eNaBLe, an On-Line Tool to Evaluate Natural Background Levels in Groundwater Bodies

Daniele Parrone [ID], Eleonora Frollini *[ID], Elisabetta Preziosi [ID] and Stefano Ghergo [ID]

IRSA-CNR, Water Research Institute—National Research Council, Via Salaria km 29.300, PB 10,
00015 Monterotondo (Rome), Italy; parrone@irsa.cnr.it (D.P.); preziosi@irsa.cnr.it (E.P.); ghergo@irsa.cnr.it (S.G.)
* Correspondence: frollini@irsa.cnr.it

Abstract: Inorganic compounds in groundwater may derive from both natural processes and anthropogenic activities. The assessment of natural background levels (NBLs) is often useful to distinguish these sources. The approaches for the NBLs assessment can be classified as geochemical (e.g., the well-known pre-selection method) or statistical, the latter involving the application of statistical procedures to separate natural and anthropogenic populations. National Guidelines for the NBLs assessment in groundwater have been published in Italy (ISPRA 155/2017), based mainly on the pre-selection method. The Guidelines propose different assessment paths according to the sample size in spatial/temporal dimension and the type of the distribution of the pre-selected dataset, taking also into account the redox conditions of the groundwater body. The obtained NBLs are labelled with a different confidence level in function of number of total observations/monitoring sites, extension of groundwater body and aquifer type (confined or unconfined). To support the implementation of the Guidelines, the on-line tool evaluation of natural background levels (eNaBLe), written in PHP and using MySQL as DBMS (DataBase Management System), has been developed. The main goal of this paper is to describe the functioning of eNaBLe and test the tool on a case study in central Italy. We calculated the NBLs of As, F, Fe and Mn in the southern portion of the Mounts Vulsini groundwater body, within the volcanic province of Latium (Central Italy), also separating the reducing and oxidizing facies. Specific results aside, this study allowed to verify the functioning and possible improvements of the online tool and to identify some criticalities in the procedure NBLs assessment at the groundwater body scale

Keywords: natural background; groundwater body; conceptual model; preselection; nitrates; confidence level

1. Introduction

The presence of inorganic potentially toxic elements in groundwater represents a significant problem in many parts of the world. They may derive from both natural processes and anthropogenic activities and natural background levels (NBLs) assessment is often used to distinguish these sources. The NBL has been defined in the Groundwater Directive (GWD) [1] as "the concentration of a substance or the value of an indicator in a groundwater body corresponding to no, or extremely limited, anthropogenic alterations, compared to unaltered conditions". The GWD requests the EU Member States to individuate appropriate threshold values (TVs) for various potentially harmful substances, taking into account NBLs when necessary, in order to evaluate the chemical status of groundwater bodies.

Currently, it is possible to distinguish a geochemical and a statistical approach for the NBLs assessment. The geochemical approach, originally called "pre-selection" (PS) within the BRIDGE (Background cRiteria for the Identification of Groundwater thrEsholds) project [2], requires the identification of groundwater with no or negligible human impact, using markers such as nitrates/ammonia in oxidizing/reducing environments, organic compounds and isotopes. Once the samples not affected by anthropogenic impact have been selected, a

representative value (usually a chosen percentile) of the natural background is derived. Many examples of PS method applications, including comparisons between groundwater bodies of different European countries [3,4], can be found in the literature [5–12].

The statistical approach involves the separation of uninfluenced and influenced populations by means of statistical procedures. For this purpose, numerous statistical techniques have been developed and tested. Some of these methods point to eliminate the outliers, assuming that the remaining data belong to the natural background. In some cases, the same methods can be used to separate different data populations. The "mean + 2σ" [13], the Median Absolute Deviation (MAD) [14], the Box and Whisker plot [15], the component separation [11,16] and other parametric or non-parametric techniques, including graphical methods as probability plots or quantile–quantile plots, have been largely used for this purpose [17–20]. Please refer to Preziosi et al. [21] for a review of the most common approaches for NBLs assessment, including many statistical techniques and the pre-selection methods.

In 2017, national Guidelines for the NBLs assessment in groundwater were published in Italy [22]. Starting from these Guidelines, Frollini et al. [23] applied the procedure to define the NBLs at the groundwater body (GWB) scale, while Parrone et al. [24] tested a multi-methodological approach on two case studies, also suggesting new criteria for the choice of the nitrate threshold to be used for the pre-selection of non-contaminated samples. An automated approach implementing both the component separation and the pre-selection methods was recently proposed by Chidichimo et al. [25].

The procedure is based mainly on the pre-selection method but different assessment paths are proposed according to the redox conditions of the GWB, the sample size in both the spatial and temporal dimension and the type of the distribution, normal or not, of the pre-selected dataset. The complexity of the schema is essential because of the great heterogeneity of the different GWB monitoring data. For this reason, moreover, the obtained NBLs are labelled with a different confidence level in function of: The number of total observations or the number of total monitoring sites (MSs), extension of groundwater body and aquifer type (confined or unconfined). For GWBs characterized by NBLs with low confidence levels, further monitoring activities are requested.

To provide support to operators involved in the use of the Italian Guidelines and to achieve a harmonization of procedures between the different structures involved, evaluation of natural background levels (eNaBLe), an on-line tool implementing the sequence of operations for NBL assessment, was developed at Water Research Institute – National Research Council (IRSA-CNR) [26]. The procedure implemented in the eNaBLe tool, written in PHP and using MySQL as DBMS, is organized into three logical blocks:

- Selection of the calculation parameters;
- NBL calculation for all the chemical parameters that show concentration values exceeding the relative TV;
- Graphical output of the results.

The main goal of this work is to test the eNaBLe tool on a case study in Central Italy. We calculated the NBLs of As, F, Fe and Mn in the southern portion of the Mounts Vulsini GWB, within the volcanic province of Latium (Central Italy), also operating a separation of data based on the different redox conditions found in the study area.

2. Materials and Methods

This section is dedicated to the detailed description of the tool and of all the steps that lead to the calculation of the natural background values. A description of the case study to which the Guidelines were applied is also included.

