Linking local natural background levels in groundwater to their generating hydrogeochemical processes in Quaternary alluvial aquifers

Chiara Zanotti[1], *Mariachiara Caschetto*[1], *Tullia Bonomi*[1], *Marco Parini*[2], *Giuseppa Cipriano*[3], *Letizia Fumagalli*[1] *and Marco Rotiroti*[1]*

[1]Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, 20126 Milan, Italy.

[2]Regione Lombardia, Direzione Generale Territorio e Protezione Civile, Struttura Risorse Idriche, Piazza Città di Lombardia 1, 20124 Milan, Italy.

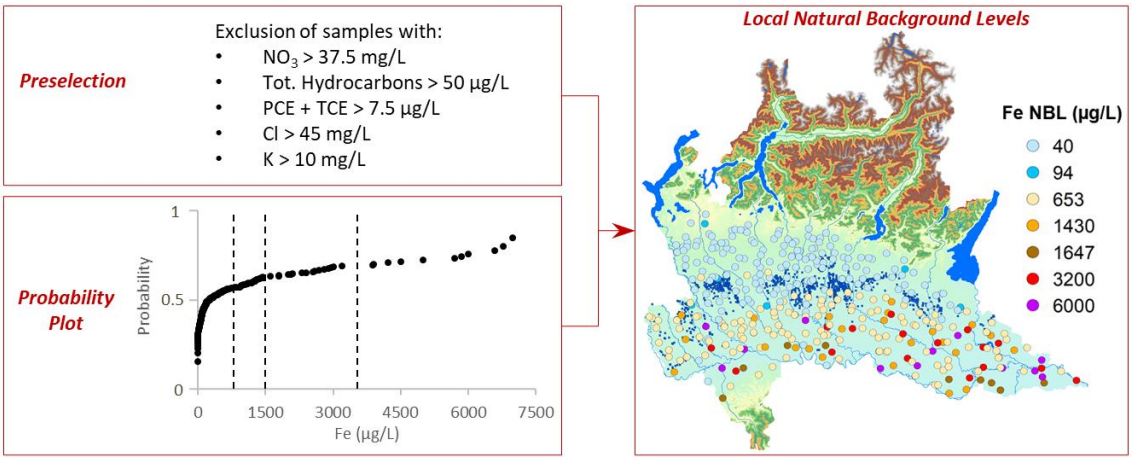
[3]Agenzia Regionale per la Protezione dell'Ambiente della Lombardia, Settore Monitoraggi Ambientali, Via Rosellini 17, 20124 Milan, Italy.

*corresponding author, email: marco.rotiroti@unimib.it

Science of the Total Environment 2022, https://doi.org/10.1016/j.scitotenv.2021.150259

Received 29 June 2021, Revised 25 August 2021, Accepted 6 September 2021, Available online 11 September 2021.

For the full version go to https://www.sciencedirect.com/science/article/pii/S0048969721053365



Abstract

Calculating natural background levels (NBLs) in groundwater is vital for supporting a sustainable use of groundwater resources. Although NBLs are often assessed through a unique concentration value per groundwater body, where hydrogeochemical features are highly variable, spatial heterogeneity needs to be accounted for, leading to the calculation of so-called "local" NBLs. Despite much research devoted to the identification of the best performing techniques for local NBLs spatialization, a deep understanding of the link between local NBL values and their generating hydrogeochemical processes is often lacking and so is addressed here for the redox-sensitive species As, NH_4, Fe and Mn in the groundwater bodies of Lombardy region, N Italy.

Local NBLs were calculated by a tired approach involving the hybridization of preselection and probability plot methods. Since the spatial variability of the target species depends mainly on redox conditions, a redox zonation was performed using multivariate statistical analysis. A conceptual model was developed and improved combing factor and cluster analysis. Results showed that NBLs for arsenic were up to 291 µg/L, reached in groundwaters under methanogenesis, a condition related to the prolonged degradation of peat buried in aquifer sediments. Ammonium NBLs up to 6.62 mg/L were generated by the upwelling of fluids from deep sediments hosting petroleum systems; ammonium NBLs up to 4.48 mg/L were generated as the accumulation of by-products of peat degradation. Iron and manganese NBLs up to, respectively, 6.0 and 1.51 mg/L were generated by the oxidation of younger and less stable Mn and Fe oxides within river valleys, mostly the Po River valley.

The evaluation of local NBLs, and their association to generating natural hydrogeochemical processes/conditions, achieves a step forward from the commonly used approach of a single NBL per groundwater body, improving decision-support tools for sustainable groundwater management and protection.

Keywords: Multivariate analysis, arsenic, ammonium, iron, manganese, Lombardy region.

1. Introduction

A proper management and protection of groundwater quality is fundamental for a sustainable use of groundwater resources. It represents not only an environmental but also a significant societal challenge, requiring science and policy makers additional efforts and more intersectoral strategies. A groundwater quality characterization, which is fundamental for supporting a sustainable groundwater management, relies on a robust evaluation of the

hydrogeological and hydrogeochemical components of the studied groundwater body (GWB). In this view, discretizing the nature of dissolved undesired species is of high relevance for tracing a contamination that can eventually turn into a pollution, thus producing adverse effects on humans and ecosystems (Chapman, 2007). Disproportionately high concentration of dissolved chemical species can reflect lithological, geochemical and hydrogeological intrinsic GWB features rather than anthropogenic pressures. It is self-evident how a false estimation of natural contributions to groundwater contamination can inevitably lead to some drawbacks on groundwater management, e.g., an improper chemical status classification of a GWB (EC, 2000, 2006). Consequently, natural background levels (NBLs) are the mainstay for a proper groundwater management plan, especially when target chemicals naturally attain concentrations above regulatory limits. The EU BRIDGE project (Background Criteria for the Identification of Groundwater Threshold; Muller et al., 2006) set a methodological guidance, at the European level, for individual Member States to assess NBLs for dissolved chemicals in a specific GWB. According to EU regulations (EC, 2000, 2006), GWBs are defined as subdivisions of large regional aquifers so that they can be properly managed. The BRIDGE project proposed the estimation of NBLs based on the component separation or the preselection methods (Hinsby et al., 2008; Muller et al., 2006; Wendland et al., 2005). An alternative commonly used method is the probability plot (Sinclair, 1974). An ongoing EU project called HOVER (Hydrogeological processes and geological settings over Europe controlling dissolved geogenic and anthropogenic elements in groundwater of relevance to human health and the status of dependent ecosystems) is proposing a method for NBLs calculation based on lithological classification and land-use analysis (Lions et al., 2021; Voutchkova et al., 2021). Most of all these methods assess a single NBL value for the entire GWB. However, GWBs have frequently complex hydrogeology and hydrogeochemistry making a single NBL somewhat speculative (Reimann and Garrett, 2005). Many efforts on boosting more reliable NBLs have been documented by the scientific community with the aim of incorporating the intrinsic local heterogeneities characterizing an aquifer system (Biddau et al., 2017; Dalla Libera et al., 2017; Ducci et al., 2016; Filippini et al., 2021; Gao et al., 2020; Guadagnini et al., 2020; Molinari et al., 2019), thus calculating variable NBLs within a GWB, that can be termed as “local NBLs” (Dalla Libera et al., 2018). Anyway, these research efforts were more dedicated toward identifying the best performing techniques for NBL spatialization rather than deeply understanding the link between NBLs values and natural hydrogeochemical processes determining them.

This work presents the assessment of local NBLs on a regional scale through a tired approach which combines preselection and probability plot methods. This methodology was applied to naturally high groundwater concentrations of As, NH_4, Fe and Mn in Lombardy region, N Italy. Since the spatial variability of these target species depends mainly on redox conditions, redox zonation was used as the main driver for local NBLs spatialization. The aim of this work is

to juxtapose a) the calculation of local NBLs with b) a thorough understanding of the natural processes which generated the NBLs.

2. Materials and methods

2.1. Preliminary conceptual model of the study area

The study covers the entire Lombardy region (23,844 km^2) in northern Italy (Fig. S1), in particular, its 30 identified GWBs (Regione Lombardia, 2016), that span from Alpine valley aquifers (10 GWBs) to alluvial aquifers in the Po Plain (20 GWBs; Fig. S1 and S2). The latter are subdivided into shallow (13 GWBs), intermediate (6 GWBs) and deep (1 GWB) aquifer systems (Regione Lombardia, 2016; Fig. S1 and S2). Within the Po Plain alluvial system, each GWB can be considered as an open system, and thus groundwater (and solutes) exchanges can occur between laterally and vertically adjacent GWBs (Fig. S2). Main human uses of groundwater resources in the study area are related to drinking water supply, irrigation, animal husbandry, industry and domestic needs.

The whole alluvial aquifer system of the Po Plain comprises Pleistocene sediments that prograded from W to E, overlain by Holocene fluvial sediments only within river valleys cut into the Pleistocene sediments (Garzanti et al., 2011; Marchetti, 2002). The aquifers pass from a mono-layer structure mainly made of sands and gravels at north, i.e., the area called the higher plain, to a multi-layer structure at south, i.e., the lower plain, that is generated by the intercalation of silty-clayey into sandy layers (Giuliano, 1995; Ori, 1993; Perego et al., 2014). Silt and clay layers in the lower plain are frequently accompanied by buried peats (Amorosi et al., 2008; Miola et al., 2006). The transition from higher to lower plain is marked by numerous lowland springs, fed by groundwater outflows, that are called “springs belt” (Balestrini et al., 2021; Fumagalli et al., 2017; Fig. S1). The shallow aquifer system has a thickness >100 m at the foot of the Alps, passes to around 50 m in the higher plain and decreases to 20-30 m in the lower plain (Fig. S1). The shallow aquifers are unconfined in the higher plain and semiconfined/confined in the lower plain (Regione Lombardia, 2016). Shallow groundwater generally flows from north to south in the higher plain and from north-west to south-east in the lower plain, however here, groundwater flow directions are altered by gaining rivers (Regione Lombardia, 2016; Fig. S2). The intermediate aquifer system has a thickness of around 50-100 m in the higher plain that progressively increases in the lower plain (Fig. S1) reaching 600 m in the south-east of the region, however here, only the first 250 m below ground surface (bgs) are tapped by water wells (Rotiroti et al., 2014b). The intermediate aquifers are semiconfined/confined in the higher plain and confined in the lower plain (Regione Lombardia, 2016). The deep aquifer

system is tapped by water wells only in the north-west part of the region where the thicknesses of the overlying intermediate and shallow aquifers are lesser (Fig. S1). The deep aquifer system is confined and its bottom is never reached by water wells (Regione Lombardia, 2016). Within the intermediate and deep aquifer systems, groundwater generally flows from north-west to south-east (Regione Lombardia, 2016; Fig. S2). The whole alluvial aquifer system of the Po Plain receives recharge waters (from local precipitation, losing rivers and irrigation) at the foot of the Alps and in the higher plain, whereas recharge from the surface is absent or very limited in the lower plain due to the widespread presence of surficial clay and silt layers (Rotiroti et al., 2019). Therefore, lower plain aquifers are mainly recharged by groundwater inflows from the upgradient higher plain aquifers. This configuration leads the higher plain aquifers to have (a) shorter groundwater flowpaths, and thus, shorter groundwater residence times, in the order of some tens of years (Musacchio et al., 2018), (b) higher vulnerability to anthropogenic activities (Azzellino et al., 2019; Stevenazzi et al., 2015) and (c) oxic groundwaters (Sacchi et al., 2013), where no contaminations by organic compounds occur. Conversely, the lower plain aquifers have longer groundwater flowpaths and residence times, up to 50k years (Martinelli et al., 2014; Zuppi and Sacchi, 2004). This promotes reducing conditions fueled by the degradation of the natural organic matter (OM) of the peat layers and the mobilization of its byproducts, such as NH_4 (Böhlke et al., 2006), and the byproducts of associated terminal electron accepting processes (TEAPs), such as Mn and Fe (Rotiroti et al., 2021). In this geochemical setting, As is mainly released to groundwater by the reductive dissolution mechanism (Ravenscroft et al., 2009; Rotiroti et al., 2021). Therefore, As, NH_4, Fe and Mn in lower plain groundwaters have high concentrations of natural origin (exceeding the regulatory limits of 10, 500, 200 and 50 µg/L, respectively; D. Lgs. 152/06, 2006; D. Lgs. 30/09, 2009). Anyway, some anthropogenic influences on their concentrations, generated by the degradation of anthropogenic organic compounds (e.g., organic leachate, hydrocarbons, etc.), were reported in some lower plain areas (Rotiroti et al., 2014a). Another important geochemical process operating in the lower plain and influencing its groundwater quality is the upwelling of the deep brines of the Po Plain (Conti et al., 2000) facilitated by some fault systems. Evidence of the operation of this process is reported for the south-west area of the region, the so called “Oltrepò Pavese” (Pilla et al., 2015, 2010), and for the south-east, the so called “Oltrepò Mantovano” (Bonori et al., 2000), in correspondence with, respectively, the fault systems named Emilia and Ferrara arcs (Castellarin et al., 2006; Michetti et al., 2012; Scardia et al., 2015).

The Alpine valley aquifers are alluvial, fluvioglacial and glacial aquifers that extend along and beneath main Alpine rivers, with a limited lateral extent (up to 1-2 km; Fig. S1). These aquifers are generally interconnected with the rivers and are unconfined, however, in some areas, they can have some local confinements generated by silt/clay lenses. The thickness of tapped aquifer by water wells is generally <100 m bgs. Groundwaters circulating in these aquifers are

mostly oxic, with very few exceptions where confining silt/clay lenses and buried OM promote reducing conditions, leading to high concentrations of reduced species (Fe, Mn and NH_4). Within the GWBs of the Valtellina area (GWB IDs IT03GWBFITE, IT03GWBFMTE and IT03GWBFSTE in Fig. S2) hosting oxic groundwaters, As concentrations were reported to exceed the limit of 10 µg/L (Regione Lombardia, 2016). The As release mechanisms operating here are likely the alkali desorption and pyrite oxidation (Ravenscroft et al., 2009), as speculated by Peña Reyes et al. (2015).

