GuEstNBL: The Software for the Guided Estimation of the Natural Background Levels of the Aquifers

Francesco Chidichimo 1,*, Michele De Biase 1, Alessandra Costabile 2, Enzo Cuiuli 3, Orsola Reillo 2, Clemente Migliorino 3, Ilario Treccosti 2 and Salvatore Straface 1

1 Department of Environmental Engineering, University of Calabria, via P. Bucci 42B, 87036 Rende, Italy; michele.debiase@unical.it (M.D.B.); salvatore.straface@unical.it (S.S.)
2 Department of Environmental Policy, Calabria Region, Cittadella Regionale, Località Germaneto, 88100 Catanzaro, Italy; alessandra.costabile@regione.calabria.it (A.C.); o.reillo@regione.calabria.it (O.R.); i.treccosti@regione.calabria.it (I.T.)
3 Regional Agency for the Protection of the Environment—ARPACAL, via Lungomare (Loc Mosca) (Giovino), 88100 Catanzaro, Italy; e.cuiuli@arpacal.it (E.C.); c.migliorino@arpacal.it (C.M.)
* Correspondence: francesco.chidichimo@unical.it

Received: 28 August 2020; Accepted: 28 September 2020; Published: 29 September 2020

Abstract: Natural background levels (NBLs) for targeted chemical elements characterize a specific groundwater body, the knowledge of which represents a fundamental information for environmental agencies responsible for the protection, management, and remediation of territory. The large number of areas subject to strong anthropogenic pressures of a different nature and magnitude makes the job of control authorities particularly difficult. The process to distinguish effective anthropogenic contamination from natural conditions and to define realistic environmental clean-up goals goes through the computation of several mutually dependent statistical methods, some of which have non-trivial resolution and interpretation. In this study, we presented a new tool designed to drive those working in the sector into an articulated path towards NBL assessment. The application software was developed in order to read environmental input data provided by a user-friendly web-based geographic information system (GIS) and to return the NBL estimate of a given chemical element following a wizard that allows for the implementation of two methodologies, i.e., component separation or pre-selection. The project was born from a collaboration between the Department of Environmental Engineering of the University of Calabria and the Department of Environmental Policies of the Calabria Region. The software was used to estimate NBLs in selected chemical species at potentially contaminated industrial sites located in Lamezia Terme, Italy. In the future, the developed calculation program will be the official evaluation tool of the Calabria Region for identifying groundwater thresholds.

Keywords: natural background levels; software implementation; parameters estimation; statistical methods; component separation method; pre-selection method

1. Introduction

Groundwater characterization activities may be marked by the observation of large concentration values related to an aquifer’s petrographical composition rather than the pollution status of an investigated site [1]. The reliable quantification of the actual natural contribution to detected concentrations, at a field scale, is underlined by the European Union (EU) Water Framework Directive (WFD 2000/60/EC, article 17), which identified significant and feasible clean-up goals. Applying remediation strategies based on compliance levels guided by current regulations might lead, in some cases, to ineffective and unaffordable cleaning targets for areas where specific natural conditions take place. In such a context,
estimating natural background levels (NBLs) in groundwater gained large attention in the last decade. The Ground Water Daughter Directive (GWDD 2006/118/EC) defines the NBL as “the concentration of a substance or the value of an indicator in a body of groundwater corresponding to no, or only very minor, anthropogenic alterations to undisturbed conditions”. Chemical and biological processes taking place in the water-rock system, as well as intakes from other water bodies and rainfall contribution, may cause local effects to give rise to the spread of background concentrations [2–6]. The correct distinction between anthropogenic and natural origin contamination is mandatory for the estimation of reliable NBLs. When this last condition is not met, the misleading classification of a site as potentially highly contaminated can be assessed by detecting large concentrations. The legal implications of such a conclusion have direct consequences, especially in cases where the possible negative effects of human activities occur in correspondence with polluting areas (industrial sites, landfills, intensive agriculture, farming systems, etc.). Optimizing remediation strategies occurs via proper NBL evaluation and the associated revision of the compliance values that can be locally modified to account for a specific context under investigation. Practical applications, based on the statistical analysis of monitored data also proposed by the EU research project BRIDGE (2007) (background criteria for the identification of groundwater thresholds [7]) have been frequently employed for NBL estimation and are available in the literature for several countries such as Germany, Italy, Spain, and South Korea [2–6,8–17].

