in South Korean groundwater using model-based statistical approaches.
673 J. Geochem. Explor. 148, 196-205. https://doi.org/10.1016/j.gexplo.2014.10.001.
674 Kim, H.R., Kim, K.H., Yu, S., Moniruzzaman, M., Hwang, S.I., Lee, G.T., Yun, S.T., 2019. Better
675 assessment of the distribution of As and Pb in soils in a former smelting area, using ordinary
676 co-kriging and sequential Gaussian co-simulation of portable X-ray fluorescence (PXRF)
677 and ICP-AES data. Geoderma 341, 26-38. https://doi.org/10.1016/j.geoderma.2019.01.031.
678 Li, P., Qian, H., Howard, K.W.F., Wu, J., Lyu, X., 2014. Anthropogenic pollution and variability of
679 manganese in alluvial sediments of the Yellow River, Ningxia, northwest China. Environ.
680 Monit. Assess. 186(3), 1385-1398. https://doi.org/10.1007/s10661-013-3461-3.
681 Liang, C.P., Hsu, W.S., Chien, Y.C., Wang, S.W., Chen, J.S., 2019. The combined use of
682 groundwater quality, drawdown index and land use to establish a multi-purpose groundwater
683 utilization plan. Water Resour. Manag. 33, 4231-4247, https://doi.org/10.1007/s11269-019-
684 02360-2.
685 Liang, C.P., Chen, J.S., Chien, Y.C., Jang, C.S., Chen, C.F., 2018. Spatial analysis of the risk to
686 human health from exposure to arsenic contaminated groundwater: a kriging approach. Sci.
687 Total Environ. 627, 1048-1057. doi: 10.1016/j.scitotenv.2018.01.294.
688 Liang, C.P., Chien, Y.C., Jang, C.S., Chen, C.F., Chen, J.S., 2017. Spatial analysis of human health
689 risk due to arsenic exposure through drinking groundwater in Taiwan's Pingtung Plain. Int.
690 J. Environ. Res. Public Health, 14, 81. doi:10.3390/ijerph14010081.
691 Machalová, J., Hron, K., Monti, G.S., 2016. Preprocessing of centred logratio transformed density
692 functions using smoothing splines. J. Appl. Stat. 43(8), 1419-1435.
693 https://doi.org/10.1080/02664763.2015.1103706.
694 Marron, J.S., Alonso, A.M., 2014. Overview of object oriented data analysis. Biometrical J. 56(5),
695 732-753. https://doi.org/10.1002/bimj.201300072.

696 Menafoglio, A., Secchi, P., Dalla Rosa, M., 2013. A Universal Kriging predictor for spatially
697 dependent functional data of a Hilbert Space. Electron. J. Statist. 7, 2209–2240.
698 https://doi.org/10.1214/13-EJS843.
699 Menafoglio, A., Guadagnini, A., Secchi, P., 2014. A Kriging approach based on Aitchison
700 geometry for the characterization of particle-size curves in heterogeneous aquifers. Stoch.
701 Environ. Res. Risk Assess. 28(7), 1835-1851. https://doi.org/10.1007/s00477-014-0849-8.
702 Menafoglio, A., Secchi, P., Guadagnini, A., 2016a. A Class-Kriging predictor for Functional
703 Compositions with application to particle-size curves in heterogeneous aquifers. Math.
704 Geosci. 48(4), 463-485. https://doi.org/10.1007/s11004-015-9625-7.
705 Menafoglio, A., Guadagnini, A., Secchi, P., 2016b. Stochastic Simulation of soil particle-size
706 curves in heterogeneous aquifer systems through a bayes space approach. Water Resour.
707 Res. 52, 5708–5726. https://doi.org/10.1002/2015WR018369.
708 Menafoglio, A., Secchi, P., 2017. Statistical analysis of complex and spatially dependent data: A
709 review of object oriented spatial statistics. Eur. J. Oper. Res. 258(2), 401-410.
710 https://doi.org/10.1016/j.ejor.2016.09.061.
711 Molinari, A., Guadagnini, L., Marcaccio, M., Guadagnini, A., 2012. Natural background levels and
712 threshold values of chemical species in three large-scale groundwater bodies in Northern
713 Italy. Sci. Total Environ. 425, 9-19. https://doi.org/10.1016/j.scitotenv.2012.03.015.
714 Molinari, A., Guadagnini, L., Marcaccio, M., Straface, S., Sanchez-Vila, X., Guadagnini, A., 2013.
715 Arsenic release from deep natural solid matrices under experimentally controlled redox
716 conditions. Sci. Total Environ. 444, 231-240.
717 https://doi.org/10.1016/j.scitotenv.2012.11.093.
718 Molinari, A., Ayora, C., Marcaccio, M., Guadagnini, L., Sanchez-Vila, X., Guadagnini, A. 2014.
719 Geochemical modeling of arsenic release from a deep natural solid matrix under alternated
720 redox conditions. Environ. Sci. Pollut. Res. 21, 1628-1637. https://doi.org/10.1007/s11356-
721 013-2054-6.