2.2. Available dataset

Hydrochemical data were made available from the archive managed by the Regional Agency for Environmental Protection of Lombardy (ARPA), referring to the regional monitoring surveys performed between 2014 and 2017. The regional monitoring network comprises 503 stations (wells and piezometers; Fig. S1), 451 of them tap groundwater from 27 out of 30 GWBs (3 Alpine GWBs have no monitoring stations; Fig. S1 and S2), whereas the remaining 52 monitoring stations tap local aquifers (Fig. S1), which are not defined as GWBs yet, and thus were then excluded from the NBLs calculation (see Sect. 2.3 for details). Lithologs for each station were available from ARPA and TANGRAM (Bonomi et al., 2014) databases. During the considered period (i.e., 2014-2017), 8 regional monitoring surveys were acquired from ARPA (one survey in 2014, two surveys in 2016 and 2017 and three surveys in 2015), leading to a total of 3383 samples for As, 3473 for Fe and Mn, and 3283 for NH_4. In each sample, the analysis of target species (As, NH_4, Fe and Mn; no metal speciation analyses were performed) was accompanied by the measurement/analysis of pH, water temperature (Temp), electrical conductivity (EC), dissolved O_2 (DO), hardness as $CaCO_3$, alkalinity as HCO_3, Mg, Ca, Cl, K, Na, SO_4, NO_3 and PO_4. Additional anthropogenic compounds (hydrocarbons, chlorinated solvents, pesticides, herbicides, etc.) were measured. The pH, temperature, EC and DO were measured in the field by ARPA. Samples for As, Fe and Mn analysis were treated in the field by 0.45 µm filtration and acidification. The chemical analyses were performed in the ARPA laboratories, which comply to ISO/IEC 17025, using standard procedures (ARPA Lombardia, 2021). Concerning the target species, the average analytical precision was ±2 µg/L for As, ±138 µg/L for Fe, ±32 µg/L for Mn and ±146 µg/L for NH_4. Data quality was evaluated verifying a charge balance error (CBE) for each sample below the threshold of 10%, according to European (Muller et al., 2006) and Italian (ISPRA, 2017) guidelines for estimating NBLs. Accordingly, samples with incomplete major ions analyses or having a CBE >|10|% were discarded. This operation led to retain a total of 2396 samples for As, 2506 for NH_4 and

2508 for Fe and Mn. A summary of statistics for these samples is reported in Table S1. Censored data, i.e., concentrations below the limit of detection (LOD), were replaced with LOD/2. Although the percentage of censored data in the total dataset (58, 67, 58 and 55% for As, NH_4, Fe and Mn, respectively; Table S1) exceeds the generic threshold of 15% recommended for the application of a simple substitution method (US EPA, 2006), it decreases (41, 44, 28 and 18%, respectively) considering only the monitoring stations located in the lower plain (i.e., the area affected by naturally high concentrations of the target species; see Sect. 2.1) below the threshold of 50% recommended for the substitution method in the case, as the present case, of large dataset size (>20 samples) with highly skewed distribution (geometric standard deviation >3) and more than one LOD (Hewett and Ganser, 2007; Hornung and Reed, 1990). The adoption of different LODs among the eight monitoring surveys must be properly managed in the LOD/2 substitution to avoid the generation of trends at time and space scales that would be the result of a fictitious variability in concentrations. The strategy adopted here was to replace the <LOD concentrations with the half of the minimum LOD used for each parameter of interest ($LOD_{min}/2$). In this way, concentrations <LOD were replaced with the value of 0.5 μg/L for As, 5 μg/L for NH_4, 2.5 μg/L for Fe and 0.5 μg/L for Mn.

2.3. Methodology for deriving local NBLs

The methodology used in this study to calculate local NBLs was consistent with recommendations of national guidelines (ISPRA, 2017; Parrone et al., 2021) and involved a tiered approach constituted by the following main steps (Fig. 1): 1) preliminary hydrochemical characterization 2) preselection, 3) redox zonation, 4) validation of hydrofacies identification, 5) outlier samples and temporal trends identification and treatment, 6) subpopulation identification by probability plot, 7) station outliers identification and treatment and 8) NBLs calculation. This approach can be viewed as a hybrid preselection-probability plot method. The preselection involves the use of indicator chemical species to discard samples with most likely anthropogenic influences and then the calculation of an upper limit (90^{th}, 95^{th} or 97^{th} percentile) of the remaining dataset as NBL. The probability plot method is based on the identification of one or more inflection points on a probability graph, which separate different subpopulations representing background and human-affected concentrations; the NBL is attributed as the concentration corresponding to one of the inflection points identified. The hybridization considered here is realized by, firstly, the use of the preselection (together with outliers and temporal trends identification and treatment) to extract a human impacts free dataset, secondly, the use of the probability plot to identify different subpopulations, which, at this point, are likely the expression of different natural

processes operating within a GWB and, thirdly, the calculation of an upper limit for each subpopulation identified (Fig. 1). In this way, local NBLs are assessed.

2.3.1. Preliminary hydrochemical characterization

The preliminary hydrochemical characterization consisted in a first elaboration of the whole dataset (concentrations of target species and major ions) in order to improve the preliminary conceptual model (Fig. 1), addressing a) main hydrogeochemical features, b) spatial patterns and c) likely anthropogenic influences. The preliminary hydrochemical characterization was done using multivariate statistical analysis, more specifically, combining a factor analysis (FA) and a hierarchical cluster analysis (CA), i.e., the CA was applied to the extracted factors from an initial FA (Azzellino et al., 2019; Liu et al., 2019; Masiol et al., 2010; Wang et al., 2017). The advantage of a FA-CA combination is the dimensional (i.e., no. of variables) reduction of the CA input matrix. The FA was based on the calculation of the correlation matrix and was made on standardized data (i.e., mean = 0 and standard deviation = 1; Judd, 1980), then a Varimax rotation was applied (Kaiser, 1958). The selection of the significant factors was done on the basis of the eigenvalues matrix: only those factors with eigenvalues ≥1 were considered as significant factors (Kaiser, 1958). The FA was applied on an input matrix made of the 18 original variables (pH, Temp, EC, DO, hardness, alkalinity, Mg, Ca, Cl, K, Na, SO_4, NO_3, PO_4, NH_4, As, Fe, Mn) and 503 samples (all monitoring stations). Matrix values were calculated as mean values, for each variable, of the time series of each monitoring station (monitoring station time series averaging). Calculation of mean values was done to prevent missing data, which must be avoided in the FA input matrix. The mean, instead of the median, was preferred since it is more sensitive to extreme values, which can represent contaminations (an important information to be registered in the preliminary characterization). The CA was done by means of the Ward method (Ward, 1963) using the Euclidean distance (Cloutier et al., 2008). Similarly to the FA, a standardization of data was applied in the CA to guarantee an equal weight for each variable in calculating the Euclidean distance matrix. The CA was applied on an input matrix made of 5 variables (the 5 significant factors obtained from the FA; see Sect. 3.1) and 503 samples.

2.3.2. Preselection

The preselection was applied to discard from the working dataset the samples likely affected by anthropogenic influences on the target species, identified by the exceedance of threshold concentrations for selected indicator species

of anthropogenic impact (Hinsby et al., 2008; Muller et al., 2006). The following indicator species and threshold concentrations were used: a) NO_3 >37.5 mg/L, b) total hydrocarbons (THC, computed as an n-hexane equivalent) >50 µg/L, c) sum of tetrachloroethene and trichloroethylene (PCE+TCE) >7.5 µg/L, d) Cl >45 mg/L and e) K >10 mg/L. These indicator species were analyzed in all samples, except for THC which was analyzed only in stations located where possible hydrocarbons spills can likely occur (e.g., urban and industrial areas). The lower number of analysis for THC is not considered to be able to bias the preselection results, since the probability to have detectable hydrocarbons in stations where THC was not measured (areas far from potential source of hydrocarbon spills) is very low. Nitrate was used as indicator of generic anthropogenic impacts, as recommended by national and European guidelines (ISPRA, 2017; Muller et al., 2006), however, since it is non-conservative under reducing conditions (due to denitrification), it played an insignificant role in the identification of anthropogenic alterations in the highly reducing groundwaters affected by the target species (As, NH_4, Fe and Mn). Obviously, the reduced aqueous species of nitrogen (i.e., NH_4) cannot be used in the present work as indicator since it is one of the target species of the NBL calculation. The threshold concentration of NO_3 (37.5 mg/L) was calculated as the 75% of the regulatory limit (50 mg/L), as recommended by national guidelines (ISPRA, 2017). The THC was used as specific indicator of anthropogenic influences on As, Fe and Mn concentrations, since the degradation of anthropogenic organic compounds can stimulate the reductive dissolution of Mn and Fe oxyhydroxides, releasing Mn, Fe and As to groundwater (Baedecker et al., 1993; Burgess and Pinto, 2005). The threshold of 50 µg/L corresponds to the minimum detected concentration from the ARPA regional monitoring. Similarly to THC, the PCE+TCE was used as specific indicator of anthropogenic organic contamination. Although the reductive dechlorination process implies the use of PCE and TCE before Mn and Fe oxyhydroxides within the ecological succession of TEAPs (McMahon and Chapelle, 2008), thus inhibiting the release of Mn and Fe to groundwater, its final product (ethene) can be then oxidated boosting the reductive dissolution of Mn and Fe oxyhydroxides. Accordingly, high Mn and Fe concentrations were reported in PCE/TCE contaminated sites (e.g., Palau et al., 2014). The threshold concentration (7.5 µg/L) was calculated as the 75% of the national regulatory limit (10 µg/L). Chloride was used as semi-specific indicator of anthropogenic influences on the target species, since it can trace generic anthropogenic impacts on groundwaters, but it is also a proxy of specific organic contaminations originated from sewage pipes, septic tanks, animal manure spreading and municipal solid waste landfill (MSWL). The use of Cl as indicator species was excluded for the samples with the brackish natural hydrofacies (see Sect. 3.2). The threshold concentration of Cl (45 mg/L) was calculated from measured Cl data (excluding the brackish natural hydrofacies), applying the interquartile range (75^{th} percentile + 1.5 × (75^{th} percentile − 25^{th} percentile); result of 45.3 mg/L) and the mean+2σ (mean + 2 × standard deviation; result of 45.6 mg/L) methods (Parrone et al., 2019). Potassium was used as

specific indicator of MSWL leachate contamination due to its release from vegetal wastes (Naveen et al., 2017; Stefania et al., 2019). The threshold concentration of 10 mg/L was chosen according to the previous work by Rotiroti et al. (2015) on the calculation of groundwaters NBLs in the Aosta Valley region of Italy. As done for Cl, the use of K as indicator species was excluded for the samples with the brackish natural hydrofacies.

2.3.3. Redox zonation and validation of hydrofacies identification

Performing a redox zonation after the exclusion of likely human-impacted samples (Fig. 1) means the segregation of natural redox states. A zonation of natural redox conditions is extremely important in the NBL calculation for redox-sensitive species (reduced species, in the present study), since it allows to separate the dataset into likely affected (reduced) and unaffected (oxidized) samples, in terms of naturally high concentrations of target species (ISPRA, 2017; Muller et al., 2006). The redox zonation was performed by means of a CA, using the same methodology reported in Sect. 2.3.1. This CA considered the following 8 variables (redox-sensitive species) in the input matrix: DO, NO_3, Mn, Fe, SO_4, NH_4, As, PO_4. A multivariate approach, with respect to the typical approach based on algorithms combining threshold concentrations of various redox-sensitive species (e.g., Lions et al., 2021; McMahon and Chapelle, 2008; Mendizabal and Stuyfzand, 2011), has the following main advantage: the definition of a priori absolute threshold concentrations is not needed, but instead, the segregation of the different redox zones is based on relative variations of the redox-sensitive species within the dataset. This could be particularly useful when the general quality of available data is not so high, e.g., due to different sampling or analytical techniques applied over time and space. However, the main disadvantage is the need to obtain necessarily clusters characterized by only redox processes. Another disadvantage is that a single cluster could represent more than one TEAP. Additionally, some limitations can emerge in the case of high concentrations of some redox-sensitive species, such as SO_4, NH_4, PO_4, originated from human influences, rather than linked to the operation of redox processes (e.g., high NH_4 indicates prolonged organic matter oxidation). However, these limitations can be overcome by a validation of obtained results (Fig. 1), as discussed below. The CA of redox sensitive species was applied on 417 samples, that represent the monitoring stations (initially 503) resulted from the preselection (437; see Sect. 3.2) from which it was subtracted the 11 stations with natural brackish hydrofacies (see Sect. 3.2), not considered in the redox zonation, and other 9 stations which resulted with no samples for PO_4 (8 stations) and As (1 station) after the preselection (the CA input matrix cannot have missing data).

The validation of hydrofacies identification consisted in checking the compliance of obtained redox hydrofacies (oxidized or reduced facies) from the redox zonation with a) the actual chemical composition of monitoring

292 stations and b) the conceptual model. The verification of the compliance of redox hydrofacies with the actual chemical
293 composition of stations was needed in those (few) cases in which the belonging to a specific CA cluster was unrelated
294 to main redox processes operating in the aquifers. Some practical cases are discussed in Sect. 3.2. The consideration of
295 the conceptual model, in terms of aquifer types or presence of local-scale processes, such as groundwater/surface water
296 interactions, or additional anthropogenic influences undetected by the preselection helped in this verification (Fig. 1). In
297 broad terms, results from a mere data-driven technique (CA) were validated with the soft knowledge (conceptual
298 model). This was needed to overcome the limitations related to the application of a CA for performing a redox zonation,
299 as discussed above.
300 The attribution of the natural hydrofacies was done for all the 503 monitoring stations, including also the
301 stations discarded from the redox zonation (see above), to support the subsequent assessment of the chemical status of
302 the GWBs (not presented in this work), according to European regulations (EC, 2006, 2000). The attribution of a
303 reduced or oxidated facies to the discarded stations was done on the basis of the conceptual model evaluating station
304 lithologs, type of anthropogenic contamination affecting the stations and station groundwater compositions; for
305 example, natural oxidized facies were attributed to stations located in the higher plain with highly-permeable and
306 unconfined aquifers but affected by organic compound contaminations which led to anoxic groundwaters.
307
308 ***2.3.4. Outlier samples and temporal trends identification and treatment***
309
310 Differently to the previous steps which comprised all available parameters (target species and major ions), the
311 identification and treatment of outlier samples and temporal trends were applied singularly to each target species dataset
312 (Fig. 1). In general, this operation was done with the aim of eliminating samples and/or stations likely affected by
313 human interferences not detected by the preselection. The idea is that impulsive anthropogenic influences may generate
314 one or more outlier samples and/or a temporal trend within the time series of a monitoring station, whereas constantly
315 impacted monitoring stations may constitute an outlier with respect to the surrounding monitoring points (outlier
316 stations, treated in Sect. 2.3.5). The present step was performed on each monitoring station time series by means of the
317 a) outlier samples identification and treatment, b) temporal trend identification and treatment, c) data distribution
318 identification, d) upper limit calculation.
319 Outlier samples were identified through the interquartile range approach (Wang et al., 2018). The small sample
320 size for time series (8 surveys, Sect. 2.2) did not allow the application of statistical tests (e.g., Rosner, Dixon, Huber,

etc.). According to the interquartile range approach, upper (UO) and lower (LO) outliers were defined as follows (Tukey, 1972):

- UO > 75^{th} percentile + 1.5 × (75^{th} percentile − 25^{th} percentile),
- LO < 25^{th} percentile − 1.5 × (75^{th} percentile − 25^{th} percentile).