2.1. The Italian Guidelines and the eNaBLe Tool

eNaBLe operates in compliance with the provisions of the Italian Guidelines for NBLs assessment, the general scheme of which is shown in Figure 1.
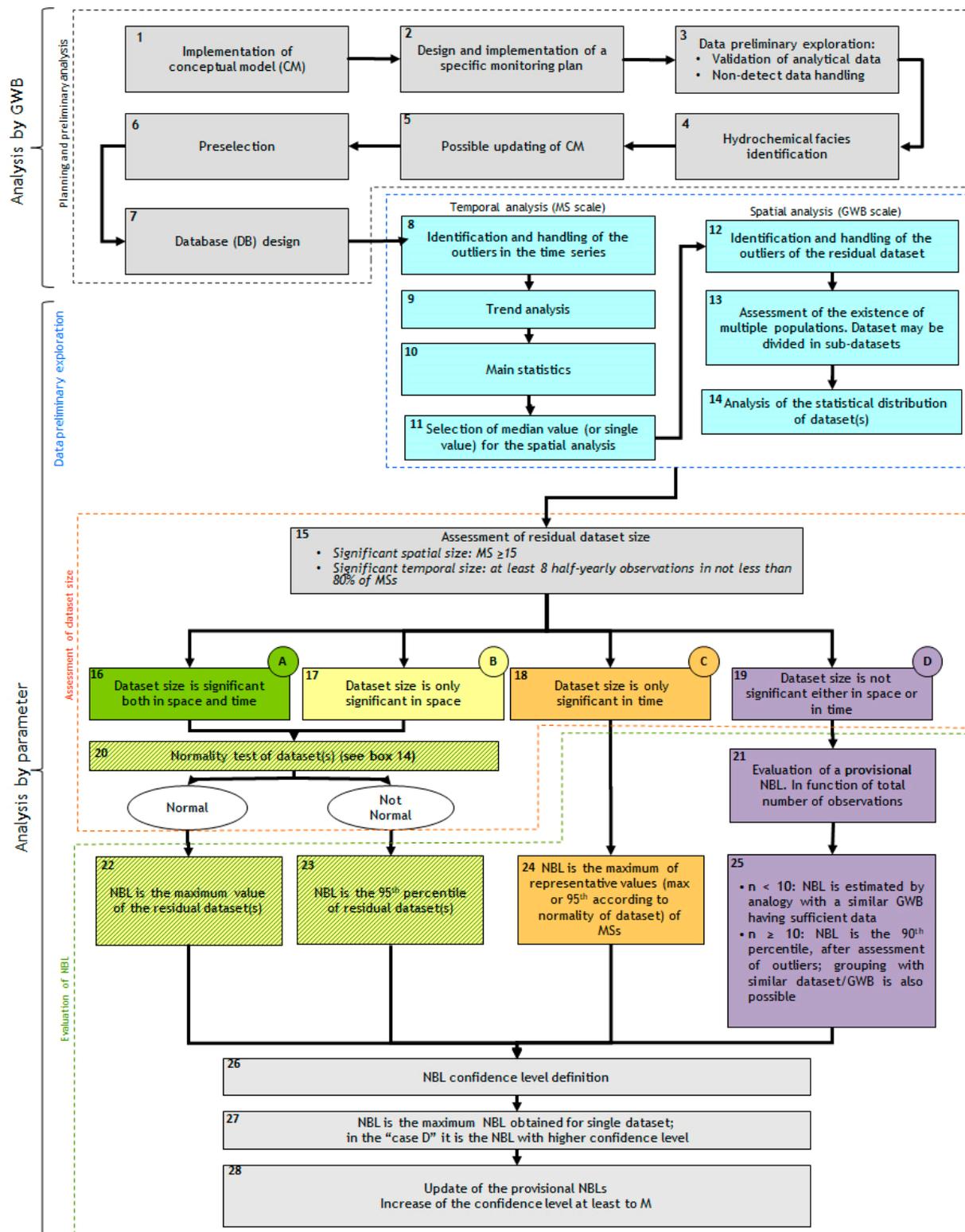
Figure 1. Italian Guidelines: Scheme of the procedure for the natural background levels (NBLs) assessment at the Groundwater Body (GWB) scale [23].

The general methodology adopted by the Guidelines is that of preselection which is accompanied by statistical evaluations in order to adapt the application of the procedure to the various groundwater bodies and allow the use of datasets of different numerical consistency. Four different evaluation paths (A, B, C, D) were therefore identified according

to the spatial and temporal consistency of the available dataset. As a consequence of the great differences between the various GWB, both in terms of extension and number of monitoring networks, it was considered appropriate to associate also an index (confidence level) to the defined NBLs linked to the size of the statistical sample, to the extension and type of the aquifer (unconfined or confined) (Table 1). For more detailed explanation of the entire procedure, see [22].

Table 1. Confidence levels. The confidence level attributed to a NBL takes into account the characteristics of the dataset (case A, B, C, D), of the total number of observations and monitoring sites (MSs), also with reference to the extension of the GWB and the aquifer type (unconfined/confined). Confidence levels: H: High; M: Medium; L: Low; and LL: Very Low.

N. Total obs.	N. Total MS	Size of GWB of Portion of GWB Represented by Dataset (km^2) Aquifer Type (Unconfined/Confined)							
		<10 km^2		10–70 km^2		70–700 km^2		>700 km^2	
		Unconf.	Conf.	Unconf.	Conf.	Unconf.	Conf.	Unconf.	Conf.
	15–25	H	H	H	H	M	H	M	H
	>25	H	H	H	H	H	H	H	H
	15–25	M	H	M	H	L	M	L	M
	>25	H	H	H	H	M	H	M	M
≤15		M	M	L	M	LL	L	LL	LL
16–30		M	H	M	M	L	M	LL	L
>30		H	H	M	H	M	H	M	M
<10		L	L	LL	LL	LL	LL	LL	LL
≥10		L	L	L	L	LL	L	LL	LL

The complexity of the Guidelines framework advised the demand for an automated procedure which, while following the general formulation and the involved provisions, could facilitate the task of the operators. In 2018, a first automated approach was developed at IRSA-CNR. Later, the procedure was implemented as the eNaBLe on-line tool. The complete procedure, written in PHP [27] and using MySQL [28] as relational databases manager, is structured in modules, each one performing specific tasks on the input data, using specific configuration parameters. Some of these parameters are inherited from the Guidelines and cannot be changed; the others can be modified by the user. Each resulting dataset is then sent to the next module. The overall breakdown of eNable tool is shown in Figure 2. The modularity of eNaBLe has allowed us to easily integrate it into the Institute's Water Resources Database [29] which now is the source for the raw data for the NBL evaluation.

The user interface for the tool is divided into three logical phases which correspond to three different procedures and therefore to three distinct associated web pages: Selection of the analytical parameters for the evaluation of the NBLs and selection of the calculation options to be used; NBLs assessment and relative confidence levels; and output of results in graphical, tabular and report form. These procedures will now be described schematically.