Recently, the Italian Institute for Environmental Protection and Research, ISPRA, has promoted initiatives aimed at addressing the methodological aspects related to determining NBLs of soil and groundwater, as well as to draw up guidelines to be followed at the national level. This target was pursued by joining the experiences and skills of the Regional Environmental Protection Agencies (ARPA). the Resolution of the Council of the National System for the Protection of the Environment (SNPA), dated 14/11/2017, laid the foundations for drafting the aforementioned guidelines collected in the manual 174/2018 [18]. the procedural iter, required by this document, can be extremely complex, especially for operators without a high level of knowledge and experience in statistical calculation or iterative optimization processes.

In this study a software was presented to estimate the NBLs in aquifers, complying with indications provided by ISPRA guidelines, including an additional feature for the component separation methodology, which is not covered by the guidelines. This project was born from a collaboration between the Department of Environmental Engineering of the University of Calabria and the Department of Environmental Policies of the Calabria Region, and aimed at providing the public bodies responsible for environmental control with a tool to simplify and automate complex calculations and estimation processes. the software was able to manage environmental data (i.e., concentration of chemical elements were analyzed in the sampling points and collected into a user-friendly web-based geographic information system (GIS)) and to provide the estimate of the background values following different methodologies indicated by the guidelines. the wizard, implemented on Java Platform, was equipped with a step-by-step graphical interface that guided the operator by showing the results obtained for each elaboration, and by allowing him or her to make choices based on his/her experience. The data was analyzed after implementing two methodologies: component separation and pre-selection [7]. At the end of the analysis, performed for the single chemical element, the software printed a PDF report where the used environmental data, the adopted methods, the pursued decisions, the generated graphs/tables, and the estimated NBL were sequentially recorded. in the future, the developed calculation program will be the official evaluation tool of the Calabria Region to identify groundwater thresholds. the work also presented a case study in which the software was used to estimate the NBL of a selected chemical species in a potentially contaminated industrial site located in Lamezia Terme, Italy. To the best of our knowledge, the computer system presented in this work is the first of its kind. It could be easily exported and used in international contexts since the methodological approaches included therein are widely recognized by the whole scientific community.
2. Materials and Methods

This section is devoted to the detailed description of the internal structure of the software-based procedure, i.e., how it operates to reach the NBLs estimation starting from the environmental data, as well as an overview of the case study.

Figure 1. Scheme of the procedure implemented to estimate the natural background levels (NBLs) of aquifers. The blocks numbering (bold characters at the top right of each box) having the CS prefix refers to the procedures included into the component separation method, while the PS prefix is related to the pre-selection method. Other terms appearing in the scheme are: Web-GIS which stands for online geographical information system; PDF which stands for probability density function; CDF which stands for cumulative distribution function; bold letters A, B, C, D indicating four different cases in which the dataset can fall; 95p which stands for 95th percentile.
2.1. Structure of the Software-Based NBL Determination Procedure

Figure 1 shows the overall scheme of the procedure, which collects all aspects planned to guide the operator to determine the NBL for selected chemical species. The action depicted in each numbered block is expanded in the following dedicated subsections.

2.1.1. Data Scheduling and Acquisition

- Block 1 (Conceptual model of the site)

The definition of the conceptual model of the site is a fundamental step for any NBL estimation procedure. It is not an internal gear of the implemented procedure, since its definition is above the latter and it is included in the scheme for the sake of completeness. The conceptual model is aimed at identifying the factors (sources and processes) that determine the distribution, in space and time, of the parameters of interest. It constitutes the cognitive framework of the area. It contains interpretative and relational elements that allow for the understanding of the processes at play in the site in relation to the presence of the targeted substances. The conceptual model guides and supports some choices during the NBL determination procedure (e.g., the data grouping, the exclusion of specific observations, the identification of the most suitable statistical indicator, etc.). A good formulation of the conceptual model cannot exclude information on the geological, geochemical, and hydrogeological nature of the investigated environmental matrices, and on the anthropic pressures that, in the past or present, have impacted the study area. All these details are necessary to ensure that the analyzed data are from homogeneous environmental horizons.