Once identified, a decision on retaining or discarding the outlier was made (outlier treatment). False outliers, considered as part of the natural background (i.e., hot spots), were retained whereas true outliers, such as data errors or samples likely affected by some inhomogeneity of sampling methods or human influences, were discarded. As a rule, outliers with target species concentrations below half of the national regulatory limit (i.e., <5, <250, <100 and <25 µg/L for As, NH_4, Fe and Mn, respectively), were considered as false outliers. The presence of human influences, not detected by the indicator species used in the preselection, was evaluated on the basis of the station information available from the conceptual model (i.e., known human impacts on the station), whereas likely inhomogeneities of sampling methods were assessed through the co-presence of outliers for other parameters (target species and major ions) in the same sample. After true outliers exclusion, the presence of a temporal trend within monitoring station time series was evaluated by the Mann-Kendall test (Kendall, 1955; Mann, 1945). The test is applicable to time series with sample size ≥3. However, considering our small sample size (≤8), which is below the recommended minimum of 10 to obtain a statistically acceptable trend identification with the Mann-Kendall test (Hu et al., 2020; Şen, 2017), inaccurate results may be obtained. To overcome this limitation, identified trends were checked against the conceptual model of the monitoring station (i.e., presence of known human impacts, co-variability of other target species or major ions, etc.). Where the temporal trend was likely related to anthropogenic influences, the whole monitoring station was discarded from the working dataset.

Afterwards, a data distribution identification through the Shapiro-Wilk test (Shapiro and Wilk, 1965) was performed for each monitoring station. The test is applicable to sample size ≥3. Results of this test were related to the upper limit calculation of each station time series: if the time series consisted of a normal population, the maximum value was chosen as upper limit, if a non-normal population was identified, the 95^{th} percentile was used as upper limit (Parrone et al., 2019). For populations with size <3, the maximum was used as upper limit. The choice of an upper limit as value representative of a monitoring station disagrees with European and national guidelines which recommend the use of the median (ISPRA, 2017; Muller et al., 2006). Anyway, we believe that the use of the median may lead to an underestimation of NBLs when the peak values of station time series are unequivocally related to natural processes, as it should be at this point of the methodology (i.e., after preselection and outlier samples treatment).

351 ***2.3.5. Subpopulation identification, outlier stations treatment and local NBLs calculation***

352

353 The subpopulation identification was applied on datasets made of the representative values of each monitoring
354 stations (Fig. 1), calculated as described in Sect. 2.3.4. Only stations tapping GWBs were considered (Sect. 2.2). These
355 datasets were built for each target species (As, NH_4, Fe and Mn), aquifer type (Po Plain and Alpine valley aquifers;
356 Sect. 2.1) and natural hydrofacies (oxidized, reduced and brackish, the last is present only in the Po Plain aquifers; Sect.
357 3.2). Therefore, for each target species, five input datasets to the subpopulation identification were considered: 1) Po
358 Plain oxidized, 2) Po Plain reduced, 3) Po Plain brackish, 4) Alpine valley oxidized and 5) Alpine valley reduced
359 groundwaters. The subpopulation identification was made through probability plots, identifying inflection points which
360 separate different subpopulations (Preziosi et al., 2014; Sinclair, 1974). At this point of the methodology (i.e., after
361 preselection and outlier samples treatment; Fig. 1), these subpopulations reflect natural background. Different
362 subpopulations can be considered as the expression of different natural processes operating in the aquifer to which
363 different NBLs can be assessed. Considering that each point of these datasets represents a monitoring station (Fig. 1),
364 identifying different data subpopulations means to detect a spatial variability of the NBLs, complying the concept of
365 local NBL (Dalla Libera et al., 2017; Ducci et al., 2016; Molinari et al., 2019). The use of the probability plot method
366 introduces a degree of subjectivity into the whole methodology for deriving the NBLs, since the choice of the inflection
367 points is dependent on "the eyes of the investigator" (Kim et al. 2015; Preziosi et al., 2014). Anyway, probability plots
368 are recommended when dealing with multimodal distributions (Kim et al. 2015; Nakic et al. 2020). Supporting the
369 choice of the inflection points by the soft knowledge (conceptual model), the uncertainty on resulted NBLs introduced
370 by the subjective interpretation of probability plots can be constrained (Preziosi et al., 2014).

371 The probability plot can also identify some outliers, which here represent outlier stations. When identified,
372 these outliers were evaluated through the conceptual model for their exclusion (true outliers) or retention (false
373 outliers). This represented the last check for discarding likely human-affected data from the working dataset (Fig. 1).
374 When the outlier station was generated by a single sample outlier in the time series, only this sample was discarded to
375 avoid the exclusion of the entire monitoring station. Once true station outliers were discarded, the NBLs were calculated
376 for each subpopulation identified using the following criteria (Parrone et al., 2019):

377 - NBL = maximum of the subpopulation, when a normal distribution was identified (by the Shapiro-Wilk test),
378 - NBL = 95th percentile of the subpopulation, when a non-normal distribution was identified.

379 When the subpopulation size was <15, the NBL was calculated a) as the 90th percentile of the dataset formed by the
380 aggregation of the single samples of all the monitoring stations constituting the subpopulation, or alternatively, b)

joining the subpopulation with another one with larger size (>15) and representing similar hydrogeochemical processes (ISPRA, 2017).

3. Results

3.1. Improved conceptual model

The FA of the preliminary hydrochemical characterization identified 5 significant factors (FAC1-5) explaining a total cumulative variance of 74.2% (Table S2). FAC1 explained 27.7% of the total variance. The original variables representing FAC1 (i.e., loading value >|0.6|) were hardness, Ca, alkalinity, EC, Mg and SO_4, thus FAC1 expresses major ions chemistry, which determines groundwater salinity (intended as the total amount of ions dissolved in groundwater). FAC2 explained 14.2% of the variance and was represented by Na and Cl, identifying brackish groundwaters. FAC3 explained 12.7% and was represented by As, PO_4 and NH_4. Since these three species can dominate after prolonged OM degradation (As is mainly released by the reductive dissolution mechanism; Rotiroti et al., 2021), FAC3 identifies groundwaters with highly reducing conditions. FAC4 explained 11.7% and was represented by Mn and Fe, thus identifying groundwaters with Mn- and Fe-reducing conditions. FAC5 explained 7.9% and was represented by Temp and DO, negatively correlated with each other, thus FAC5 identifies oxic groundwaters (please note the negative loading value for DO implying more oxic conditions with lower FAC5 scores).

The combined CA showed that monitoring stations can be grouped into eight clusters (C1-8). The histogram of centroids for the eight clusters is shown in Fig. S3; average values of the 18 original variables for each cluster are shows in Table S3. Cluster C1 (136 monitoring stations) was mainly characterized by negative Z-scores of FAC5 and FAC1, therefore stations forming C1 tap groundwaters with oxic conditions and low salinity (average EC and DO of 386 µS/cm and 8.1 mg/L; Table S3). Cluster C2 (148 stations) was characterized by negative Z-scores of FAC5 and positive Z-scores of FAC1, thus it groups stations tapping groundwaters with oxic conditions and higher salinity, generated by stronger anthropogenic influences (average EC, DO and NO_3 of 616 µS/cm, 7.7 and 33.5 mg/L, respectively; Table S3). Cluster C3 (124 stations) is characterized by positive Z-scores of FAC5 and FAC4 (although low) and negative Z-scores of FAC1, thus it groups stations tapping groundwaters with Mn-/Fe-reducing conditions and low salinity (average EC, Mn and Fe of 376 µS/cm, 63 and 178 µg/L, respectively; Table S3). Cluster C4 (31 stations) is characterized by positive and high Z-scores of FAC3, thus it groups stations tapping groundwaters with highly reducing conditions (average As and NH_4 of 41 and 2034 µg/L; Table S3). Cluster C5 (13 stations) is characterized by positive

411 Z-scores of FAC2, thus it groups stations tapping brackish groundwaters (average EC and Cl of 1000 µS/cm and 114
412 mg/L; Table S3). Cluster C6 (41 stations) is characterized by positive Z-scores of FAC5, FAC4 and FAC1, thus it
413 groups stations tapping groundwaters with Mn-/Fe-reducing conditions and higher salinity, likely generated by
414 anthropogenic influences (average Mn, Fe and Cl of 251, 876 µg/L and 23 mg/L, respectively; Table S3). Cluster C7
415 groups only 8 stations and it is similar to C6 but having higher Mn and Fe (average of 750 and 4265 µg/L, respectively;
416 Table S3). Cluster C8 groups only 2 stations and it is similar to C5 (brackish water) but having higher EC, Cl and, in
417 addition, NH_4 (average of 1889 µS/cm, 401 mg/L and 3539 µg/L, respectively; Table S3). In summary, three main
418 hydrofacies can be identified: a) oxidized (C1 and C2), b) reduced (C3-4 and C6-7) and c) brackish (C5 and C8)
419 groundwaters.
420 Analyzing the spatial distribution of these clusters (Fig. S4) and integrating the information available from the
421 preliminary conceptual model, an improved conceptual model, presented below, can be drawn (Fig. 1). The higher plain
422 aquifers are generally oxic (C1 and C2). Less (or non) human-impacted groundwaters (C1) are limited to the north and
423 north-west areas in the shallow aquifers, are prevalent in the intermediate aquifers and dominate the deep aquifers. By
424 contrast, human-impacted groundwaters (C2), mainly affected by nitrate pollution (Martinelli et al., 2018), dominates
425 the shallow aquifers, are less present in the intermediate aquifers and are sporadic in the deep aquifers. A few stations
426 with reduced groundwaters (C3-4 and C6) were found in the higher plain in the shallow aquifers. This can be likely
427 related to the presence of anthropogenic impacts rather than natural conditions (as confirmed by the preselection, see
428 Sect. 3.2). The lower plain aquifers are generally anoxic (C3-4 and C6-7) with a trend from Mn-/Fe-reducing conditions
429 (C3 and C6-7) at north-west to highly reducing conditions (C4) at south-east. Human-impacted groundwaters (C6-7) are
430 conspicuous in the shallow, rare in the intermediate and absent in the deep aquifers. A few stations with oxic
431 groundwaters (C2) were found in the lower plain in the shallow aquifers. This can be related to the presence of surficial
432 windows of coarser deposits, as revealed by the analysis of station lithologs, allowing, locally, some recharge of oxic
433 water from the surface. Stations with brackish groundwaters (C5 and C8) are mainly located in the southern area with
434 only a few exceptions spread over the entire region. The 11 southern stations near the fault systems called Emilia and
435 Ferrara arcs (Sect. 2.1, Fig. S5) were considered to tap natural brackish groundwater (see Sect. 2.1) whereas the 4
436 northern stations far from these tectonic structures were considered affected by human impacts, generating
437 anthropogenic brackish groundwater. In two stations (IDs PO098012NR0011 and PO017029NR0001; Fig. S5), the
438 human impact was confirmed by the exceedance of regulatory limit for Cr(VI) and many anthropogenic compounds
439 (PCE+TCE, trichloromethane, atrazine and bentazon).
440

3.2. Natural redox zones and hydrofacies

The preselection led to discard a total number of 520 samples, more specifically, 318 samples were discarded for exceeding the threshold of NO_3, 11 for THC, 167 for PCE+TCE, 98 for Cl and 7 for K (the sum of these samples is greater than the total number of excluded samples since some samples had multiple exceedances). Locations and number of discarded samples (total and partial for each indicator species) for each monitoring stations are shown in Fig. S6a-f. The monitoring stations with all samples discarded (discarded stations; Fig. S6g) were 66. Most of the discarded samples were in the higher plain. This agrees with the conceptual model which underlined the greater vulnerability to anthropogenic impacts of the higher plain aquifers due to their higher permeability and recharge rates. The number of samples resulted from the preselection (preselected samples) for each target species were 2655 for As, 2604 for NH_4 and 2739 for Fe and Mn (Table S1).

The CA of the redox zonation identified six clusters of monitoring stations ($C1_{redox}$-6_{redox}). The histogram of centroids for the six clusters is shown in Fig. S7 whereas average values of the 8 original variables for each cluster are shows in Table S4. Cluster $C1_{redox}$ (221 monitoring stations) was characterized by positive Z-scores of DO and NO_3, thus it groups stations tapping oxic groundwaters (average DO and NO_3 of 8.0 and 18.4 mg/L; Table S4). Cluster $C2_{redox}$ (32 monitoring stations) was characterized by positive Z-scores of NO_3, Mn, Fe and SO_4; it can be considered to group stations tapping groundwaters ranging between denitrification and Mn-/Fe-reducing conditions (average NO_3, Mn and Fe of 14.1 mg/L, 140 and 677 µg/L; Table S4). The increase of average SO_4 from $C1_{redox}$ (26 mg/L) to $C2_{redox}$ (74 mg/L) can be due to the operation of denitrification coupled to iron sulfides oxidation (Schwientek et al., 2008; Vaclavkova et al., 2014; Zhang et al., 2009). Cluster $C3_{redox}$ (9 stations) is characterized by positive and high Z-scores of Mn and Fe, thus it can be considered to group stations tapping groundwaters with strong Mn-/Fe-reducing conditions (average Mn and Fe of 817 and 3256 µg/L; Table S4). Cluster $C4_{redox}$ (122 stations) is characterized by negative Z-scores of SO_4; it can be considered to group stations tapping groundwaters dominated by SO_4-reducing conditions due to the following reasons: a) average SO_4 and Fe in $C4_{redox}$ (17 mg/L and 230 µg/L) are lower than $C2_{redox}$ (74 mg/L and 677 µg/L), marking the reduction of sulfate with the precipitation of iron sulfides (Rotiroti et al., 2021); b) average NH_4 in $C4_{redox}$ (294 µg/L) is higher than $C1_{redox}$ and $C2_{redox}$ (19 and 128 µg/L, respectively) marking the progression of OM degradation and, accordingly, the hierarchical succession of TEAPs (i.e., from Mn-/Fe-reducing to SO_4-reducing conditions); c) average As is still relatively low (5.7 µg/L), likely due to its coprecipitation in iron sulfides (Rotiroti et al., 2021). Cluster $C5_{redox}$ (31 stations) is characterized by positive Z-scores of NH_4, As, PO_4 and Fe; it can be considered to group stations tapping groundwaters dominated by methanogenesis since, under this process, a co-occurrence of these four

471 species (average NH_4, As, PO_4 and Fe of 2002, 22.7, 780 µg/L and 1.23 mg/L; Table S4) is expected (Rotiroti et al.,
472 2021). Cluster $C6_{redox}$ groups only 2 stations and it is similar to $C5_{redox}$ but having higher NH_4, As, PO_4 and Fe (average
473 of 3392, 236, 1374 µg/L and 2.53 mg/L, respectively; Table S4). The spatial distribution of the clusters is shown in Fig.
474 2. According to the conceptual model, oxic waters ($C1_{redox}$,) dominated the higher plain aquifer whereas a progression to
475 lower redox states (from $C2_{redox}$ to $C6_{redox}$) can be seen in the lower plain groundwaters moving towards south-east and
476 over depth, corresponding to increasing groundwater ages (Sect. 2.1). In summary, the CA attributed an oxidized
477 natural hydrofacies ($C1_{redox}$) to 221 stations and a reduced natural hydrofacies ($C2\text{-}6_{redox}$) to 196 stations.
478 The validation of hydrofacies identification led to change the hydrofacies attributed from the CA for 27 out of
479 417 stations. In broader terms, this can be viewed as the error estimation of the redox zonation, that corresponds to the
480 acceptable value of 6.5%. Stations which experienced the hydrofacies change are shown in Fig. 2. The change was from
481 oxidized to reduced hydrofacies for 10 stations and from reduced to oxidized for 17 stations. In most cases, the change
482 from reduced to oxidized facies was needed when the falling in $C2_{redox}$ (characterized by higher SO_4) for a station with
483 an actual oxidized groundwater composition (i.e., high DO, low Fe and Mn) was due to a relatively high SO_4 of likely
484 anthropogenic origin. On the other hand, most changes from oxidized to reduced facies were done when the falling in
485 $C1_{redox}$ for a station with an actual reduced groundwater composition (i.e., high Fe and Mn) was due to an inaccurate
486 DO measurement (e.g., without using a flow cell under low flow conditions) resulting in relatively high DO.
487 After validation and attribution of natural hydrofacies to discarded stations (see Sect. 2.3.3 for details), the
488 final hydrofacies attribution resulted in 293 stations with a natural oxidized hydrofacies, 199 stations with a natural
489 reduced hydrofacies and 11 stations with a natural brackish hydrofacies. The spatial distribution of natural hydrofacies
490 for stations tapping GWBs is shown in Fig. S8.
491
492 **3.3. Subpopulations and local NBLs**
493
494 Results of sample outlier and temporal trend identification are reported in Table S1. A few sample outliers (7
495 out of 60 for As, 22 out of 79 for NH_4, 13 out of 75 for Fe and 13 out of 71 for Mn) were considered as true outliers
496 (errors or likely affected by human influences; according to criteria discussed in Sect. 2.3.4) and thus were discarded
497 from the working dataset. A few stations were identified to have a temporal trend (4 for As and NH_4, 3 for Fe and 13 for
498 Mn) and a small portion of these (1 for As and Fe, 2 for NH_4 and 3 for Mn) was considered as affected by human
499 influences and discarded. Results of data distribution identification are reported in Table S1; it is noted that many

stations (207 out of 436 for As, 228 out of 435 for NH_4, 198 out of 436 for Fe and 203 out of 434 for Mn) had an unidentifiable distribution due to a sample size <3.