722 Molinari, A., Guadagnini, L., Marcaccio, M., Guadagnini, A., 2015. Arsenic fractioning in natural
723 solid matrices sampled in a deep groundwater body. Geoderma 247-248, 88-96.
724 https://doi.org/10.1016/j.geoderma.2015.02.011.
725 Molinari, A., Guadagnini, L., Marcaccio, M., Guadagnini, A., 2019. Geostatistical multimodel
726 approach for the assessment of the spatial distribution of natural background concentrations
727 in large-scale groundwater bodies. Water. Res. 149, 522-532.
728 https://doi.org/10.1016/j.watres.2018.09.049.
729 Panno, S.V., Kelly, W.R., Martinsek, A.T., Hackley, K.C., 2006. Estimating background and
730 threshold nitrate concentrations using probability graphs. Groundwater 44(5), 697-709.
731 https://doi.org/10.1111/j.1745-6584.2006.00240.x
732 Pebesma, E. J., 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences
733 30, 683-691. https://doi.org/10.1016/j.cageo.2004.03.012.
734 Perulero Serrano, R., Guadagnini, L., Riva, M., Giudici, M., Guadagnini, A., 2014. Impact of two
735 geostatistical hydro-facies simulation strategies on head statistics under non-uniform
736 groundwater flow. J. Hydrology 508, 343-355.
737 https://doi.org/10.1016/j.jhydrol.2013.11.009.
738 Pi, K., Wang, Y., Postma, D., Ma, T., Su, C., Xie, X., 2018. Vertical variability of arsenic
739 concentrations under the control of iron-sulfur-arsenic interactions in reducing aquifer
740 systems. J. Hydrol. 561, 200-210. https://doi.org/10.1016/j.jhydrol.2018.03.049.
741 Redman, A.D., Macalady, D. L., Ahmann, D., 2002. Natural organic matter affects arsenic
742 speciation and sorption onto hematite. Environ. Sci. Technol. 36, 2889-2896.
743 https://doi.org/10.1021/es0112801.
744 Regione Emilia-Romagna, 2010. Council Decree (Delibera di Giunta) n. 350 of 8/02/2010,
745 Approval of the activities of the Emilia-Romagna Region related to the implementation of
746 Directive 2000/60/CE aiming at the design and adoption of the Management Plans of the
747 hydrographic districs Padano, Appennino settentrionale and Appennino centrale

(Approvazione delle attività della Regione Emilia-Romagna riguardanti l'implementazione della Direttiva 2000/60/CE ai fini della redazione ed adozione dei Piani di Gestione dei Distretti idrografici Padano, Appennino settentrionale e Appennino centrale). http://ambiente.regione.emilia-romagna.it/acque/temi/piani%20di%20gestione (In Italian, accessed 21 February 2020).

Regione Emilia-Romagna, ENI-AGIP, 1998. Riserve idriche sotterranee della Regione Emilia-Romagna. S.EL.CA, Firenze.

Reimann, C., Garrett, R.G., 2005. Geochemical background: concept and reality. Sci. Total Environ. 350, 12-27. https://doi.org/10.1016/j.scitotenv.2005.01.047.

Ricci Lucchi, F., 1984. Flysh, molassa, clastic deposits: traditional and innovative approaches to the analysis of north Apeninic basins (Flysch, molassa, cunei clastici: tradizione e nuovi approcci nell'analisi dei bacini orogenici dell'Appennino settentrionale). Cento Anni di Geologia Italiana. Volume Giubilare 1° centenario Soc. Geol. Ital., 279-295 (in Italian).

Short, M., Guadagnini, L., Guadagnini, A., Tartakovsky, D. M., Higdon, D., 2010. Predicting vertical connectivity within an aquifer system. Bayesian Analysis 5(3), 557-582. https://doi.org/10.1214/10-BA522.

Smith, J.V.S, Jankowski, J., Sammut, J., 2003. Vertical distribution of As (III) and As (V) in a coastal sandy aquifer: Factors controlling the concentration and speciation of arsenic in the Stuarts Point groundwater system, northern New South Wales. Australia. Appl. Geochem. 18(9), 1479-1496. https://doi.org/10.1016/S0883-2927(03)00063-5.

Urresti-Estala, B., Carrasco-Cantos, F., Vadillo-Pérez, I., Jiménez-Gavilán, P., 2013. Determination of background levels on water quality of groundwater bodies: A methodological proposal applied to a Mediterranean river basin (Guadalhorce River, Málaga, southern Spain). J. Environ. Manage. 117, 121-130. https://doi.org/10.1016/j.jenvman.2012.11.042.

Van den Boogaart, K. G., Egozcue, J.J., Pawlowsky-Glahn, V., 2014. Bayes Hilbert spaces. Aus. N.Z. J. Stat. 56, 171-194. https://doi.org/10.1111/anzs.12074.

774 Walter, T., 2008. Determining natural background values with probability plots. EU Groundwater
775 Policy Developments Conference, UNESCO, Paris, France, 13-15 Nov 2008.
776 https://orbi.uliege.be/bitstream/2268/76101/1/actes_colloques_cfhaih_nov08.pdf (accessed
777 21 February 2020).
778 Wendland, F., Hannappel, S., Kunkel, R., Schenk, R., Voigt, H.J., Wolter, R., 2005. A procedure to
779 define natural groundwater conditions of groundwater bodies in Germany. Water Sci.
780 Technol. 51(3-4), 249-257. https://doi.org/10.2166/wst.2005.0598.
781 Winter, C.L., Tartakovsky, D.M., Guadagnini, A., 2003. Moment differential equations for flow in
782 highly heterogeneous porous media. Surv. Geophys. 24 (1), 81-106.
783 https://doi.org/10.1023/A:1022277418570.

***conflict of Interest Statement**

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

*Credit Author Statement

Credit Statement

All Authors have contributed to the study and preparation of the manuscript.