- Block 2 (Dataset organization (Web-GIS))

This is the first step of the procedure falling within the guided system. It relies on a web-based geographic information system setup by the Department of Environmental Policies of the Calabria Region to contain the environmental data of the Water Protection Plan. This facility, providing a validated and constantly updated database, was equipped with specific additional accessories in order to select the data, follow appropriate screening criteria, and make them readily portable into the NBL estimation procedure. In particular, the system is designed to display the regional map of the monitoring wells, each of which contains the time series of the chemical analysis carried out on the collected samples. The system can be initially queried by entering specific spatial filters (e.g., provincial or municipal limits) in order to have a first rough delineation of the investigation area together with the included sampling stations. Subsequently, the knowledge gained from the conceptual model of the area, together with the use of the Web-GIS overlapping layers, containing the geological information of the site, or the land use map (industrial zones location, urban or agricultural areas, dumps position, etc.) allows to further circumscribe the monitoring points falling within portions of territory with homogeneous characteristics. This operation can be done by means of a “lasso tool” with which freehand selected areas are drawn on the map and the position of the inside monitoring wells, together with their data, are automatically downloaded into an Excel file. This file has a format suitable to be the input of the GuEstNBL software (implemented by the Department of Environmental Engineering—University of Calabria and owned by the Calabria Region, Italy). The tabular format of the file allows the operator to preliminary open it, to easily inspect and edit it if necessary. In this way, new data, not yet included in the GIS, can be added and analyzed together with the pre-existing ones, and those data, pertaining to chemical species having evident correlations with the chemical-physical characteristics of the sampled water, can be divided into homogeneous datasets according to these characteristics. This is particularly the case of redox-sensitive elements such as As, Fe, and Mg.
2.1.2. GuEstNBL

- **Block 3 (Selection of the chemical element)**

  When the software is launched, a starting window appears on the PC desktop and a loading icon allows the operator to browse the computer and find the input file generated by the Web-GIS. Once the data are visualized into the GuEstNBL software, the graphical interface of the program shows a tabular representation of the concentrations of all the analyzed species. The operator can easily select the data associated with a chemical element just clicking on the specific name.

- **Block 4 (Preliminary analysis of the data)**

  The concentrations of some parameters may be lower than the limit of quantification (LOQ) or the limit of detection (LOD) of the analytical method with which they were analyzed. The software automatically identifies these occurrences within the dataset and assigns to the <LOQ or <LOD observations a concentration value equal to LOQ and LOD, or $\frac{1}{2}$LOQ and $\frac{1}{2}$LOD, respectively [7,18]. A box shows the number of “non-detected” concentrations that are now provided with a value and can be part of the subsequent analysis. At the end of this step, the operator selects the methods for the NBL determination and the guided wizard shows the next interactive window. The numbering of the following blocks having the CS prefix refers to the procedures included into the component separation method, while the PS prefix is related to the pre-selection method.

2.1.3. Component Separation Method

The methodology follows the philosophy according to which the concentration of a chemical species in groundwater is due to the combination of natural and anthropogenic (where it exists) components [6]. The natural component is associated with the hydro-geochemical characteristics of the aquifer and to solid-water interaction processes. The anthropogenic component is due to the effects of human activities concerning specific substances whose detection have no causal relationship with the characteristics of a given site and the natural phenomena occurring in it [19–24].

- **Blocks CS1 and CS2 (Calculation of the median values and construction of the observed frequency distribution)**

  The data related to the chosen chemical element are displayed in a table in which they appear chronologically and are associated with their own monitoring well. The software calculates the median values from the available concentration time series at each monitoring well and displays them into another table placed next to the previous one. The observed frequency distribution of the median values is automatically reconstructed and showed into a graph.