Probability plots, after station outlier treatment (Table S1), for stations tapping Po Plain reduced groundwaters are shown in Fig. 3. The plots for Po Plain oxidized and brackish and Alpine valley oxidized groundwaters are reported in Fig. S9-11; no plots were done for Alpine valley reduced groundwaters due to the small sample size (3; Table S1). Fig. 3 shows that three inflection points, identifying four subpopulations (Red1-4), can be observed for As, NH_4 and Fe, whereas four inflection points (5 subpopulations; Red1-5) were observed for Mn. Table S5 shows the results of subpopulation identification for all the datasets. Two subpopulations were identified for Po Plain oxidized (Ox1-2) and brackish (Brck1-2) groundwaters, except for the former for As and NH_4, for which three subpopulations were identified (Ox1-3). Concerning the Alpine valley aquifers, a unique population for each dataset was identified, except for As for oxidized groundwaters (2 subpopulations; Ox1-2) and Fe for reduced groundwaters (2 subpopulations; Red1-2). The subpopulation Ox2 for As in the Alpine valley aquifers is related to As mobilization mechanisms operating under oxidizing conditions, as previously reported by Peña Reyes et al. (2015). Table S5 reports also results of data distribution identification through the Shapiro-Wilk test and criteria used for calculating the NBL, depending on distribution type and subpopulation size (Sect. 2.3.5). Finally, Table S5 shows the resulted NBLs for each subpopulation. Spatial distributions of NBLs for As, NH_4, Fe and Mn are shown in Fig. 4, 5, 6 and 7, respectively.

4. Discussion

4.1. Linking NBLs to redox evolution in Po Plain groundwaters

The evolution of redox processes in lower Po Plain groundwaters along regional flowpaths can be inferred by combining results of redox zonation CA (Fig. 2) and regional groundwater flow (Fig. S2). In the shallow aquifers, denitrification/Mn-/Fe-reducing conditions ($C2_{redox}$) dominate in upstream areas (i.e., just downstream of the springs belt and the piedmont of the Apennines) and pass to sulfate reducing ($C4_{redox}$) along groundwater flow (Fig 2a); methanogenesis ($C5_{redox}$) is found largely in the south-east area which is the terminal part of regional groundwater flowpaths (Fig 2a and S2). Strong Mn-/Fe-reducing conditions ($C3_{redox}$) are mostly found in the shallow aquifers in the south-east and south-west areas, corresponding to Holocene river valleys, mostly the Po River valley (Fig 2a). This may indicate that strong Mn and Fe reduction can operate where younger, less stable and more reactive oxides are present. These oxides may be originated from the weathering of serpentinites and peridotites coming from the Apennines and/or

transported by the Po River (Pettine et al., 1994; Sacchi et al., 2020). In the underlying intermediate aquifers, sulfate reducing conditions ($C4_{redox}$) dominate in upstream areas and pass to methanogenesis ($C5_{redox}$) in the south-east area (Fig 2b). Interestingly, strong methanogenesis ($C6_{redox}$) is not found in the far south-east, at the terminal part of the regional groundwater flowpath, but, instead, it is found in the middle of it (Fig. 2b), probably due to a higher content of peat in this area.

The link between redox zonation and obtained NBLs is shown in Fig. 8a which plots the average NBL corresponding to each cluster of the redox zonation. Since the transition from $C1_{redox}$ to $C6_{redox}$ represents the succession from higher to lower redox states, as discussed in Sect. 3.2, Fig. 8a marks the variation of NBLs along the evolution of redox processes in Po Plain groundwaters. Average NBLs of NH_4 progressively increase from 52 µg/L in $C1_{redox}$, representing oxic conditions, to 3785 µg/L in $C6_{redox}$, representing strong methanogenesis. The increase agrees with the accumulation of by-products of OM (in this case, peat) degradation (e.g., NH_4) over decreasing redox states (Böhlke et al., 2006; Rotiroti et al., 2021). This interpretation helps also to decipher the spatial distribution of NH_4 NBLs (Fig. 5), which increase towards the south-east according to the spatial distribution of redox states discussed above. It is noted that the highest NH_4 NBL (6620 µg/L) is not related to a reduced hydrofacies but, instead, to a brackish hydrofacies. The maximum NH_4 NBL found in the brackish groundwater is attributed to a high NH_4 content featuring the upwelling fluids from deep (>1 km of depth) Po Basin sediments, hosting petroleum systems (Lindquist, 1999), as NH_4 is excreted during hydrocarbon formation and maturation (Williams et al., 1992). Average NBLs of Fe (Fig. 8a) have a peak (4190 µg/L) with strong Mn-/Fe- reducing conditions, then decrease (872 µg/L) during sulfate reduction and finally increase again (up to 1927 µg/L) during methanogenesis. This profile agrees with the evolution of groundwater Fe over decreasing redox states (Appelo and Postma, 2005; Rotiroti et al., 2021), more specifically, the decrease during sulfate reduction is attributed to the precipitation of iron sulfides whereas the increase with methanogenesis is related to the likely co-operation of Fe-oxides reduction and methanogenesis (Rotiroti et al., 2021). The spatial distribution of Fe NBLs (Fig. 6) reflects these hydrogeochemical features, leading to have the highest Fe NBLs (3200 and 6000 µg/L) in the shallow aquifers in the south-east and south-west areas, where Fe-oxide reduction is supposed to be stronger due to the likely presence of younger and less-stable oxides, as discussed above. Average NBLs of Mn (Fig. 8a) have a peak (1009 µg/L) with strong Mn-/Fe- reducing conditions, decrease (205 µg/L) during sulfate reduction and then remain quite stable (216 and 238 µg/L) during methanogenesis. This profile is consistent with the conceptual model developed for groundwater Mn in the Po Plain aquifer systems (Rotiroti et al., 2021): the peak of groundwater Mn generated by Mn-oxides reduction is then attenuated by the precipitation of rhodochrosite ($MnCO_3$) which then reaches an equilibrium condition for precipitation/dissolution, leading to quite constant concentrations in groundwater.

Accordingly, the spatial distribution of Mn NBLs (Fig. 7) reflects that of Fe, with the highest NBLs (912 and 1514 µg/L) reached in the shallow aquifers in the south-east and south-west areas. Average NBLs of As (Fig. 8a) have a first (smaller) peak (21 µg/L) with strong Mn-/Fe-reducing conditions, then decrease (16 µg/L) due to sulfate reduction and finally increase obtaining a second (higher) peak (up to 291 µg/L) during methanogenesis. This profile is consistent with the "two-peaks" conceptual model explaining the evolution of groundwater As during ongoing degradation of peat in the Po Plain aquifers (Rotiroti et al., 2021): the first peak is generated by As release from reductive dissolution of Fe-oxides after prolonged Fe-oxide reduction and it is diminished by sulfate reduction with co-precipitation in iron sulfides; the second peak occurs with concomitant Fe-reduction and methanogenesis, during which process As is released with no or little attenuation, so it can reach the highest concentrations. The spatial distribution of As NBLs (Fig. 4) reflects the redox zonation resulting in the highest As NBLs (71 and 291 µg/L) in the south-east areas where methanogenesis dominates.

4.2. Deciphering the subpopulations

The approach used in this work is based on the assumption that the different data subpopulations identified through the probability plot method represent different processes, to which different NBLs can be assessed. Below, we try to decipher which redox process (or processes) each subpopulation can represent. Fig. 8b shows, for Po Plain reduced groundwaters, a histogram representing, for each subpopulation, the number of monitoring stations (%) per cluster of the redox zonation (C1-6_{redox}) over the entire cluster size. Fig. 8b for As tells us that the subpopulation Red1 groups As concentrations mainly obtained under denitrification, Mn-/Fe-oxides reduction (both represented by $C2_{redox}$) or sulfate reduction ($C4_{redox}$). In other words, Red1 groups lower As concentrations which can be generated by, at least, three different processes: a) the inhibition of As release (via reductive dissolution) due to the fact that denitrification outcompetes Mn-/Fe-oxides reduction; b) the re-sorption onto residual oxides, during Mn-/Fe-oxides reduction (McArthur et al., 2004; Welch et al., 2000); c) the co-precipitation into iron sulfides, during sulfate reduction. Red2 represents As concentrations obtained with methanogenesis ($C5_{redox}$) or strong Mn-/Fe- reducing conditions ($C4_{redox}$), the latter can be related to the first As peak discussed in Sect. 4.1. Red3 groups As concentrations with methanogenesis ($C5_{redox}$) and Red4 represents those concentrations obtained under strong methanogenesis ($C5_{redox}$), the second As peak discussed in Sect. 4.1. Concerning NH_4, Red1 groups NH_4 concentrations obtained as the product of OM (peat) oxidation coupled to reduction reactions ranging from denitrification to sulfate reduction (C2-4_{redox}). Red2 collects NH_4 concentrations reached with methanogenesis ($C5_{redox}$), whereas Red3 and Red4 group those concentrations obtained

590 with strong methanogenesis ($C6_{redox}$). Concerning Fe, Red1 groups lower Fe concentrations mainly generated under
591 sulfate reduction (via iron sulfides precipitation; $C4_{redox}$), Red2 collects Fe concentrations obtained with Fe-oxides
592 reduction ($C2_{redox}$) or methanogenesis ($C5_{redox}$). Red3 groups Fe concentrations related to strong methanogenesis
593 ($C6_{redox}$) whereas Red4 collects those concentrations obtained under strong Fe-oxides reduction ($C3_{redox}$). Concerning
594 Mn, Red1 represents lower Mn concentrations obtained during denitrification, early stages of Mn-oxides reduction
595 (both represented by $C2_{redox}$) or precipitation of rhodochrosite that can take place under sulfate reduction ($C4_{redox}$),
596 although there is no causal relationship between them. Red2 collects Mn concentrations related to the
597 precipitation/dissolution equilibrium of rhodochrosite occurring under strong methanogenesis ($C6_{redox}$), methanogenesis
598 ($C5_{redox}$) or sulfate reduction ($C4_{redox}$). Red3 groups Mn concentrations generated by Mn-oxides reduction ($C2_{redox}$),
599 whereas Red4 and Red5 collect those concentrations related to strong Mn-oxides reduction ($C3_{redox}$).
600 In summary, it emerges that, with such complex hydrogeochemical reactions involved, a single concentration
601 subpopulation, identified through the probability plot method, can be the expression of different processes, all
602 contributing to a specific concentration range of the target species. In other cases, mostly for high concentrations of the
603 target species, a subpopulation can represent a single main process.
604
605 **4.3. To interpolate or not to interpolate**
606
607 The ideal final step of this work would be the spatial interpolation of obtained discrete (point) NBLs to produce a
608 continuous distribution of NBLs over the entire aquifer system, taking full advantage of the concept of local NBL
609 within a GWB. A continuous NBL distribution would be of great importance within the scope of contaminated site
610 management and remediation because a NBL can substitute the reference level established by law for the identification
611 of anthropogenic contaminations and/or remediation goals. Several spatial interpolation methods exist, falling into three
612 main categories (Li and Heap, 2014): non-geostatistical interpolators (e.g., spline, nearest neighbor, etc.), geostatistical
613 interpolators (e.g., ordinary kriging, cokriging, stratified kriging, etc.) and combined methods (e.g., regression kriging,
614 etc.). Some methods can assist the interpolation of target (primary) variable with auxiliary information, in the form of,
615 for example, secondary variables (e.g., cokriging method) or strata (i.e., stratified kriging). The choice of the most
616 suitable method depends on the assumption and properties of each method, sample size or sample density and
617 distribution of the primary variable, the availability of secondary information, etc. (Li and Heap, 2014). Beyond the
618 choice of any method, one should be firstly questioning on the meaningfulness of expected results of the interpolation,
619 on the basis of the characteristics of available data (e.g., sample quality, density and size of primary and secondary