2.1.1. Choice of the GWB, of the Parameters and the Calculation Options to Be Used for the NBLs Assessment.

Performing a query on the Water Resource Database [29], eNaBLe produces a table in which all the analytical parameters and their number are listed, proposing for the NBLs evaluation those parameters in which at least one exceedance of the relative TVs has been found. It is therefore possible to select the upper limits to be used in the validation, preselection and calculation procedures.
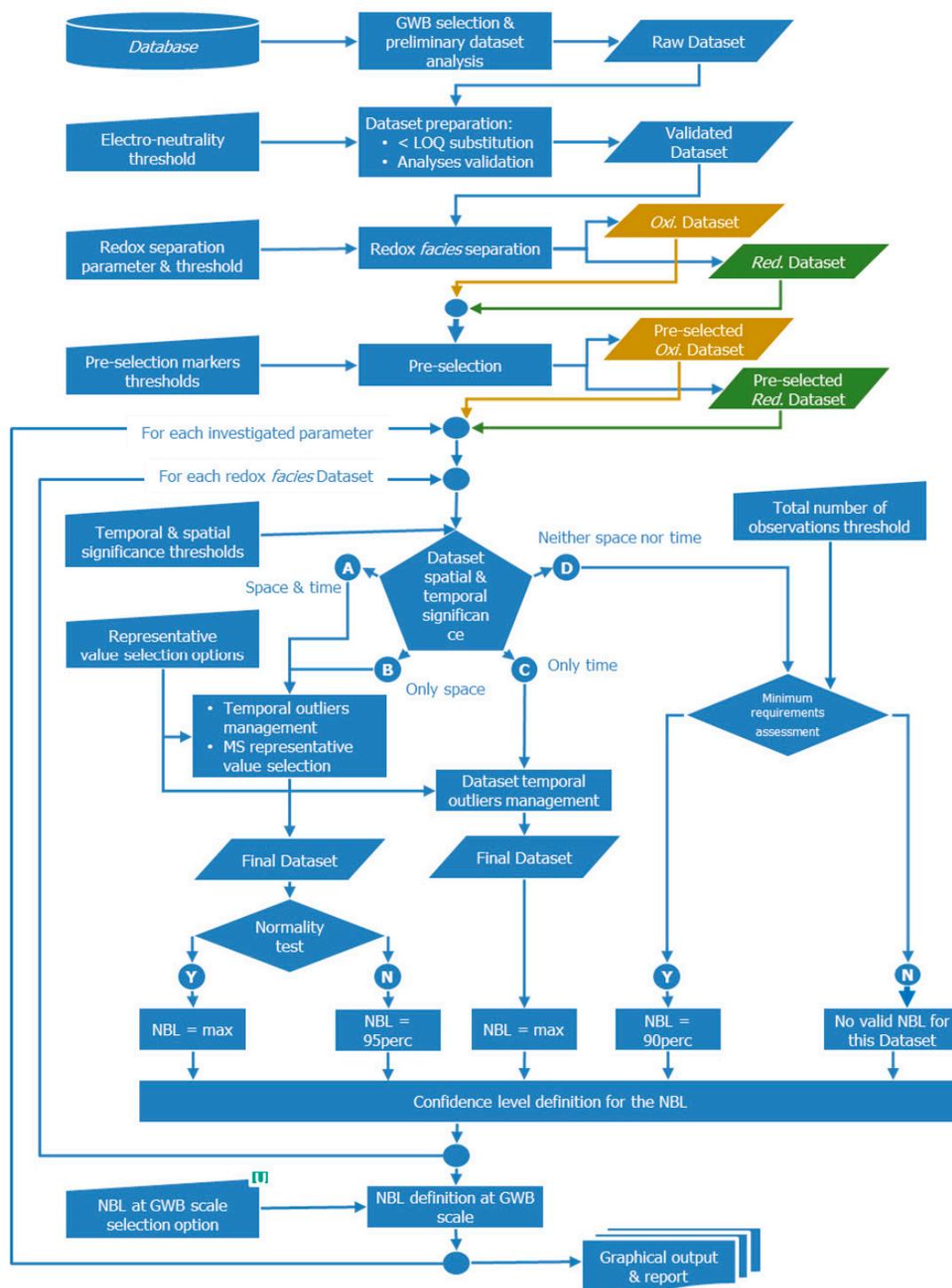
Figure 2. General scheme of the evaluation of natural background levels (eNaBLe) tool.

In particular, it is possible to define:

- The inclusion or not of those stations that, due to their overall characteristics, have been judged unsuitable for NBLs evaluation (checkbox "Use MPs exclusion list").
- A threshold for the control of the electrical balance, typically 5% [30] or 10% [2].
- The appropriate time interval for the data to be processed.
- Parameters to be used for the redox facies separation. The redox potential (ORP) or the dissolved oxygen concentration (DO) can be selected. The predefined limits proposed by the system for these two parameters (e.g., DO 3 mg/L or ORP 100 mV) can be

modified. It is also possible, by deselecting the checkbox, to completely disable the redox facies separation.

- The limits to be applied for the preselection process to nitrates or ammonia in relation to the oxidizing or reducing facies. The system proposes values equal to 75% of the expected quality standards.
- Methodology for the management of the time series in order to calculate the representative values for the monitoring stations. It is possible to select the option of the simple calculation of the median; else an analysis and elimination of the outliers and the choice of the maximum value after the elimination. In this last case, it is also possible to select the method for the identification of the outliers (Huber non-parametric test or boxplot) [31];
- Selection of the method for identifying and eliminating the outliers of the representative values of the MSs (Huber test or boxplot). It is also possible to disable the procedure of outliers elimination.
- Selection of the method of NBLs assessment at GWB scale. If the MSs have been classified based on the redox facies, it is possible to select the highest value or the one with the highest confidence level.

A screenshot of the configuration option selection window is shown in Figure 3.

Figure 3. Selection of the calculation parameters within the eNaBLe tool.

2.1.2. NBLs Assessment and Relative Confidence Levels

The general NBLs assessment procedure can be summarized in the following operational phases:

1. Query to the database to retrieve data of the selected GWB, relative to the selected parameters and to the parameters linked to the facies separation.

2. MSs list filtering using the selected time interval and, possibly, the exclusion list.
3. Processing of the values below the limit of quantification (LOQ). All analytical values reported in the dataset as lower than the LOQ, are replaced with half of the LOQ value.
4. Validation of each sample analysis using the threshold for electrical balance entered in the configuration phase.
5. Redox facies separation using the parameter and threshold selected in the configuration phase.
6. Preselection. As regards the oxidizing facies, the monitoring stations with the median of the nitrate values higher than the selected limit in the configuration phase or with missing values are rejected. For the reducing facies, MSs with the median of ammonia values higher than the selected indicated in the configuration phase or with missing values are rejected. If redox facies separation has been disabled, stations with values greater than the respective limits of nitrates and/or ammonia or stations with missing data will be excluded from the dataset.
7. Analysis of the time series and calculation of the representative values for each MS using the methodology selected during the configuration phase, rejecting, if required, the outliers data.
8. Analysis of the representative values of the MSs using the methodology selected for the management of the outliers and verifying the normality of the resulting dataset distribution using the Shapiro–Wilk test [32].
9. Evaluation of the consistency of the dataset of representative values to identify the assessment path for NBL calculation. In order to distinguish different levels of spatial and temporal consistency of the data available for a given GWB, once the processing described in the previous sections has been concluded, the consistency of the individual datasets is definitively assessed. Therefore, 4 cases are identified:

    - Case A: Sample size adequate to describe the temporal and spatial variability of the parameter for the dataset in question;
    - Case B: Sample size adequate to describe the spatial variability of the parameter for the dataset in question, but not the temporal variability;
    - Case C: Sample size adequate to describe the temporal variability of the parameter for the dataset in question, but not spatial variability;
    - Case D: Sample size inadequate to describe the spatial and temporal variability of the parameter for the dataset in question.

    For sample size adequate to describe, from a purely statistical point of view, the spatial variability of the system in question, it means a minimum of 15 monitoring stations adequately distributed ($N \geq 15$). For sample size adequate to describe, from a purely statistical point of view, the temporal variability of the system in question, it means a minimum of 8 observations ($n \geq 8$) distributed regularly over at least 2 years for each station, over at least 80% of the monitoring stations. These requirements are those listed in the Guidelines and are not modifiable by the tool users.