- **Block CS3 (Interpretation of the observed frequencies)**

  The frequency distribution of the observed median concentration is interpreted by means of the following frequency distributions combinations:

  $$f_{obs}(c) = f_{nat}(c) + f_{inf}(c) = k\{A \cdot pdf_{LogNorm} + (1 - A) t_N \cdot pdf_{Norm}\},$$

  where $f_{obs}$ are the observed frequencies and $c$ is the concentration of a given environmental parameter. According to Wendland et al., (2005) [6], the first term of the sum is associated with the natural component ($f_{nat}$) described by a log-normal frequency distribution ($pdf_{LogNorm}$), while the second term, represented by a normal distribution ($pdf_{Norm}$), constitutes the influenced component ($f_{inf}$). Moreover, $k$ is the average width of the classes associated with the observed frequencies, $A$ is a weight coefficient, and $t_N$ is a truncation factor (equal to 0.5).

  The two probability density functions (PDFs) are both characterized by a standard deviation and a mean that are estimated by imposing an optimization criterion (nonlinear least square method)
within a calibration procedure automatically performed by the software. the best fitting of the observed data is showed as a graph.

- **Block CS4 (NBL calculation)**

  Parameter calibration is then followed by the automatic computation of the range of concentrations comprised between the 10th (NBL10) and the 90th (NBL90) percentile of the identified log-normal PDF (a graph shows the cumulative distribution function (CDF)), then the NBL is set to be equal to NBL90 \([6,7,18]\), and its value is displayed in a box next to the CDF graph.

- **Blocks CS5-CS7 (NBL reliability)**

  NBL determination must take into account the spatial and temporal variability of the hydro-chemical characteristics of aquifers. the reliability of an estimated NBL depends on the coverage ratio of the available data: a reasonable sample size, suitable to describe the spatial variability of the system under examination, should have a minimum of 15 adequately distributed monitoring points; a reasonable sample size, adequate to describe the temporal variability of the system under examination, should have a minimum of 8 observations, regularly distributed over a period of 2 years on the 80% of the monitoring points. If the dataset meets both requirements, a high level of confidence is attributed to the determined NBL; otherwise, it is considered as provisional pending further data collection \([6,7,18]\).

2.1.4. **Pre-Selection Method**

The pre-selection global statistical method, outlined in the BRIDGE project just like the component separation method, is based on the assumption that the concentration of specific indicator substances, detected into the analyzed samples, is strictly related to anthropogenic influence. If the concentration of these species is higher than well-defined values, the involved samples are excluded from the NBLs estimation procedure, as they are affected by human impact. the opportunity of excluding the data, because they are considered as not representative of the natural system under investigation, is based on the following criteria: the presence of concentrations of organic contaminants, or of other substances related to anthropogenic activity, greater than 75% of the threshold value provided by the current regulations; the presence of nitric or ammonia nitrogen concentrations whose values exceed 10.0 mg/L for nitrates (NO\(_3^-\)) and 0.1 mg/L for ammonia (NH\(_4^+\)) \([7,18]\).

- **Block PS1 (Dataset Pre-selection)**

  Based on what was said above, the software displays an interactive window in which the anthropogenic markers can be selected to exclude the samples with values higher than the established thresholds. the program is provided with a list containing the anthropogenic species together with their control values, and automatically displays those included in the dataset. When a marker is selected and confirmed, the dataset is filtered by those samples not meeting the above requirements \([7,18]\).