620 variables) with respect to the nature and spatial structure of the processes driving the spatial variability of the target
621 variable. Many previous studies addressed the interpolation of NBL within a GWB using different methods. Ducci et al.
622 (2016) and Dalla Libera et al. (2018) used, respectively, indicator kriging and indicator cokriging to spatialize the
623 probability of exceeding the NBL of target species. Dalla Libera et al. (2017) used collocated cokriging to obtain a
624 piecewise distribution of As NBL. Molinari et al. (2019) used ordinary kriging with a stochastic approach, varying
625 variogram models, to obtain a continuous distribution of As and NH_4 NBLs. Anyway, in the preset work, we decided
626 not to perform an interpolation of our point NBLs, the motivation is discussed below. Essentially, we think that the
627 spatial density of our point NBLs is not sufficiently high to obtain meaningful results, particularly in the areas with high
628 hydrogeological heterogeneity (i.e., springs belt and surficial higher-permeability windows in the lower plain) that
629 produces steep or sharp spatial changes of redox states, and thus steep or sharp variations of our target redox-sensitive
630 species. An accurate prediction of target species in these areas would require a very high data density (Li and Heap,
631 2014), that is currently not available. The risk of obtaining inaccurate and meaningless results in the area surrounding
632 the springs belt, where NBLs steeply pass from below the current regulatory limits (10, 500, 200 and 50 µg/L for As,
633 NH_4, Fe and Mn, respectively) in the higher plain to above it in the lower plain (Fig. 4-7), would be very high,
634 complicating (rather than facilitating) the management of potentially anthropogenic contaminated sites: high (higher
635 than regulatory limits) predicted NBLs in the higher plain hosting oxic groundwaters would underestimate
636 anthropogenic contaminated sites, leading to environmental drawbacks, whereas low (lower than regulatory limits)
637 predicted NBLs in the lower plain hosting reduced groundwaters would overestimate anthropogenic contaminated sites,
638 overcommitting in vain public offices in charge of contaminated sites management and generating unnecessary costs for
639 site holders. A strategy to overcome this issue would be the use of the cokriging method considering secondary
640 variable(s) with sample density high enough to catch the steep variability in the springs belt area. Possible secondary
641 variables for the interpolation of As and NH_4 NBLs would be the hydraulic conductivity of shallow layers (e.g., first 5
642 m of depth) and/or the content of OM (peat) in the subsurface, that can be derived from lithologs, that, importantly, are
643 likely better sampled than the primary variables. Theoretically, these two secondary variables would be able to catch the
644 spatial structure of As and NH_4 NBLs, i.e., the difference between higher (high-conductive shallow layers) and lower
645 (low-conductive shallow layers) plains and, within the lower plain, the trend of increasing NBLs with increasing peat
646 content. Anyway, these secondary variables would not be able to represent the other important process affecting As and
647 NH_4 NBLs, that is the upwelling of deep brine leading to brackish groundwaters with higher (than surrounding lower
648 plain areas) NH_4 NBLs (as discussed in Sect. 4.1) and lower As NBLs (Fig. 4). So, this prevents an accurate
649 interpolation of the NBLs over the entire study area. The spatial variability of Fe and Mn NBLs is governed by different

processes (highest NBLs with strong Mn-/Fe- reducing conditions mainly occurring within river valley sediments, as discussed in Sect. 4.1), so stratified kriging (considering Holocene river valleys as a stratum) rather than cokriging would be more suitable. Anyway, a stratified kriging may not be able to properly address the issue of low sample density in the highly variable area surrounding the springs belt, as discussed above. So, in conclusion, none of the methods considered above seem suitable for the interpolation of our point NBLs (for the reasons discussed above), so the interpolation on a regional scale was not performed in this study. This choice agrees with the previous work of Biddau et al. (2017) that calculated point NBL at regional scale (Sardinia region, Italy). Future works will be addressed on the interpolation of NBLs of these target species in smaller pilot areas with suitable characteristics (sufficient sample density of primary and secondary viarable(s) with respect to the nature and spatial structure of processes governing the spatial variability of the primary variable) for a meaningful interpolation.

5. Conclusions

The present work assessed local NBLs within 27 GWBs in Lombardy region, N Italy, for groundwater As, NH_4, Fe and Mn. These NBLs were then associated to corresponding natural hydrogeochemical processes/conditions that generated/favored them. More specifically, for the lower Po Plain aquifers, we found that:

- natural concentrations of As up to ~300 µg/L can be reached in groundwaters under methanogenesis, a condition related to the prolonged degradation of peat buried in sediments of the lower plain;
- natural concentrations of NH_4 up to ~6.6 mg/L can be generated by the upwelling of fluids from deep sediments hosting petroleum systems; natural concentrations of NH_4 up to ~4.4 mg/L can be generated as the accumulation of by-products of peat degradation;
- natural concentrations of Fe and Mn up to, respectively, ~6 and ~1.5 mg/L can be generated by the oxidation of younger (Holocene) and less stable Mn and Fe oxides within river valleys.

Concerning methodological aspects, the following key points were highlighted:

- cluster analysis of redox-sensitive species seems a valid alternative to algorithms combining threshold concentrations for performing redox zonation – advantages: segregation of redox zones based on relative variations of the redox-sensitive species, without using a priori absolute threshold concentrations; disadvantages: a single cluster could represent more than one TEAP; a validation of obtained results is needed.
- hybridization of preselection and probability plot methods leads to calculate local NBLs – probability plots of human impact free datasets, obtained through preselection (together with outliers and temporal trends

treatment), lead to identify different natural background subpopulations, linked to different natural processes operating within a GWB, so leading to different (local) NBLs within a GWB;

- spatial interpolation of local point NBLs is recommended only when available data characteristics (sample quality, density and size of primary and secondary variables), nature and spatial structure of processes driving the spatial variability of the target variable, and assumption/properties of the suitable interpolation method allow to obtain meaningful results;
- conceptual model has a fundamental role in calculating NBLs and interpreting/explaining them – the understanding of the main processes driving mobilization and spatial variability of the target species is key for accurately assessing NBLs; the conceptual model should be viewed as an evolving object along the methodological flow: a preliminary conceptual model, derived from literature data/info, should be developed through data-driven techniques (e.g., multivariate analysis) to improve the knowledge of factors/processes driving target species in the specific study area; finally, a robust conceptual model can support the interpretation of each calculated NBL, possibly leading to attributing to each NBL the likely natural generating processes.

In conclusion, this work underlined that the evaluation of site- and redox- specific NBLs achieves a step forward from the commonly used approach of a single NBL per GWB, helping to improve decision-support tools for integrated groundwater resources management and protection.

Funding

This work was supported by Regione Lombardia [grant number FEC 57/2018]. Results of this work were implemented in the regional regulation (D.G.R. 23 novembre 2020 n. XI/3903).

Acknowledgements

We thank Viviane Iacone, Marina Bellotti and Paolo Casciano of Regione Lombardia and Valeria Marchesi, Andrea Fazzone and Cinzia Monti of ARPA Lombardia for supporting and contributing to this study.

Appendix A: Supporting Information

Supporting information to this article can be found online at ...

References

Amorosi, A., Pavesi, M., Ricci Lucchi, M., Sarti, G., Piccin, A., 2008. Climatic signature of cyclic fluvial architecture from the Quaternary of the central Po Plain, Italy. Sediment. Geol. 209, 58–68. https://doi.org/10.1016/J.SEDGEO.2008.06.010

Appelo, C.A.J., Postma, D., 2005. Geochemistry, Groundwater and Pollution, second. ed. Balkema Publishers, Leiden.

ARPA Lombardia, 2021. ARPA Lombardia Lab Analysis Website. https://www.arpalombardia.it/siti/arpalombardia/trasparenza/Pagine/Trasparenza_Pubblicato.aspx?l1=6&l2=32&l3=599 (accessed 12 June 2021).

Azzellino, A., Colombo, L., Lombi, S., Marchesi, V., Piana, A., Andrea, M., Alberti, L., 2019. Groundwater diffuse pollution in functional urban areas: The need to define anthropogenic diffuse pollution background levels. Sci. Total Environ. 656, 1207–1222. https://doi.org/10.1016/j.scitotenv.2018.11.416

Baedecker, M.J., Cozzarelli, I.M., Eganhouse, R.P., Siegel, D.I., Bennett, P.C., 1993. Crude oil in a shallow sand and gravel aquifer—III. Biogeochemical reactions and mass balance modeling in anoxic groundwater. Appl. Geochemistry 8, 569–586. https://doi.org/10.1016/0883-2927(93)90014-8

Balestrini, R., Delconte, C.A., Sacchi, E., Buffagni, A., 2021. Groundwater-dependent ecosystems as transfer vectors of nitrogen from the aquifer to surface waters in agricultural basins: The fontanili of the Po Plain (Italy). Sci. Total Environ. 753, 141995. https://doi.org/10.1016/j.scitotenv.2020.141995

Biddau, R., Cidu, R., Lorrai, M., Mulas, M.G., 2017. Assessing background values of chloride, sulfate and fluoride in groundwater: A geochemical-statistical approach at a regional scale. J. Geochemical Explor. 181, 243–255. https://doi.org/10.1016/j.gexplo.2017.08.002

Böhlke, J.K., Smith, R.L., Miller, D.N., 2006. Ammonium transport and reaction in contaminated groundwater: Application of isotope tracers and isotope fractionation studies. Water Resour. Res. 42. https://doi.org/10.1029/2005WR004349

Bonomi, T., Fumagalli, L., Rotiroti, M., Bellani, A., Cavallin, A., 2014. The hydrogeological well database TANGRAM©: a tool for data processing to support groundwater assessment. Acque Sotter. - Ital. J. Groundw. 3. https://doi.org/10.7343/as-072-14-0098

Bonori, O., Ciabatti, M., Cremonini, S., Di Giovambattista, R., Martinelli, G., Maurizzi, S., Quadri, G., Rabbi, E., Righi,

P., Tinti, S., Zantedeschi, E., 2000. Geochemical and geophysical monitoring intectonically active areas of the Po Valley (Northern Italy). Case histories linked to gas emission structures. Geogr. Fis. e Din. Quat. 23, 3–20.

Burgess, W.G., Pinto, L., 2005. Preliminary observations on the release of arsenic to groundwater in the presence of hydrocarbon contaminants in UK aquifers. Mineral. Mag. 69, 887–896. https://doi.org/10.1180/0026461056950296

Castellarin, A., Vai, G.B., Cantelli, L., 2006. The Alpine evolution of the Southern Alps around the Giudicarie faults: A Late Cretaceous to Early Eocene transfer zone. Tectonophysics 414, 203–223. https://doi.org/10.1016/j.tecto.2005.10.019

Chapman, P.M., 2007. Determining when contamination is pollution — Weight of evidence determinations for sediments and effluents. Environ. Int. 33, 492–501. https://doi.org/10.1016/j.envint.2006.09.001

Cloutier, V., Lefebvre, R., Therrien, R., Savard, M.M., 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

Conti, A., Sacchi, E., Chiarle, M., Martinelli, G., Zuppi, G.M., 2000. Geochemistry of the formation waters in the Po plain (Northern Italy): an overview. Appl. Geochemistry 15, 51–65. https://doi.org/10.1016/S0883-2927(99)00016-5

D. Lgs. 152/06, 2006. Decreto Legislativo n. 152 del 3 aprile 2006 sulle norme in materia ambientale “Legislative Decree on environmental regulations”.

D. Lgs. 30/09, 2009. Decreto Legislativo n. 30 del 16 marzo 2009 sull’attuazione della direttiva 2006/118/CE relativa alla protezione delle acque sotterranee dall’inquinamento e dal deterioramento “Legislative Decree on the implementation of Directive 2006/118/EC on the protection of groundwater against pollution and deterioration".

Dalla Libera, N., Fabbri, P., Mason, L., Piccinini, L., Pola, M., 2018. A local natural background level concept to improve the natural background level: a case study on the drainage basin of the Venetian Lagoon in Northeastern ItalyA local natural background level concept to improve the natural background level: a case study. Environ. Earth Sci. 77, 487. https://doi.org/10.1007/s12665-018-7672-3

Dalla Libera, N., Fabbri, P., Mason, L., Piccinini, L., Pola, M., 2017. Geostatistics as a tool to improve the natural background level definition: An application in groundwater. Sci. Total Environ. 598, 330–340. https://doi.org/10.1016/J.SCITOTENV.2017.04.018

Ducci, D., de Melo, M.T.C., Preziosi, E., Sellerino, M., Parrone, D., Ribeiro, L., 2016. 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

EC, 2006. Directive 2006/118/EC of the European Parliament and of the Council of 12 December 2006 on the protection of groundwater against pollution and deterioration.

EC, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy.

Filippini, M., Zanotti, C., Bonomi, T., Sacchetti, V.G., Amorosi, A., Dinelli, E., Rotiroti, M., 2021. Deriving Natural Background Levels of Arsenic at the Meso-Scale Using Site-Specific Datasets: An Unorthodox Method. Water 13, 452. https://doi.org/10.3390/w13040452

Fumagalli, N., Senes, G., Ferrario, P.S., Toccolini, A., 2017. A minimum indicator set for assessing fontanili (lowland springs) of the Lombardy Region in Italy. Eur. Countrys. 9, 1–16. https://doi.org/10.1515/euco-2017-0001

Gao, Y., Qian, H., Huo, C., Chen, J., Wang, H., 2020. Assessing natural background levels in shallow groundwater in a large semiarid drainage Basin. J. Hydrol. 584, 124638. https://doi.org/10.1016/j.jhydrol.2020.124638

Garzanti, E., Vezzoli, G., Andò, S., 2011. Paleogeographic and paleodrainage changes during Pleistocene glaciations (Po Plain, Northern Italy). Earth-Science Rev. 105, 25–48. https://doi.org/10.1016/J.EARSCIREV.2010.11.004

Giuliano, G., 1995. Ground water in the Po basin: some problems relating to its use and protection. Sci. Total Environ. 171, 17–27. https://doi.org/10.1016/0048-9697(95)04682-1

Guadagnini, L., Menafoglio, A., Sanchez-Vila, X., Guadagnini, A., 2020. Probabilistic assessment of spatial heterogeneity of natural background concentrations in large-scale groundwater bodies through Functional Geostatistics. Sci. Total Environ. 740, 140139. https://doi.org/10.1016/j.scitotenv.2020.140139

Hewett, P., Ganser, G.H., 2007. A Comparison of Several Methods for Analyzing Censored Data. Ann. Occup. Hyg. 51, 611–632. https://doi.org/10.1093/annhyg/mem045

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, 1–20. https://doi.org/10.1016/j.scitotenv.2008.03.018

Hornung, R.W., Reed, L.D., 1990. Estimation of Average Concentration in the Presence of Nondetectable Values. Appl. Occup. Environ. Hyg. 5, 46–51. https://doi.org/10.1080/1047322X.1990.10389587

Hu, Z., Liu, S., Zhong, G., Lin, H., Zhou, Z., 2020. Modified Mann-Kendall trend test for hydrological time series under the scaling hypothesis and its application. Hydrol. Sci. J. 65, 2419–2438. https://doi.org/10.1080/02626667.2020.1810253

ISPRA, 2017. Linee guida recanti la procedura da seguire per il calcolo dei valori di fondo naturale per i corpi idrici sotterranei (DM 6 luglio 2016) “Guidelines for calculating natural background levels for groundwater bodies

(Ministerial Decree 6 July 2016).” Manuali e Linee Guida 155/2017. ISPRA, Rome.

Kaiser, H.F., 1958. The varimax criterion for analytic rotation in factor analysis. Psychometrika 23, 187–200. https://doi.org/10.1007/BF02289233

Kendall, M.G., 1955. Rank Correlation Methods. Griffin, London.