10. Calculation of NBLs. The NBL is given for datasets of type A, B or C by the maximum value among the representative values of the MSs (in case the dataset shows a normal distribution) or by the 95th percentile of the representative values of the MSs (in case of non-normality of the dataset). For the type D datasets, if the number of total observations available is $\geq$10, the provisional NBL will be given by the 90th percentile of the total available observations. When the total number of observations is less than 10, the calculation will not be carried out and a provisional NBL can be obtained by analogy with other GWBs, or portions of GWBs, characterized by similar conditions in terms of geochemical facies, hydrogeological context and anthropic pressures.
11. Determination of the confidence level to be associated to the calculated NBLs. It was considered appropriate to associate an index (confidence level) to the NBLs, as a function of the size of the statistical sample on which the calculation was based, of

the dimensional (area) and typological (unconfined or confined) characteristics of the GWB (Table 1).

12.   Assignment of a NBL at the groundwater body scale using the methodology selected in the configuration phase.

Steps 6 to 11 are performed by eNaBLe for the different redox facies and steps 7 to 12 for each of the analytical parameters selected.

2.1.3. Output of Results

At the end of the calculation procedures, eNaBLe produces a summary with the configuration options and the results (TV, validated data, minimum and maximum representative values, calculation model and normality of distribution) and the calculated NBLs with relative confidence levels for the investigated parameters, differentiated by redox facies. Finally, the system produces a table which shows the NBLs relative to the entire GWB. If during the configuration phase the separation of redox facies has been deactivated, only the NBLs calculated for the entire GWB are produced.

By the appropriate links contained in the results page, files in CVS format, containing the intermediate and final datasets produced during the data processing, are also accessible.

The tool will finally produce graphical reports of the selected parameters consisting of a table with the main statistical data (minimum and maximum value, mean, median, MAD, 95th percentile and normality of distribution), quantile-quantile plots and the georeferenced spatial distribution of the monitoring stations. A printable PDF files in which are summarized all the configuration parameters and the resulting NBLs, is also available.

2.2. The Mounts Vulsini Groundwater Body

The investigated area extends for about 60 km². It is located on the southern flank of the Mounts Vulsini groundwater body, an unconfined aquifer hosted in the Pleistocene volcanites of the Vulsini volcanic district (Central Italy). Groundwater in the study area flows from N to SW and ESE (Figure 4).
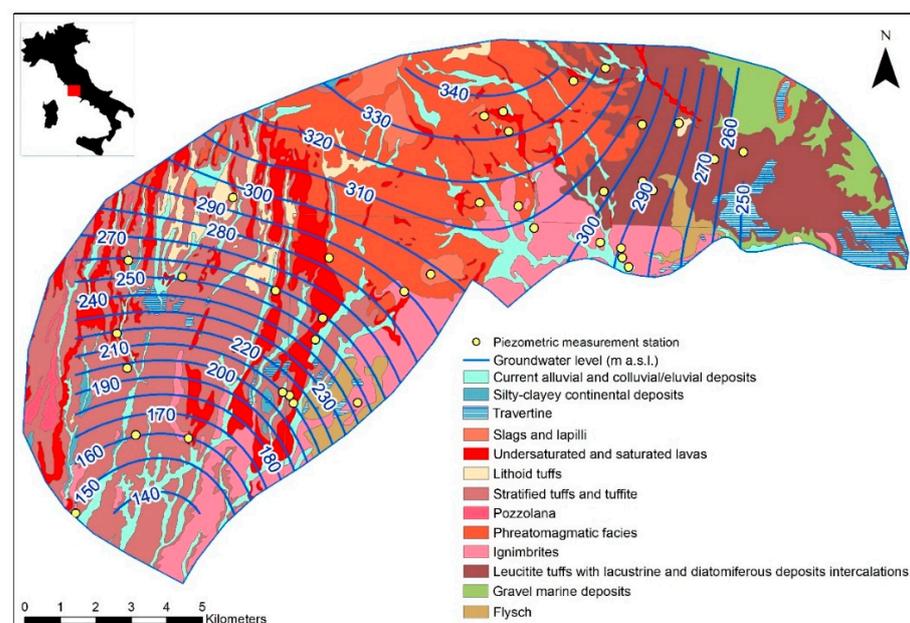

Figure 4. Groundwater level contour map of the southern Mounts Vulsini GWB and outcropping lithologies.

The main anthropogenic pressures are agriculture and animal husbandry. Urban areas and industrial sites including waste management facilities are also present in the area.

Groundwaters are mainly of the alkaline-earth bicarbonate and alkaline bicarbonate type; due to different natural phenomena (water–rock interaction, upwelling of geothermal fluids along fracture/fault systems and presence of mineral deposits), As and F are known to be widespread and co-present in the area [33–40], mainly in oxidizing conditions, with values often higher than the standards set for human consumption by WHO [41] and Directive 98/83/EC. Fe and Mn, on the other hand, are mainly linked to reducing conditions that are found locally, causing the reductive dissolution of their oxi-hydroxides and significant concentrations of the two elements in groundwater.

A total of 50 groundwater samples were collected from private wells in July 2017, 9 of which have been resampled twice more in January and July 2020 to increase the numerosity of the peculiar reducing facies existing in the area. Groundwater samples were collected following standardized sampling protocols [42,43]. Particular attention was paid to the in-line measurement of physical–chemical parameters, whose correct determination is crucial for the definition of the aquifer conceptual model and in the setting of the calculation parameters included in the procedure for NBLs assessment (e.g., DO and ORP accurate measurements for the redox facies separation).

Physical–chemical parameters and chemical data were used to build the starting database for the calculation of NBLs, operated by using the eNaBLe tool.

3. Results

3.1. Configuration and Preliminary Analysis on the Monitoring Stations

The first stage of the procedure includes a series of operations to be applied to the 50 monitoring stations (MSs), all belonging to the Mounts Vulsini groundwater body. All analytical values below the limit of quantification should be replaced with a value equal to half the LOQ. However, none of the parameters of interest for this study (As, F, Fe, Mn) showed values lower than the LOQ. As regards the data validation, the threshold of the electrical balance (5%) did not result in the elimination of any sample (maximum error = 3.8%).

The next step concerns the redox facies separation, for which a DO threshold of 3.0 mg/L was used. This value led to the identification of an oxidizing facies (41 water points), with DO > 3.0 mg/L, largely dominant and widely present throughout the study area, and a reducing facies (9 water points), with DO < 3.0 mg/L, not spatially diffused but rather linked to a few isolated points (Figure 5).

The third and last step of this phase of the procedure is represented by the preselection of the MSs, operated using two different markers of anthropogenic contamination, according to the redox conditions: NO_3^- (<37.5 mg/L) for the oxidizing facies and NH_4^+ (<0.375 mg/L) for the reducing facies. The selected concentration limits correspond to the 75% of the expected quality standards/TVs. As for the oxidizing facies, 15 samples were discarded, so the preselected dataset is composed of 26 MSs useful for the NBLs calculation (Figure 5). No points were discarded for the reducing facies, as all NH_4^+ values are below the chosen threshold. The preselected dataset is therefore composed of the 9 initial stations.

Before continuing with the next parameter-specific phase, it was decided to separately evaluate the correlation between the elements of interest and the redox parameters (DO and ORP), in order to evaluate their sensitivity to the redox conditions. The Pearson parametric and Spearman non-parametric correlation indexes (Table 2) show a significant negative correlation with the redox parameters for Fe and Mn (mutually proportional), while As and F are well correlated with each other but not redox sensitive elements.