- **Block PS2 (Temporal outlier identification and management)**

  In this step, the temporal analysis of the data is carried out to identify potential outliers. the outliers represent concentration values that strongly differ from those within the temporal series of each monitoring well. the challenge is to recognize their nature according to the phenomenon under investigation. Concentration spikes could be caused by anthropic contamination or by the strong presence of a specific element in the minerals of a portion of the aquifer solid matrix. in the first case, the high values are certainly outside the objective of the study, while in the second case they could be considered representative of the natural background. Therefore, the removal of potential outliers must be carefully evaluated. These extreme concentration values, frequently occurring in environmental data, are highlighted by the software through graphic methods or through specific statistical tests.
In particular, when the outliers analysis is run, the program allows to evaluate, simultaneously or well by well, the results of the most used graphical methods (normal quantile–quantile (Q-Q) plots and box plots) and those of the best known statistical methods such as the Discordance [24], Huber [25], Walsh [26], Dixon [27], and Rosner [28] tests. Graphical and statistical results must be evaluated in the light of the outcomes of the conceptual model, so that the operator can decide to keep or discard the single detected outlier, providing a justification for the second operation.

- **Block PS3 (Trend analysis)**

  The concentration time series collected in each monitoring well are analyzed, at this point, looking for the occurrence of a trend in the data. The software estimates the slope of a possible trend line through the implementation of the Mann–Kendall test [29,30]. Depending on the obtained results, the following actions are proposed: the analysis does not show, for the single monitoring well, a significant trend in the observed time window, so the fluctuation in the data can be attributed to seasonal variations under natural conditions; the analysis shows a significant trend in the time series of a specific monitoring well, so the investigated element is suspected of being subject to non-natural control factors. In the first case, the monitoring well is suitable to move forward in the guided procedure. In the second case, an assessment of whether to exclude the monitoring well from the NBL estimation must be done, also considering the indications coming from the conceptual model. The trend analysis window displays the results in the form of graphs, showing the chronological distribution of the observed data in each monitoring well, and the presence of a trend line where this occurs. A table summarizes the results and allows, via checkboxes, to decide the exclusion or not of the wells having a trend in their data.

- **Block PS4 (Calculation of the median values)**

  As described for block CS1, relating to the component separation method, the software calculates the median values from the available concentration time series that have passed the previous steps, assigns them to the respective monitoring well, and displays them into a table.

- **Block PS5 (Spatial outlier identification and management)**

  The same approaches and the same methods adopted in Block PS2 are now repeated at this stage to identify and manage the potential outliers which can be found among the median values.

- **Block PS6 (Analysis of the statistical distribution of the dataset)**

  This fundamental step is presented at this point of the procedure, but it is performed by the software even before the study for the identification of both temporal and spatial outliers. The reason why this analysis recurs more than once lies in the fact that the applicability of certain methodologies, like those previously described for the outlier identification and those coming afterwards, for the NBL estimation, depends on the probability function that best approximates the available observed data. Moreover, given that the statistic sample can undergo the loss of data, precisely because of the potential exclusion of outliers, it is essential to have the possibility to repeat the analysis in this eventuality. This operation is carried out by applying appropriate tests, such as normal quantile–quantile (Q-Q) plots, Shapiro and Wilk [31], D’Agostino [32], and Lilliefors [33], all of which included and were automatically performed by the software.

- **Block PS7 (Spatial and temporal consistency of the dataset)**

  This evaluation, which is also performed in the Component Separation method and it is already described in the Blocks CS5-CS7, distinguishes different levels of spatial and temporal “coverage” of the available data and guides the NBL estimation process. There are 4 distinct cases (A, B, C, and D), which are described below and are automatically identified by the software and proposed to the operator according to the spatial and temporal coverage of the data left over by the previous steps.
Block PS8-PS12 (NBL determination for the datasets falling into cases a and B)

These datasets exhibit an adequate spatial dimension. Case A, unlike case B, also shows an adequate temporal coverage. There is no substantial difference in the NBLs estimation between these two types of dataset. The only distinction is reflected in a higher level of confidence to be attributed to the NBLs determined for the dataset of case A. The software assigns to the NBL value of the maximum observed median, provided that the dataset is normally distributed. If the dataset shows a non-normal distribution, the NBL is given by the 95th percentile of the identified PDF. In particular, the software automatically sets out whether the observed medians are best approximated by a log-normal or a gamma distribution or whether it is best to normalize the data or finally if a non-parametric distribution is the right solution. The parameters of the recognized probability density function are then estimated to fit the observed data within the same automatic calibration procedure described in Block CS3 and the NBL is calculated following the implementations defined in Block CS4. In the case of a distribution suitable for a normalization process, the data are transformed through the Box-Cox transformation [34] and then best fitted by a normal PDF to move forward with the same process described above. Non-parametric datasets (set of data that are not satisfactorily approximated by any distribution) are processed through a graphical method whereby the software draws the cumulative frequency curve of the data and identifies, as representative of the NBL value, the one corresponding to the 90th percentile. Finally, the software can also evaluate the NBL through well-known parameters such as the upper tolerance limit (UTL) and upper prediction limit (UPL) [35].