Kim, K.-H., Yun, S.-T., Kim, H.-K., Kim, J.-W., 2015. Determination of natural backgrounds and thresholds of nitrate in South Korean groundwater using model-based statistical approaches. J. Geochemical Explor. 148, 196–205. https://doi.org/10.1016/j.gexplo.2014.10.001

Li, J., Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences: A review. Environ. Model. Softw. 53, 173–189. https://doi.org/10.1016/j.envsoft.2013.12.008

Lindquist, S.J., 1999. Petroleum systems of the Po Basin Province of northern Italy and the northern Adriatic Sea; Porto Garibaldi (biogenic), Meride/Riva di Solto (thermal), and Marnoso Arenacea (thermal), U.S. Geological Survey Open-File Report 99–50-M. https://doi.org/10.3133/ofr9950M

Lions, J., Devau, N., Elster, D., Voutchkova, D.D., Hansen, B., Schullehner, J., Petrović Pantić, T., Samolov, K.A., Camps, V., Arnó, G., Herms, I., Rman, N., Cerar, S., Grima, J., Giménez-Forcada, E., Luque-Espinar, J.A., Malcuit, E., Gourcy, L., 2021. A Broad-Scale Method for Estimating Natural Background Levels of Dissolved Components in Groundwater Based on Lithology and Anthropogenic Pressure. Water 13, 1531. https://doi.org/10.3390/w13111531

Liu, Guowei, Ma, F., Liu, Gang, Zhao, H., Guo, J., Cao, J., 2019. Application of Multivariate Statistical Analysis to Identify Water Sources in A Coastal Gold Mine, Shandong, China. Sustain. 11, 3345. https://doi.org/10.3390/su11123345

Mann, H.B., 1945. Nonparametric Tests Against Trend. Econometrica 13, 245–259. https://doi.org/10.2307/1907187

Marchetti, M., 2002. Environmental changes in the central Po Plain (northern Italy) due to fluvial modifications and anthropogenic activities. Geomorphology 44, 361–373. https://doi.org/10.1016/S0169-555X(01)00183-0

Martinelli, G., Chahoud, A., Dadomo, A., Fava, A., 2014. Isotopic features of Emilia-Romagna region (North Italy) groundwaters: Environmental and climatological implications. J. Hydrol. 519, 1928–1938. https://doi.org/10.1016/J.JHYDROL.2014.09.077

Martinelli, G., Dadomo, A., De Luca, D.A., Mazzola, M., Lasagna, M., Pennisi, M., Pilla, G., Sacchi, E., Saccon, P., 2018. Nitrate sources, accumulation and reduction in groundwater from Northern Italy: Insights provided by a nitrate and boron isotopic database. Appl. Geochemistry 91, 23–35. https://doi.org/10.1016/j.apgeochem.2018.01.011

Masiol, M., Rampazzo, G., Ceccato, D., Squizzato, S., Pavoni, B., 2010. Characterization of PM10 sources in a coastal area near Venice (Italy): An application of factor-cluster analysis. Chemosphere 80, 771–778. https://doi.org/10.1016/j.chemosphere.2010.05.008

McArthur, J.., Banerjee, D.., Hudson-Edwards, K.., Mishra, R., Purohit, R., Ravenscroft, P., Cronin, A., Howarth, R.., Chatterjee, A., Talukder, T., Lowry, D., Houghton, S., Chadha, D.., 2004. Natural organic matter in sedimentary basins and its relation to arsenic in anoxic ground water: the example of West Bengal and its worldwide implications. Appl. Geochemistry 19, 1255–1293. https://doi.org/10.1016/J.APGEOCHEM.2004.02.001

McMahon, P.B., Chapelle, F.H., 2008. Redox Processes and Water Quality of Selected Principal Aquifer Systems. Groundwater 46, 259–271. https://doi.org/10.1111/j.1745-6584.2007.00385.x

Mendizabal, I., Stuyfzand, P.J., 2011. Quantifying the vulnerability of well fields towards anthropogenic pollution: The Netherlands as an example. J. Hydrol. 398, 260–276. https://doi.org/10.1016/j.jhydrol.2010.12.026

Michetti, A.M., Giardina, F., Livio, F., Mueller, K., Serva, L., Sileo, G., Vittori, E., Devoti, R., Riguzzi, F., Carcano, C., Rogledi, S., Bonadeo, L., Brunamonte, F., Fioraso, G., 2012. Active compressional tectonics, Quaternary capable faults, and the seismic landscape of the Po Plain (northern Italy). Ann. Geophys. 55, 969–1001. https://doi.org/10.4401/ag-5462

Miola, A., Bondesan, A., Corain, L., Favaretto, S., Mozzi, P., Piovan, S., Sostizzo, I., 2006. Wetlands in the Venetian Po Plain (northeastern Italy) during the Last Glacial Maximum: Interplay between vegetation, hydrology and sedimentary environment. Rev. Palaeobot. Palynol. 141, 53–81. https://doi.org/10.1016/J.REVPALBO.2006.03.016

Molinari, A., Guadagnini, L., Marcaccio, M., Guadagnini, A., 2019. Geostatistical multimodel approach for the assessment of the spatial distribution of natural background concentrations in large-scale groundwater bodies. Water Res. 149, 522–532. https://doi.org/10.1016/j.watres.2018.09.049

Muller, D., Blum, A., Hart, A., Hookey, J., Kunkel, R., Scheidleder, A., Tomlin, C., Wendland, F., 2006. Final proposal for a methodology to set up groundwater threshold values in Europe. Report to EU Project BRIDGE, Deliverable D18.

Musacchio, A., Rotiroti, M., Re, V., Oster, H., Leoni, B., Sacchi, E., 2018. Estimation of nutrients' dynamics and residence time in groundwater from the Lombardy plain (northern Italy), in: Proceedings of the 45th IAH Congress. Daejeon, Korea, p. 486.

Nakić, Z., Kovač, Z., Parlov, J., Perković, D., 2020. Ambient Background Values of Selected Chemical Substances in Four Groundwater Bodies in the Pannonian Region of Croatia. Water 12, 2671. https://doi.org/10.3390/w12102671

Naveen, B.P., Mahapatra, D.M., Sitharam, T.G., Sivapullaiah, P. V, Ramachandra, T. V, 2017. Physico-chemical and biological characterization of urban municipal landfill leachate. Environ. Pollut. 220, 1–12. https://doi.org/10.1016/j.envpol.2016.09.002

Ori, G.G., 1993. Continental depositional systems of the Quaternary of the Po Plain (northern Italy). Sediment. Geol. 83, 1–14. https://doi.org/10.1016/S0037-0738(10)80001-6

Palau, J., Marchesi, M., Chambon, J.C.C., Aravena, R., Canals, À., Binning, P.J., Bjerg, P.L., Otero, N., Soler, A., 2014. Multi-isotope (carbon and chlorine) analysis for fingerprinting and site characterization at a fractured bedrock aquifer contaminated by chlorinated ethenes. Sci. Total Environ. 475, 61–70. https://doi.org/10.1016/j.scitotenv.2013.12.059

Parrone, D., Frollini, E., Preziosi, E., Ghergo, S., 2021. eNaBLe, an On-Line Tool to Evaluate Natural Background Levels in Groundwater Bodies. Water 13, 74. https://doi.org/10.3390/w13010074

Parrone, D., Ghergo, S., Preziosi, E., 2019. A multi-method approach for the assessment of natural background levels in groundwater. Sci. Total Environ. 659, 884–894. https://doi.org/10.1016/j.scitotenv.2018.12.350

Peña Reyes, F.A., Crosta, G.B., Frattini, P., Basiricò, S., Della Pergola, R., 2015. Hydrogeochemical overview and natural arsenic occurrence in groundwater from alpine springs (upper Valtellina, Northern Italy). J. Hydrol. 529, 1530–1549. https://doi.org/10.1016/j.jhydrol.2015.08.029

Perego, R., Bonomi, T., Fumagalli, M.L., Benastini, V., Aghib, F., Rotiroti, M., Cavallin, A., 2014. 3D reconstruction of the multi-layer aquifer in a Po Plain area. Rend. Online Soc. Geol. Ital. 30, 41–44. https://doi.org/10.3301/ROL.2014.09

Pettine, M., Camusso, M., Martinotti, W., Marchetti, R., Passino, R., Queirazza, G., 1994. Soluble and particulate metals in the Po River: Factors affecting concentrations and partitioning. Sci. Total Environ. 145, 243–265. https://doi.org/10.1016/0048-9697(94)90118-X

Pilla, G., Torrese, P., Bersan, M., 2015. The Uprising of Deep Saline Paleo-Waters into the Oltrepò Pavese Aquifer (Northern Italy): Application of Hydro-Chemical and Shallow Geophysical Surveys, in: Lollino, G., Arattano, M., Rinaldi, M., Giustolisi, O., Marechal, J.-C., Grant, G.E. (Eds.), Engineering Geology for Society and Territory - Volume 3. Springer, Cham, pp. 393–397. https://doi.org/10.1007/978-3-319-09054-2_82

Pilla, G., Torrese, P., Bersan, M., 2010. Application of hydrochemical and preliminary geophysical surveys within the study of the saltwater uprising occurring in the Oltrepò Pavese plain aquifer. Boll. di Geofis. Teor. ed Appl. 51, 301–323.

Preziosi, E., Parrone, D., Del Bon, A., Ghergo, S., 2014. Natural background level assessment in groundwaters:

probability plot versus pre-selection method. J. Geochemical Explor. 143, 43–53. https://doi.org/10.1016/j.gexplo.2014.03.015

Ravenscroft, P., Brammer, H., Richards, K., 2009. Arsenic Pollution: A Global Synthesis. Wiley-Blackwell, Chichester. https://doi.org/10.1002/9781444308785

Regione Lombardia, 2016. Programma di tutela e uso delle acque (PTUA 2016) "Programme for the protection and use of water". General Report, 388 pp.

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

Rotiroti, M., Bonomi, T., Sacchi, E., McArthur, J.M., Jakobsen, R., Sciarra, A., Etiope, G., Zanotti, C., Nava, V., Fumagalli, L., Leoni, B., 2021. Overlapping redox zones control arsenic pollution in Pleistocene multi-layer aquifers, the Po Plain (Italy). Sci. Total Environ. 758, 143646. https://doi.org/10.1016/j.scitotenv.2020.143646

Rotiroti, M., Bonomi, T., Sacchi, E., McArthur, J.M., Stefania, G.A., Zanotti, C., Taviani, S., Patelli, M., Nava, V., Soler, V., Fumagalli, L., Leoni, B., 2019. The effects of irrigation on groundwater quality and quantity in a human-modified hydro-system: The Oglio River basin, Po Plain, northern Italy. Sci. Total Environ. 672, 342–356. https://doi.org/10.1016/J.SCITOTENV.2019.03.427

Rotiroti, M., Fumagalli, L., Bonomi, T., 2014a. How to manage potential groundwater contaminations by As, Fe and Mn in lower Po Plain: a proposal from the case study of Cremona. Acque Sotter. - Ital. J. Groundw. 3, 9–16. https://doi.org/10.7343/as-070-14-0096

Rotiroti, M., Fumagalli, L., Frigerio, M.C., Stefania, G.A., Simonetto, F., Capodaglio, P., Bonomi, T., 2015. Natural background levels and threshold values of selected species in the alluvial aquifers in the Aosta Valley Region (N Italy). Rend. Online Soc. Geol. Ital. 35. https://doi.org/10.3301/ROL.2015.114

Rotiroti, M., Sacchi, E., Fumagalli, L., Bonomi, T., 2014b. Origin of arsenic in groundwater from the multilayer aquifer in cremona (Northern Italy). Environ. Sci. Technol. 48, 5395–5403. https://doi.org/10.1021/es405805v

Sacchi, E., Acutis, M., Bartoli, M., Brenna, S., Delconte, C.A., Laini, A., Pennisi, M., 2013. Origin and fate of nitrates in groundwater from the central Po plain: Insights from isotopic investigations. Appl. Geochemistry 34, 164–180. https://doi.org/10.1016/j.apgeochem.2013.03.008

Sacchi, E., Brenna, S., Fornelli Genot, S., Leoni, A., Sale, V.M., Setti, M., 2020. Potentially Toxic Elements (PTEs) in Cultivated Soils from Lombardy (Northern Italy): Spatial Distribution, Origin, and Management Implications. Miner. 10, 298. https://doi.org/10.3390/min10040298

Scardia, G., Festa, A., Monegato, G., Pini, R., Rogledi, S., Tremolada, F., Galadini, F., 2015. Evidence for late Alpine

tectonics in the Lake Garda area (northern Italy) and seismogenic implications. GSA Bull. 127, 113–130. https://doi.org/10.1130/B30990.1

Schwientek, M., Einsiedl, F., Stichler, W., Stögbauer, A., Strauss, H., Maloszewski, P., 2008. Evidence for denitrification regulated by pyrite oxidation in a heterogeneous porous groundwater system. Chem. Geol. 255, 60–67. https://doi.org/10.1016/j.chemgeo.2008.06.005

Şen, Z., 2017. Hydrological trend analysis with innovative and over-whitening procedures. Hydrol. Sci. J. 62, 294–305. https://doi.org/10.1080/02626667.2016.1222533

Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples)†. Biometrika 52, 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Sinclair, A.J., 1974. Selection of threshold values in geochemical data using probability graphs. J. Geochemical Explor. 3, 129–149. https://doi.org/10.1016/0375-6742(74)90030-2

Stefania, G.A., Rotiroti, M., Buerge, I.J., Zanotti, C., Nava, V., Leoni, B., Fumagalli, L., Bonomi, T., 2019. Identification of groundwater pollution sources in a landfill site using artificial sweeteners, multivariate analysis and transport modeling. Waste Manag. 95, 116–128. https://doi.org/10.1016/j.wasman.2019.06.010

Stevenazzi, S., Masetti, M., Nghiem, S. V, Sorichetta, A., 2015. Groundwater vulnerability maps derived from a time-dependent method using satellite scatterometer data. Hydrogeol. J. 23, 631–647. https://doi.org/10.1007/s10040-015-1236-3

Tukey, J.W., 1972. Some graphic and semigraphic displays, in: Bancroft, T.A., Brown, S.A. (Eds.), Statistical Papers in Honor of George W. Snedecor. Iowa State University Press, Ames, pp. 293–316.

US EPA, 2006. Data Quality Assessment: Statistical Methods for Practitioners, EPA QA/G-9S. US EPA, Washington.