Consequently, in the NBLs assessment we have decided to proceed in a diversified way, keeping the two redox facies separate for Fe and Mn and instead defining a single value, relative to the entire preselected dataset (35 water points), for As and F, not affected by the redox conditions.
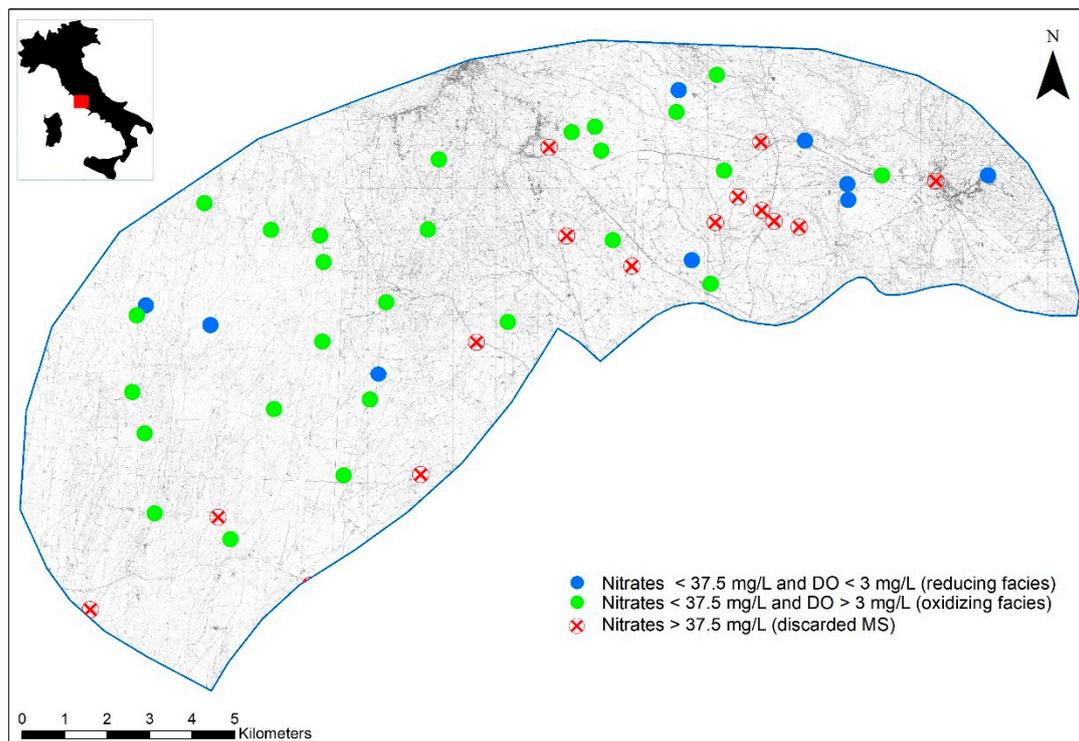
Figure 5. Redox facies subdivision based on dissolved oxygen concentration (DO) concentration and preselected MSs.

Table 2. Pearson (r) and Spearman (r_s) correlation coefficients for the redox parameters and the investigated elements. Statistically significant values in bold ($\alpha = 0.01$).

	Pearson (r)					
DO		**0.58**	−0.25	**−0.72**	**−0.69**	−0.15
ORP	**0.43**		−0.39	**−0.45**	**−0.42**	−0.21
F	−0.21	−0.29		0.18	0.21	**0.51**
Mn	**−0.66**	**−0.48**	0.16		**0.99**	0.13
Fe	**−0.53**	**−0.42**	0.28	**0.76**		0.11
As	−0.04	−0.09	**0.59**	−0.12	0.04	
	DO	**ORP**	**F**	**Mn**	**Fe**	**As**
			Spearman (r_s)			

3.2. NBLs Calculation for As and F

For the 26 stations of the oxidizing facies, the available data relate to a single sampling (2017), for which only the spatial analysis was carried out. On the other hand, as regards the 9 water points in reducing facies, the dataset also includes the samples of January and July 2020. For the two parameters, a temporal processing was therefore carried out, defining a representative value for each station, given by the median of the values measured in the three campaigns. Consequently, the As and F dataset used for the spatial analysis consists of 35 data, of which 26 individual values of the oxidizing facies and 9 median values of the reducing facies.

For both parameters, the presence of anomalous data was then analyzed using the Huber's non-parametric test, which identified two outliers for the F and one for the As. In the current state of knowledge, however, there are no valid scientific reasons to exclude these samples, which were therefore included in the final dataset for the definition of NBLs, consisting of 35 data. It is therefore the case B foreseen in the Guidelines, in which there is a significant spatial but not temporal dimension.

The Shapiro–Wilk test was then applied to verify the normality of data, showing that only F follows a Gaussian distribution. Consequently, according to the Guidelines, the NBL

for F is given by the maximum of the statistical sample (3.56 mg/L), while for As it is equal to 95th percentile of the same statistical sample (20.5 µg/L).

Finally, considering the aquifer type (unconfined), its extension (60 km^2) and the number of total samples (35), for both parameters it was possible to associate a high confidence level to the defined NBLs.

3.3. NBLs Calculation for Fe and Mn (Oxidizing Facies)

As previously observed, for the 26 MSs in oxidizing facies, the available data belongs to a single sampling survey (2017), therefore only the spatial analysis was carried out. Huber's test found 2 outliers for Fe and 4 for Mn. However, also in this case there any scientific elements were identified to discard these data, hence the final dataset for the definition of NBLs remains the original one of 26 data. This is again the case B described in the Guidelines, in which there is a significant spatial but not a temporal dimension.

As expected, the subsequent Shapiro–Wilk test showed how the distributions of the two elements are far from normal, so in both cases the NBL should be set to the 95th percentile of the statistical sample (105.1 µg/L for Fe and 9.1 µg/L for Mn) and for both parameters a high confidence level was associated to the calculated NBL.

3.4. NBLs Calculation for Fe and Mn (Reducing Facies)

As regards the reducing facies, the dataset consists of only 9 water points, corresponding to the case D indicated in the Guidelines, the worst in terms of available information, in which neither a spatial nor a significant temporal dimension is reached. In these situations, the Guidelines suggest to estimate a provisional NBL, combining all available observations (in space and time). For Fe and Mn the total observations available are 26. Huber's test did not highlight possible outliers, so the 90th percentile of the observations for both parameters was calculated. The estimated NBLs are equal to 7364.7 µg/L for Fe and 805.0 µg/L for Mn. The associated confidence level is low and new monitoring observations are needed to improve the reliability of the dataset.

The calculated NBLs and the associated confidence levels are shown in Table 3. As indicated in the Guidelines, the NBL for each parameter was expressed with the same unit of measurement and rounded up with the same number of decimals as the relative limit set by Decree 152/2006 [44]. At present, as indicated by the Guidelines, for Fe and Mn the NBL for the groundwater body will correspond to that defined for the oxidizing facies, which has the highest confidence level.

Table 3. Calculated NBLs and associated confidence levels.