Block PS13-PS14 (NBL determination for the datasets falling into case C)

This type of dataset shows an adequate temporal dimension but a poor spatial coverage. In this case, the procedures described for datasets a and B used to treat the medians of all monitoring wells, which are applied on the data of the single observation point. An NBL value is thus estimated for each of them. The final NBL representing the entire dataset is given by the maximum value among the estimated ones.

Block PS15-PS16 (NBL determination for the datasets falling into case D)

When the data does not have a significant dimension in either time or space, it is expected that further data and information must be collected and, in the meantime, an estimate of a provisional NBL can be made if a total number of observations $\geq 10$ is available. In this case, the NBL is equal to the 90th percentile of the whole dataset.

2.1.5. Final Report

The software automatically generates an output pdf file at the end of the estimation procedure, whether it has been carried out using the component separation or the pre-selection method. This file contains a detailed report in which all the operations and the choices, made by the operator to reach the final result, are recorded. All the graphic and numerical outcomes are also collected and displayed chronologically.

2.2. Study Area

GuEstNBL has been used in the proposed study to analyze the data recorded within the industrial area of the town of Lamezia Terme located in the Calabria Region, Italy. The site, which is part of the S. Eufemia plain, lies south of the town of Lamezia and overlooks the Tyrrenian Sea, involving an area of approximately 12,000.00 m$^2$ (Figure 2).
The study area, with a roughly rectangular shape, was bounded in order to have a sufficiently homogeneous zone from the morphological, geological, and hydrogeological point of view. The shallow aquifer flowing at little depth from the ground surface is the subject of the investigation. Some monitoring wells, pertaining to the facilities operating within the aforementioned area, have been affected by anomalous values of As, Fe, and Mn concentration. The general purpose of the project was to therefore monitor the portion of the aquifer underlying the study area and to evaluate which part of the observed contamination is the result of natural phenomena, as well as which is attributable to industrial activities. The latter relate to the following sectors: treatment, transformation, and recovery of special and hazardous waste; oil refining and biomass production; wastewater treatment; carpentry and paintwork; farming activities.

In geological terms, the study area is characterized by gravelly-sandy alluvial deposits with silts and clays of continental and marine origin. The analysis of the available stratigraphies highlighted the presence of significant peat beds as a peculiar geological feature of the area. In general, the peat layers found during the drilling operations of the observation wells, most of which have been deepened up to 10 m from the ground level, have a thickness ranging from 0.5 to 3 m. It is likely to believe that the peat presence is due to an extensive swamp that anciently covered the coastal stretch of the plain.

In total, 17 monitoring wells (Figure 3) were used for groundwater periodic sampling and hydraulic levels recording by following the U.S. EPA indications [36], the wells, owned by the facilities operating in the area, are part of the Integrated Environmental Authorization (IEA), an environmental administrative intervention procedure that must be carried out on certain industrial activities that are likely to affect the safety, health of people, or the environment. The IEA was established with the Council Directive 96/61/EC to comply with the principles of Integrated Pollution Prevention and Control (IPPC) dictated by the European Union since 1996. The hydro-geochemical investigation of the area concerned the analysis of heavy metals (As, Fe, Mn, Cu, Zn, Pb, Ni, Hg, Cd, Cr), cations (Na\(^+\), K\(^+\), Ca\(^{2+}\), Mg\(^{2+}\)), anions (Cl\(^-\), SO\(_4^{2-}\), NO\(_3^-\), NO\(_2^-\)), Ammonia Nitrogen (NH\(_4^+\)), Total Phosphorus (P), Phosphates (PO\(_4^{3-}\)), Hydrocarbons (C > 12), and pesticides.
Furthermore, the redox potential (Eh), the pH, and the temperature values were measured in place during the sampling procedures, as well as the electrical conductivity, the dissolved oxygen (O₂), and the chemical oxygen demand (COD). The sampling frequency was on a quarterly basis for a total period of 2 years. Eight withdrawals were hence collected and analyzed for each monitoring well, for a total of 136 samples.