Vaclavkova, S., Jørgensen, C.J., Jacobsen, O.S., Aamand, J., Elberling, B., 2014. The Importance of Microbial Iron Sulfide Oxidation for Nitrate Depletion in Anoxic Danish Sediments. Aquat. Geochemistry 20, 419–435. https://doi.org/10.1007/s10498-014-9227-x

Voutchkova, D.D., Ernstsen, V., Schullehner, J., Hinsby, K., Thorling, L., Hansen, B., 2021. Roadmap for Determining Natural Background Levels of Trace Metals in Groundwater. Water 13, 1267. https://doi.org/10.3390/w13091267

Wang, C., Caja, J., Gómez, E., 2018. Comparison of methods for outlier identification in surface characterization. Measurement 117, 312–325. https://doi.org/10.1016/j.measurement.2017.12.015

Wang, J., Zuo, R., Caers, J., 2017. Discovering geochemical patterns by factor-based cluster analysis. J. Geochemical Explor. 181, 106–115. https://doi.org/10.1016/j.gexplo.2017.07.006

Ward, J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. J. Am. Stat. Assoc. 58, 236–244.

https://doi.org/10.1080/01621459.1963.10500845

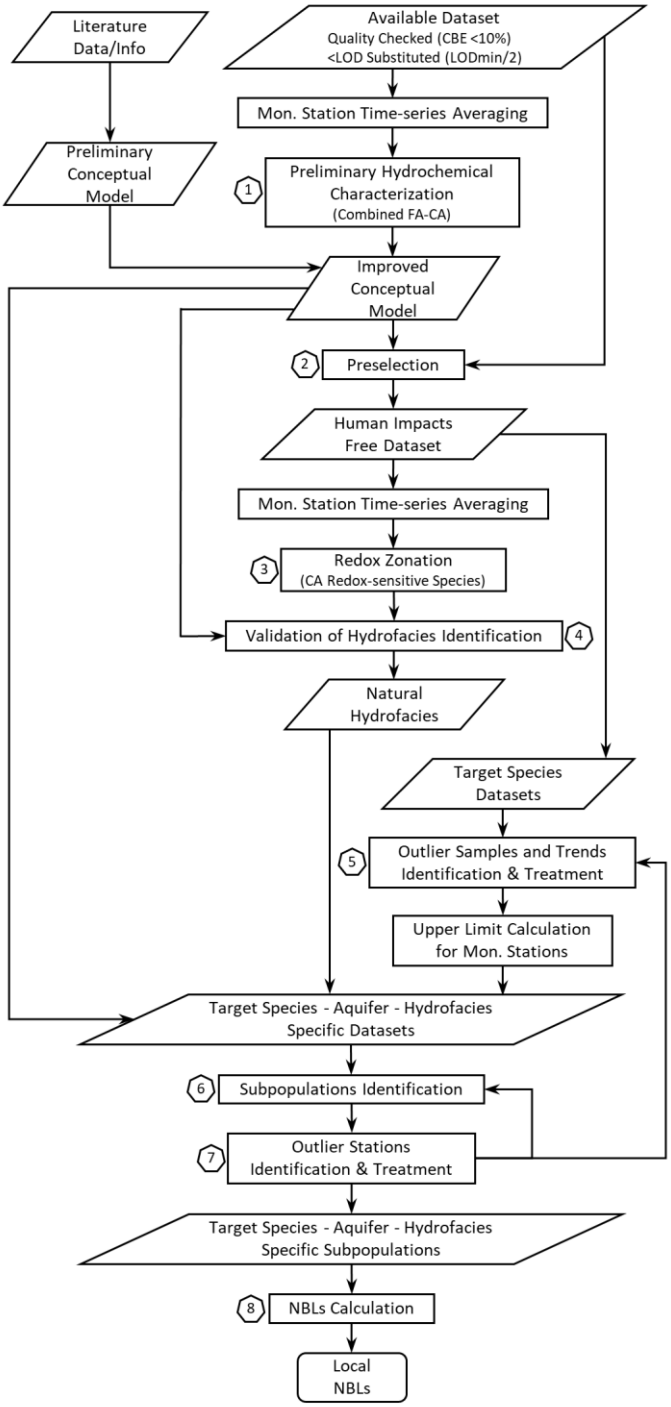
Welch, A.H., Westjohn, D.B., Helsel, D.R., Wanty, R.B., 2000. Arsenic in Ground Water of the United States: Occurrence and Geochemistry. Groundwater 38, 589–604. https://doi.org/10.1111/j.1745-6584.2000.tb00251.x

Wendland, F., Hannappel, S., Kunkel, R., Schenk, R., Voigt, H.J., Wolter, R., 2005. A procedure to define natural groundwater conditions of groundwater bodies in Germany. Water Sci. Technol. 51, 249–257. https://doi.org/10.2166/wst.2005.0598

Williams, L.B., Wilcoxon, B.R., Ferrell, R.E., Sassen, R., 1992. Diagenesis of ammonium during hydrocarbon maturation and migration, Wilcox Group, Louisiana, U.S.A. Appl. Geochemistry 7, 123–134. https://doi.org/10.1016/0883-2927(92)90031-W

Zhang, Y.-C., Slomp, C.P., Broers, H.P., Passier, H.F., Cappellen, P. Van, 2009. Denitrification coupled to pyrite oxidation and changes in groundwater quality in a shallow sandy aquifer. Geochim. Cosmochim. Acta 73, 6716–6726. https://doi.org/10.1016/j.gca.2009.08.026

Zuppi, G.M., Sacchi, E., 2004. Hydrogeology as a climate recorder: Sahara–Sahel (North Africa) and the Po Plain (Northern Italy). Glob. Planet. Change 40, 79–91. https://doi.org/10.1016/S0921-8181(03)00099-7



980

981 Figure 1. Flowchart of the tired approach composing the hybrid preselection-probability plot method for assessing local

982 NBLs.

983

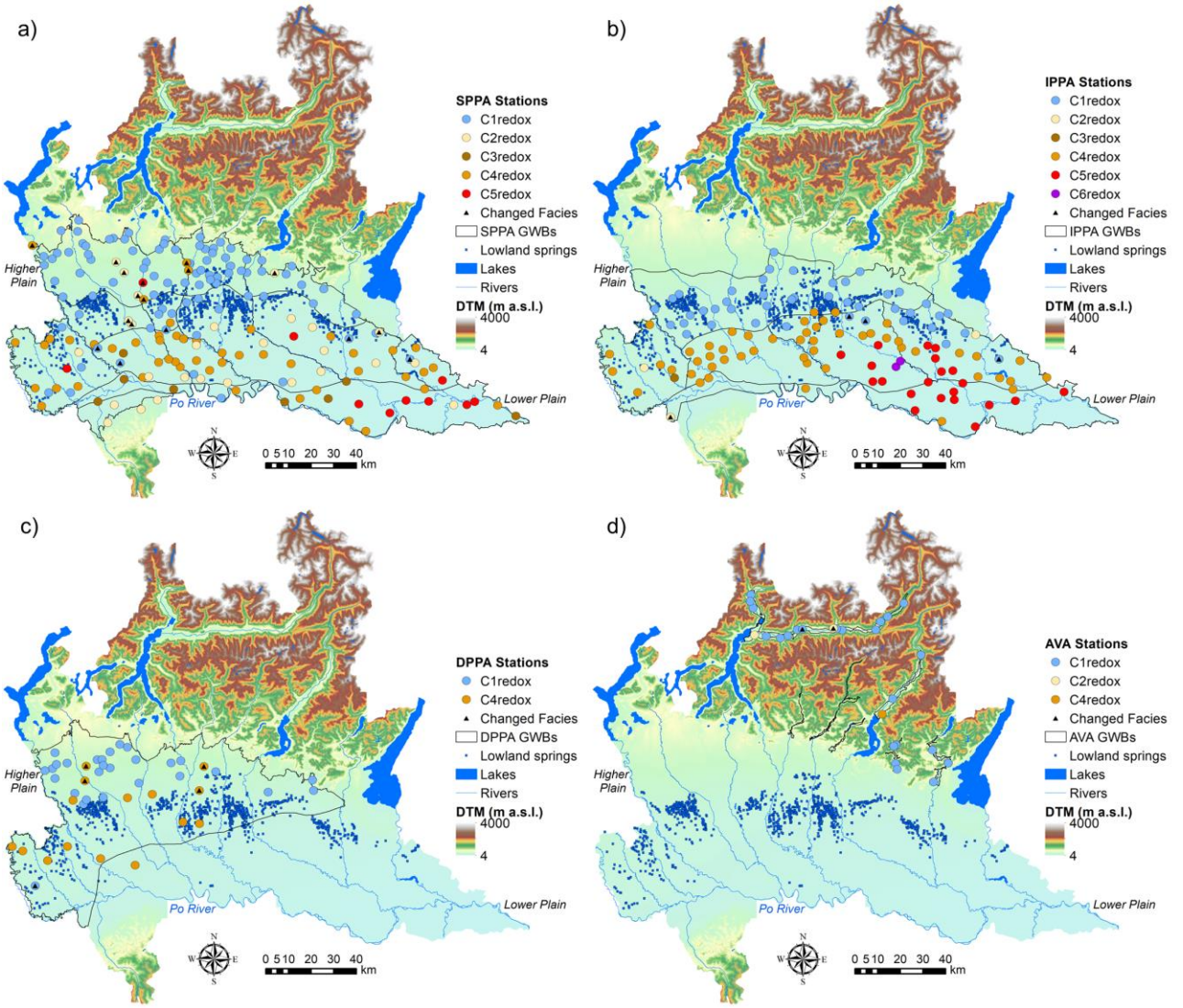


Figure 2. Results of CA of the redox zonation and stations which experienced the hydrofacies change during validation for a) shallow Po Plain aquifers (SPPA), b) intermediate Po Plain aquifers (IPPA), c) deep Po Plain aquifers (DPPA) and d) Alpine valley aquifers (AVA).

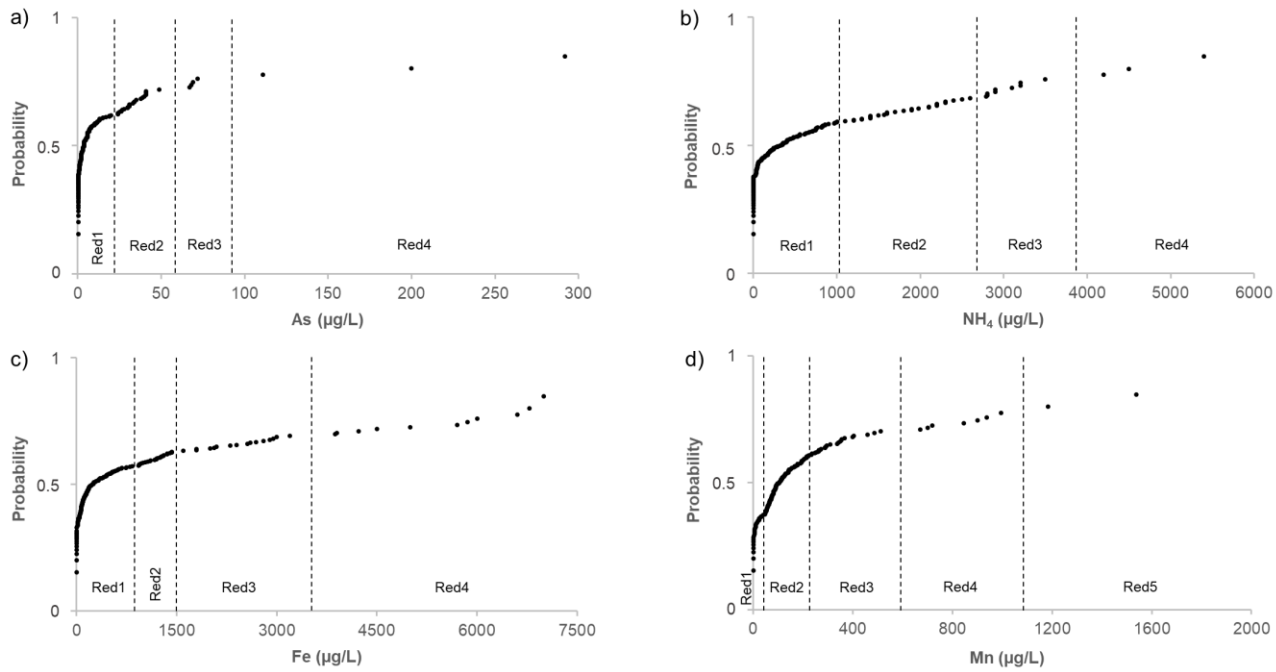


Figure 3. Probability plot for stations tapping Po Plain reduced groundwaters for a) As, b) NH_4, c) Fe and d) Mn; inflection points are marked by dotted line; labels indicate subpopulation names.

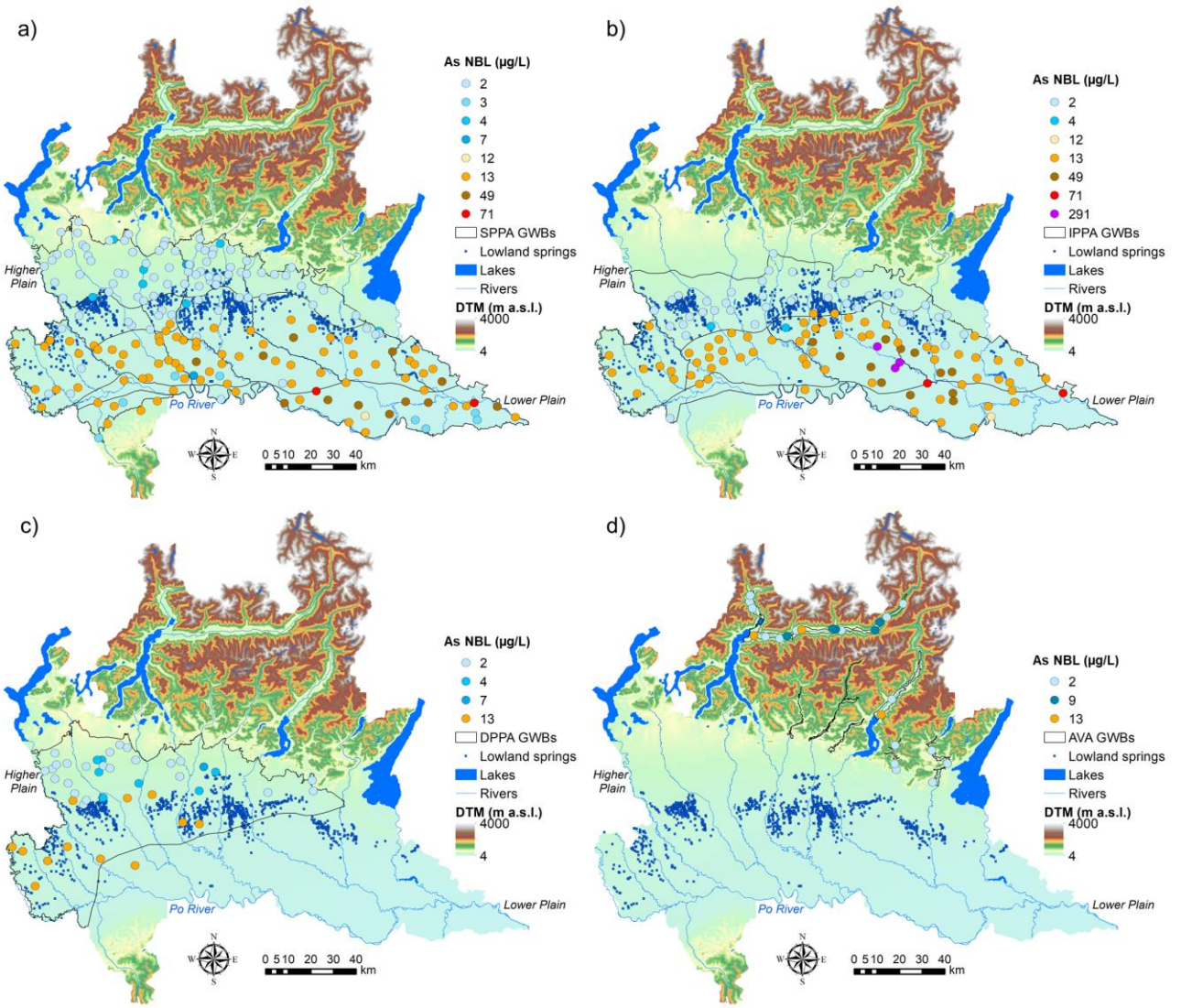


Figure 4. NBLs for As for a) shallow Po Plain aquifers (SPPA), b) intermediate Po Plain aquifers (IPPA), c) deep Po Plain aquifers (DPPA) and d) Alpine valley aquifers (AVA).