	F (mg/L)	Mn (µg/L)	Fe (µg/L)	As (µg/L)	Confidence Level
Preselected Dataset	3.6	-	-	21	High
Oxidizing Facies	-	9	105	-	High
Reducing Facies	-	805	7365	-	Low

4. Discussion

The study on the natural background values for the southern Mounts Vulsini groundwater body confirmed the presence of high concentrations of As and F, which show numerous exceedances of the TVs set at the national level by Decree 30/2009 (1.5 mg/L and 10 µg/L) [45] and are naturally present in groundwater. The relatively low range of concentration, the existence of single data populations and the normal (for fluoride) or close to normal (for arsenic) distributions suggest that the presence of the two elements in groundwater can be largely attributed to the water–rock interaction processes within the volcanic aquifer. The statistical analysis does not highlight further overlapping phenomena and this also translates into the definition of a particularly reliable NBL for the investigated groundwater body. Fe and Mn show a wide range of concentrations, depending on the redox conditions found in the aquifer. In the most common oxidizing conditions, the

natural concentrations are generally low, while in the reducing conditions that can be found locally, the values are considerably higher and well above the limits set by Decree 152/2006 for the assessment of pollution while monitoring impacts on groundwater of e.g., industrial activities. The nature of these few peculiar points should be further investigated, in order to exclude any contribution of anthropogenic origin (currently not identifiable and quantifiable). Unlike the oxidizing facies, the few points belonging to the reducing facies do not show a particular spatial distribution within the GWB. This results in a geochemical population that can hardly be better characterized, due to the difficulty in finding other equally peculiar sampling points, which would make the dataset statistically more robust. The reliability of the NBL in this case can be improved mainly by monitoring these MSs, thus increasing the number of observations over time.

During the application of the Italian Guidelines for the NBLs assessment, some critical issues emerged, partly associated with the Guidelines themselves, others more specific of the online tool.

With regard to the first phase of the procedure, MSs specific, the software allows to perform the separation of the redox facies or alternatively to keep a single dataset. However, this choice cannot be applied in a different way depending on the parameters whose NBL is to be determined. Consequently, the procedure must be repeated twice, for the total dataset and for the separate facies.

Moreover, it is not clear in the Guidelines whether a temporal analysis of the pre-election markers (nitrates/ammonia) should be done. Currently, eNaBLe calculates the median of the temporal data and if this is < the selected limit (e.g., 37.5 mg/L for nitrates or 0.375 mg/L for ammonia), the MS is considered useful for the calculation of NBLs; otherwise, it is discarded. However, this entails the risk of considering stations that have exceedances of these limits. In addition, since an analysis of temporal trends is not envisaged, stations that show ascending temporal trends of the markers, suggesting contamination in progress, could be included in the preselected dataset. In this regard, an official Guideline including different statistical methods for trend analysis, estimating concentration scenarios and identification of trend reversal, have been published in Italy in 2017 [46]. The use of these techniques, currently to be applied externally to the tool, could be helpful to integrate the temporal analysis of data, in particular for substances of clear anthropogenic origin such as nitrates. About this, Frollini et al. [47] have recently applied a slightly modified version of the Guideline to a groundwater body in Northern Italy featuring nitrate pollution, discussing its advantages and limitations.

Furthermore, the software does not currently allow to evaluate the correlation between chemical elements and redox parameters or even other chemical-physical parameters; it is therefore not possible to statistically evaluate whether they are redox-sensitive (operation currently executable only externally to the tool) and make a justified choice of the facies separation.

As regards the second phase of the procedure, parameter-specific, the software evaluates the presence of data outliers through application of the non-parametric Huber test (or alternatively, extrapolating them from a simple boxplot). This is clearly a simplification of the procedure, as non-parametric methods do not require any knowledge or assumptions about the form of data distribution. These are robust techniques, as they are applicable to any situation. However, the tool could also provide for the possibility of applying parametric tests (e.g., Rosner test), that are statistical procedures based on the assumption of normality of the data, more powerful and preferable to non-parametric tests, in particular when the data follow the hypothesized distribution. In the case study, for example, the Huber test identifies 2 outliers for F. However, the distribution of data is normal, both including and excluding the outliers, and the Rosner test [48], recommended for Gaussian distributions with more than 25 data, applied externally to the tool, does not detect anomalous values. Hence in the definition of the NBL we have included the outliers, since their exclusion is not supported by evident scientific reasons. However, the adoption of a parametric test, in this case, would have simplified the path of definition of the NBL.

Furthermore, again regarding the statistical study, it is limited only to the evaluation of the normality of the data and the presence of anomalous values, which can be correctly discarded or not. The evaluation of the existence of multiple populations, which would lead to a subdivision of the dataset before calculating the NBLs as required by the Guidelines, should be conducted aside. eNaBLe shows the distribution of the data at the end of the procedure, through Q-Q plot, but any subdivision of the dataset before the calculation is not allowed. In this regard, the possibility of implementing further partitions of the dataset, based on the study of the conceptual model of the GWB and conducted outside the tool, is currently being evaluated.

About the temporal analysis of data, currently the software allows to calculate a representative value for each monitoring station through the calculation of the median or, alternatively, through the elimination of any outliers and using the maximum value of the residual distribution. Therefore, at the moment the evaluation of temporal trends (step foreseen in the Guidelines), which could lead to the elimination of MSs that do not show outliers but simply increasing trends for certain parameters, suggesting a possible contamination in progress, should be performed aside. In the direction of improving the tool in this part, the implementation of a statistical test for the estimation of the slope of ascending trends (e.g., Mann–Kendall test) [49,50] is planned.

Finally, for the cases in which there is neither spatial nor temporal significance (Case D), without further specification by the Guidelines, eNaBLe evaluates the outliers putting together all the observations. Indeed, Case D refers to a GWB with a small number of MSs (to the limit of one only MS); therefore, if we reduce the multiple time observations to one representative value (median) for each station, we could not perform a robust statistical evaluation of the outliers.

Following the Guidelines, for each parameter only one NBL can be defined even in presence of multiple datasets, e.g., when different geochemical facies or redox conditions have been recognized and evaluated separately in the GWB. The Guidelines suggest choosing the highest value among those of the single datasets, only in case D will the NBL with the highest confidence level be selected. At present, therefore, the NBLs for Fe and Mn assigned to the entire GWB would be those, rather low, specific to the oxidizing facies. However, it is apparent that the high concentrations of these metals are to be associated with the peculiar reducing conditions that are found locally within the GWB. In order to assign them a NBL more representative of the conditions which promote their presence in solution, it is necessary to continue the temporal monitoring until reaching case C (significant temporal dimension), or to increase the number of MSs until reaching case B (significant spatial dimension). Only when the confidence levels are all the same, it will be possible to select the NBL given by the maximum value among the different datasets, therefore the one (clearly higher) associated with the reducing facies.

5. Conclusions

eNaBLe is a versatile and user-friendly instrument, meant to facilitate the assessment of the NBLs following the national Guidelines. It performs automatically the sequence of operations as indicated by the Guidelines, easily allowing a very quick calculation of NBLs at the groundwater body scale, even in the presence of a limited amount of data. By setting a few configuration parameters, an assessment is rapidly reached which appears sufficiently representative of the natural state of groundwater. However, it should be noted that these settings must be made in a reasoned way, starting from a thorough knowledge of the conceptual model of the aquifer. In particular, the redox facies separation assumes that redox conditions of groundwater are known. Its determination is commonly based on physical-chemical indicators such as DO or ORP whose measure is often critical. Further, this information is frequently missing or scarcely reliable in datasets of groundwater quality monitoring. The validation of the data and their organization into coherent datasets are also fundamental to obtain an evaluation that is as significant as possible.