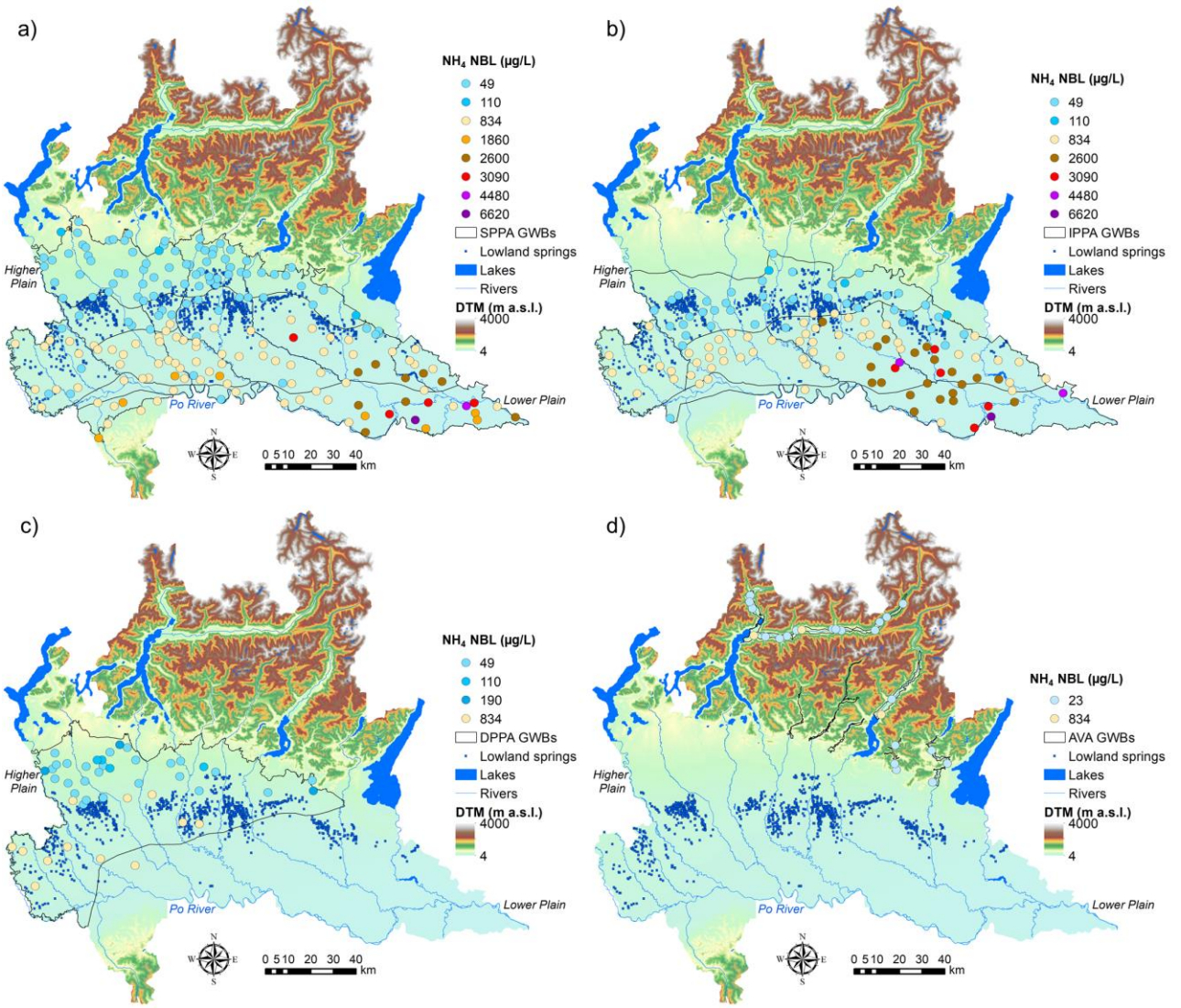


Figure 5. NBLs for NH_4 for a) shallow Po Plain aquifers (SPPA), b) intermediate Po Plain aquifers (IPPA), c) deep Po Plain aquifers (DPPA) and d) Alpine valley aquifers (AVA).

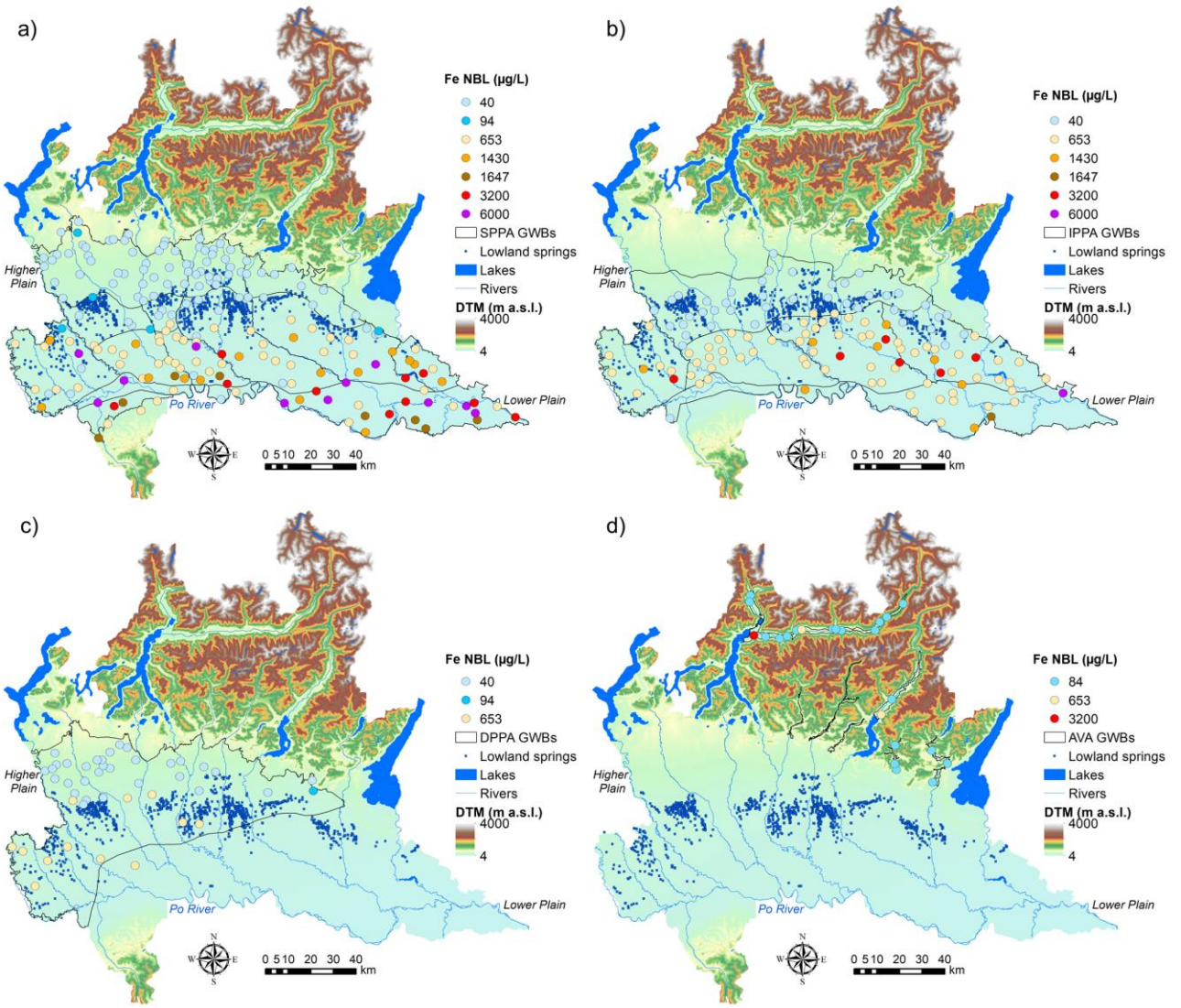


Figure 6. NBLs for Fe for a) shallow Po Plain aquifers (SPPA), b) intermediate Po Plain aquifers (IPPA), c) deep Po Plain aquifers (DPPA) and d) Alpine valley aquifers (AVA).

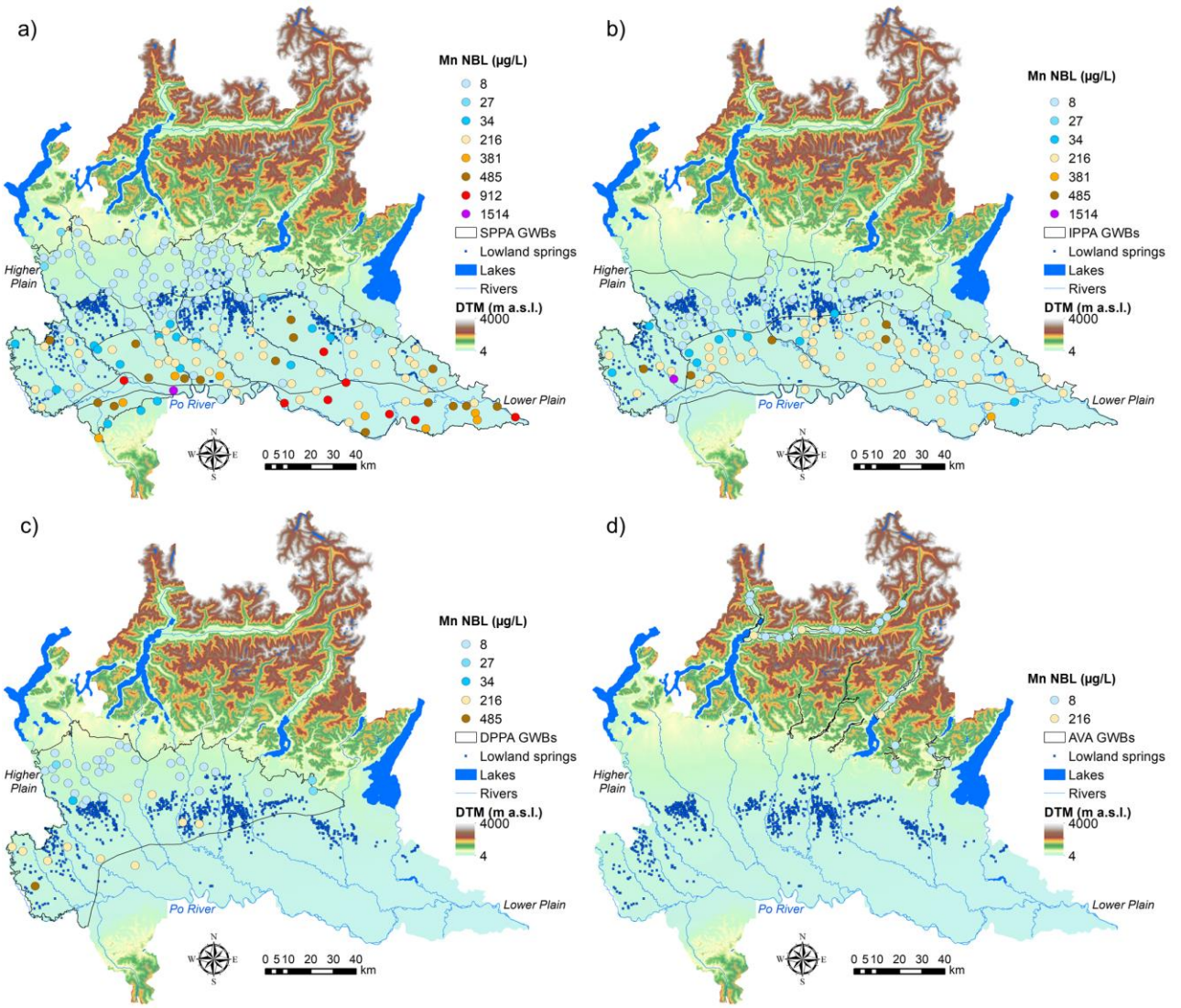


Figure 7. NBLs for Mn for a) shallow Po Plain aquifers (SPPA), b) intermediate Po Plain aquifers (IPPA), c) deep Po Plain aquifers (DPPA) and d) Alpine valley aquifers (AVA).

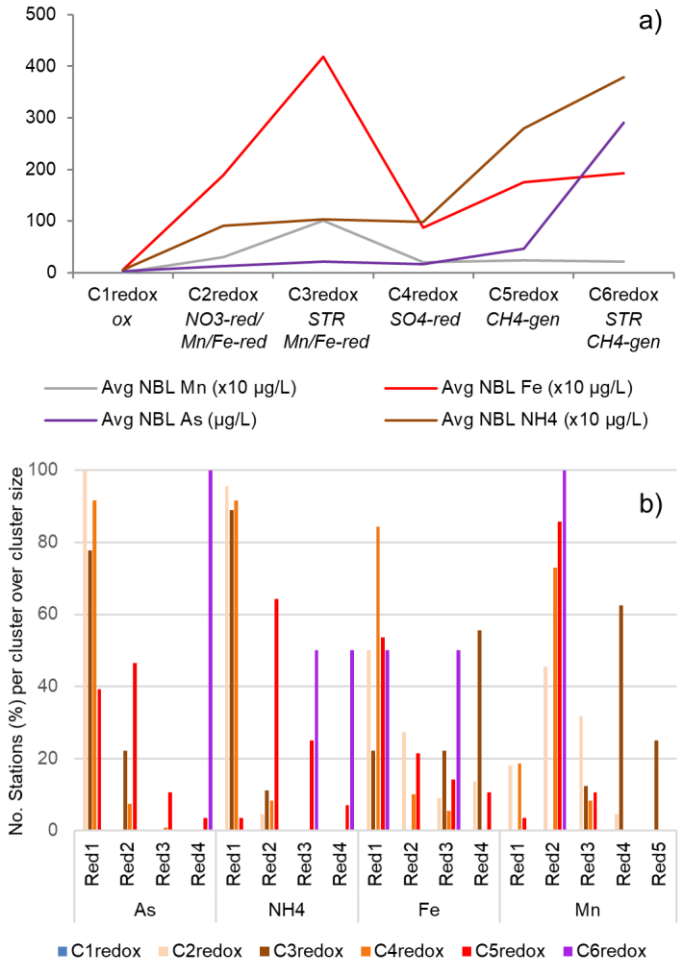


Figure 8. a) Average NBLs for each cluster of the redox zonation (C1-6_{redox}). b) Percentage of monitoring stations per cluster of the redox zonation over the entire cluster size for the subpopulations identified for Po Plain reduced groundwaters.