Some critical issues have emerged in the statistical analysis path, even in part deriving from the reference Guidelines, but the software can be improved and refined in every its part starting from the indications and criticalities deriving from these first applications. The tool is still under testing for a thorough verification of all the possible variants in dataset structures and GWB conceptual models, also taking into account the suggestions of the stakeholders.

Author Contributions: Conceptualization, D.P., E.F., E.P. and S.G.; methodology, D.P., E.P. and S.G.; software, S.G.; validation, D.P.; formal analysis, D.P.; investigation, D.P., E.F., E.P. and S.G.; data curation, S.G.; writing—original draft preparation, D.P.; writing—review and editing, D.P., E.F., E.P. and S.G.; visualization, D.P. and E.F.; supervision, E.P. and S.G. All authors have read and agreed to the published version of the manuscript.

References

1. GWD. *Groundwater Directive 2006/118/CE, Directive of the European Parliament and of the Council on the Protection of Groundwater Against Pollution and Deterioration*; Official Journal of European Union L372; European Union: Luxembourg, 2006; pp. 19–31.
2. Muller, D.; Blum, A.; Hart, A.; Hookey, J.; Kunkel, R.; Scheidleder, A.; Tomlin, C.; Wendland, F. Final proposal for a methodology to set up groundwater threshold values in Europe. In *Report to the EU project "BRIDGE"*; 2006; Deliverable D18. Available online: http://nfp-at.eionet.europa.eu/Public/irc/eionet-circle/bridge/library?l=/deliverables/bridge_groundw-205pdf/_EN_1.0_&a=d (accessed on 30 December 2020).
3. Hinsby, K.; de Melo, M.T.C.; Dahl, M. 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.* **2008**, *401*, 1–20. [CrossRef] [PubMed]
4. Ducci, D.; de Melo, M.T.C.; Preziosi, E.; Sellerino, M.; Parrone, D.; Ribeiro, L. Combining natural background levels (NBLs) assessment with indicator kriging analysis to improve groundwater quality data interpretation and management. *Sci. Total Environ.* **2016**, *569–570*, 569–584. [CrossRef] [PubMed]
5. Griffioen, J.; Passier, H.F.; Klein, J. Comparison of selection methods to deduce natural background levels for groundwater units. *Environ. Sci. Technol.* **2008**, *42*, 4863–4869. [CrossRef] [PubMed]
6. Marandi, A.; Karro, E. Natural background levels and threshold values of monitored parameters in the Cambrian–Vendian groundwater body, Estonia. *Environ. Geol.* **2008**, *54*, 1217–1225. [CrossRef]
7. Wendland, F.; Berthold, G.; Blum, A.; Elsass, P.; Fritsche, J.G.; Kunkel, R.; Wolter, R. Derivation of natural background levels and threshold values for groundwater in the Upper Rhine Valley (France, Switzerland and Germany). *Desalination* **2008**, *226*, 160–168. [CrossRef]
8. European Commission. *Guidance on Groundwater Status and Trend Assessment, Guidance Document no 18*; Technical Report; European Communities: Luxembourg, 2009; ISBN 978-92-79-11374-1.
9. Gemitzi, A. Evaluating the anthropogenic impacts on groundwaters; a methodology based on the determination of natural background levels and threshold values. *Environ. Earth Sci.* **2012**, *67*, 2223–2237. [CrossRef]
10. Preziosi, E.; Giuliano, G.; Vivona, R. Natural background levels and threshold values derivation for naturally As, V and F rich groundwater bodies: A methodological case study in Central Italy. *Environ. Earth Sci.* **2010**, *61*, 885–897. [CrossRef]
11. Molinari, A.; Guadagnini, L.; Marcaccio, M.; Guadagnini, A. Natural background levels and threshold values of chemical species in three large-scale groundwater bodies in Northern Italy. *Sci. Total Environ.* **2012**, *425*, 9–19. [CrossRef]
12. Rotiroti, M.; Fumagalli, L. Derivation of preliminary natural background levels for naturally Mn, Fe, As and NH4 rich groundwater: The case study of Cremona area (Northern Italy). *Rend. Online Soc. Geol. Ital.* **2013**, *24*, 284–286.
13. Hawkes, H.E.; Webb, J.S. *Geochemistry in Mineral Exploration*; Harper & Row: New York, NY, USA, 1962; p. 415.
14. Tukey, J.W. *Exploratory Data Analysis*; Addison-Wesley Publishing Company: Boston, MA, USA, 1977; p. 688.
15. Reimann, C.; Filzmoser, P.; Garrett, G. Background and threshold: Critical comparison of methods of determination. *Sci. Total Environ.* **2005**, *346*, 1–16. [CrossRef]
16. Wendland, F.; Hannappel, S.; Kunkel, R.; Schenk, R.; Voigt, H.J.; Wolter, R. A procedure to define natural groundwater conditions of groundwater bodies in Germany. *Water Sci. Technol.* **2005**, *51*, 249–257. [CrossRef] [PubMed]

17. Edmunds, W.M.; Shand, P.; Hart, P.; Ward, R.S. The natural (baseline) quality of groundwater: A UK pilot study. *Sci. Total Environ.* **2003**, *310*, 25–35. [CrossRef]
18. Panno, S.V.; Kelly, W.R.; Martinsek, A.T.; Hackley, K.C. Estimating background and threshold nitrate concentrations using probability graphs. *Ground Water* **2006**, *44*, 697–709. [CrossRef] [PubMed]
19. Edmunds, W.M.; Shand, P. *Natural Groundwater Quality*; Wiley-Blackwell: Hoboken, NJ, USA, 2008; p. 488. ISBN 978-14051-5675-2.
20. Walter, T. Determining natural background values with probability plots. In Proceedings of the EU Groundwater Policy Developments Conference, UNESCO, Paris, France, 13–15 November 2008.
21. Preziosi, E.; Parrone, D.; Del Bon, A.; Ghergo, S. Natural background level assessment in groundwaters: Probability plot versus pre-selection method. *J. Geochem. Explor.* **2014**, *143*, 43–53. [CrossRef]
22. ISPRA—Istituto Superiore per la Protezione e la Ricerca Ambientale. *Linee Guida Recanti la Procedura da Seguire per il Calcolo dei Valori di Fondo Naturale per i Corpi Idrici Sotterranei (DM 6 luglio 2016)*; Manuali e Linee Guida 155/2017; ISPRA: Rome, Italy, 2017; ISBN 978-88-448-0830-3. Available online: https://www.isprambiente.gov.it/it/pubblicazioni/manuali-e-linee-guida/linee-guida-recanti-la-procedura-da-seguire-per-il-calcolo-dei-valori-di-fondo-per-i-corpi-idrici-sotterranei-dm-6-luglio-2016 (accessed on 26 November 2020).
23. Frollini, E.; Parrone, D.; Ghergo, S.; Preziosi, E. Natural background level assessment in groundwater: Application of the Italian national guideline; Geophysical Research Abstracts Vol.20, EGU2018-15674-1. In Proceedings of the EGU General Assembly 2018, Vienna, Austria, 8–13 April 2018. eISSN 1607-7962.
24. Parrone, D.; Ghergo, S.; Preziosi, E. A multi-method approach for the assessment of natural background levels in groundwater. *Sci. Total Environ.* **2019**, *659*, 884–894. [CrossRef] [PubMed]
25. Chidichimo, F.; De Biase, M.; Costabile, A.; Cuiuli, E.; Reillo, O.; Migliorino, C.; Treccosti, I.; Straface, S. GuEstNBL: The software for the guided estimation of the natural background levels of the aquifers. *Water* **2020**, *12*, 2728. [CrossRef]
26. Ghergo, S.; Parrone, D.; Frollini, E.; Preziosi, E. eNaBLe, an on-line tool to evaluate Natural Background Levels in groundwater following the Italian Guidelines. Geophysical Research Abstracts Vol.21, EGU2019-7443. In Proceedings of the EGU General Assembly 2019, Vienna, Austria, 7–12 April 2019. eISSN 1607-7962.
27. HP: Hypertext Preprocessor. Available online: https://www.php.net (accessed on 30 December 2020).
28. MySQL. Available online: https://www.mysql.com (accessed on 30 December 2020).
29. Water Resources Database. Available online: http://www.irsa.cnr.it/WRdB (accessed on 30 December 2020).
30. Appelo, C.A.J.; Postma, D. *Geochemistry, Groundwater and Pollution*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2005; p. 668.
31. Huber, P.J. *Robust Statistics*; John Wiley: New York, NY, USA, 1981; p. 320.
32. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]
33. Angelone, M.; Cremisini, C.; Piscopo, V.; Proposito, M.; Spaziani, F. Influence of hydrostratigraphy and structural setting on the arsenic occurrence in groundwater of the Cimino-Vico volcanic area (Central Italy). *Hydrogeol. J.* **2009**, *17*, 901–914. [CrossRef]
34. Baiocchi, A.; Coletta, A.; Espositi, L.; Lotti, F.; Piscopo, V. Sustainable groundwater development in a naturally arsenic-contaminated aquifer: The case of the Cimino-Vico volcanic area (central Italy). *Ital. J. Eng. Geol. Environ.* **2013**, *1*, 5–18.
35. Dall'Aglio, M.; Giuliano, G.; Amicizia, D.; Andrenelli, M.C.; Cicioni, G.B.; Mastroianni, D.; Sepicacchi, L.; Tersigni, S. Assessing drinking water quality in Northern Latium by trace elements analysis. In *Proceedings of Water Rock Interaction (WRI-10), Villasimius, Italy, 10–15 June 2001*; Cidu, R., Ed.; A.A. Balkema: Exton, PA, USA, 2001; Volume 2, pp. 1063–1066.
36. De Rita, D.; Cremisini, C.; Cinnirella, A.; Spaziano, F. Fluorine in the rocks and sediments of volcanic area in central Italy: Total content, enrichment and leaching process and a hypothesis on the vulnerability of the related aquifers. *Environ. Monit. Assess.* **2012**, *184*, 5781–5796. [CrossRef]
37. Parrone, D.; Ghergo, S.; Frollini, E.; Rossi, D.; Preziosi, E. Arsenic-fluoride co-contamination in groundwater: Background and anomalies in a volcanic-sedimentary aquifer in central Italy. *J. Geochem. Explor.* **2020**, *217*, 106590. [CrossRef]
38. Preziosi, E.; Vivona, R.; Giuliano, G. Microinquinanti di origine naturale negli acquiferi vulcanici: Un approccio integrato quantitativo e qualitativo nel Lazio settentrionale [Micropollutants of natural origin in volcanic aquifers: An integrated quantitative and qualitative approach in northern Latium]. *IGEA* **2005**, *20*, 3–14.
39. Vivona, R.; Preziosi, E.; Giuliano, G.; Mastroianni, D.; Falconi, F.; Madé, B. Geochemical characterization of a volcanic-sedimentary aquifer in Central Italy. In *Proceedings of the 11th International Symposium on Water-Rock Interaction WRI-11*; Wanty, R., Seal, R.I.I., Eds.; Saratoga Springs: New York, NY, USA, 2004; pp. 513–517.
40. Vivona, R.; Preziosi, E.; Madé, B.; Giuliano, G. Occurrence of minor toxic elements in volcanic-sedimentary aquifers: A case study in central Italy. *Hydrogeol. J.* **2007**, *15*, 1183–1196. [CrossRef]
41. WHO. *Guidelines for Drinking-Water Quality*, 4th ed.; World Health Organization: Geneva, Switzerland, 2011; p. 564. ISBN 978-92-4-154815-1.
42. Preziosi, E.; Ghergo, S.; Frollini, E.; Parrone, D. Buone Pratiche per il Campionamento Delle Acque Sotterranee: Proposta di un Protocollo. (Best Practices for Groundwater Sampling: Proposal of a Protocol). Notiziario dei Metodi Analitici IRSA-CNR. 2017, Volume 1, pp. 23–36. Available online: http://www.irsa.cnr.it/images/docs/Notiz/notiz2017_1.pdf (accessed on 26 June 2018).
43. Frollini, E.; Rossi, D.; Rainaldi, M.; Parrone, D.; Ghergo, S.; Preziosi, E. A proposal for groundwater sampling guidelines: Application to a case study in southern Latium. Rend. *Online Soc. Geol. It.* **2019**, *47*, 46–51. [CrossRef]
44. *Decreto Legislativo 3 Aprile 2006, n. 152. Norme in Materia Ambientale. Gazzetta Ufficiale n. 88 del 14-04-2006*; Istituto Poligrafico e Zecca dello Stato s.p.a.: Rome, Italy, 2006.

45.   *Decreto Legislativo 16 Marzo 2009, n. 30. Attuazione Della Direttiva 2006/118/CE, Relativa Alla Protezione Delle Acque Sotterranee Dall'inquinamento e dal Deterioramento. Gazzetta Ufficiale n. 79 del 4-4-2009*; Istituto Poligrafico e Zecca dello Stato s.p.a.: Rome, Italy, 2009.

46.   ISPRA-CNR.IRSA. *Linee Guida per la Valutazione Delle Tendenze Ascendenti e D'inversione Degli Inquinanti nelle Acque Sotterranee (DM 6 Luglio 2016)*; Manuali e Linee Guida 161/2017; ISPRA: Rome, Italy, 2017; ISBN 978-88-448-0844-0.

47.   Frollini, E.; Preziosi, E.; Calace, N.; Guerra, M.; Guyennon, N.; Marcaccio, M.; Menichetti, S.; Romano, E.; Ghergo, S. Groundwater quality trend and trend reversal assessment in the European Water Framework Directive context: An example with nitrates in Italy. *Environ. Sci. Pollut. Res.*. In press. [CrossRef]

48.   Rosner, B. Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics* **1983**, *25*, 165–172. [CrossRef]

49.   Mann, H.B. Nonparametric tests against trend. *Econometrica* **1945**, *13*, 245–259. [CrossRef]

50.   Kendall, M.G. *Rank Correlation Methods*; Grin: London, UK, 1975.