3. Results

In this section, the results obtained by the interpretation of the acquired data is discussed. The qualitative trend of the water depths contours, obtained by means of an ordinary kriging performed on the median values of the observed hydraulic head values, showed a prevailing flow component that moves from east to west (Figure 3). The groundwater contour lines (this is even more evident at a larger scale) show a narrower arrangement in the innermost areas to the west. This trend highlights zones with shorter hydraulic paths and higher hydraulic gradients, which are typical of lower permeability geological formations. In these areas, the aquifer is fed by the rivers flowing through the plain. Moving to the east, toward the coast, the contour lines distance progressively increases, causing a consequent raise of the hydraulic paths and the decrease of hydraulic gradients. This suggests the occurrence of zones with higher permeability. The hydraulic heads distribution of these zones indicates that the aquifer is recharged in the innermost areas while the coastal strip of the plain acts as a drainage zone. This evidence correlates with the hydro-geo-chemical observations since the innermost monitoring wells MW16 and MW17 have higher values of dissolved oxygen associated with positive values of redox potential and pH values between 7 and 7.4, attributable to an oxidizing environment and to a groundwater recharge area supplied by superficial waterbodies. The same parameters, monitored in the other wells closer to the shoreline, show instead conditions attributable to a reducing environment: negative redox potential, $6.7 < \text{pH} < 7.0$, lower values of dissolved oxygen (Figure 4). The aforementioned conditions can be related to a drainage area characterized by a slow water flow with long residence times.
Based on the aforementioned findings, wells MW16 and MW17 have not been included in the dataset for the NBLs estimation. The remaining 15 observation points, falling within a system with rather homogeneous hydro-geochemical properties, and providing both an appropriate spatial and temporal coverage, are the points whose data have set up the input for the GuEstNBL software. The resulting NBLs are therefore assigned to a reduced aquifer portion and delimited by the sampling points involved in the guided estimation procedure. The available data are almost all affected by the presence of ammonia nitrogen (NH$_4^-$) with concentrations exceeding the limit of 0.1 mg/L. The pre-selection method is therefore not compatible with the dataset widely influenced by one of the principle anthropogenic markers (see Section 2.1.4). The application of the component separation (CS) method on the chemical species under consideration is examined below.

### 3.1. Natural Background Levels Estimation

#### 3.1.1. Arsenic

The outcomes stemming from the application of the software guided procedure for arsenic are shown in Figure 5. The reference threshold value for this element is set at 10 µg/L by the Italian environmental laws [37].

The distribution used to interpret the experimental data is characterized by a natural component that displays a sharp and narrow peak at about 15 µg/L. The log-normal (natural) component is prevalent for concentrations lower than 150 µg/L, while for larger concentration values the contribution of the Gaussian (anthropogenic) component takes hold and outclasses the former (Figure 5a). The estimated NBL, namely NBL$_{90}$ corresponding to the 90th percentile (Figure 5b), together with the mixture weight $A$ and the estimated standard deviations ($\sigma_1$, $\sigma_2$) and means ($\mu_1$, $\mu_2$) of the two PDFs, are listed in Table 1. The estimated background value is larger than the reference limit and hence, in line with what was proposed by Muller et al. (2006), a threshold value corresponding to the background level must be assigned to the area.

### 3.1.2. Iron

The curve fitting of the observed relative frequencies associated with the iron concentration values derived from the water samples collected in the monitoring wells, is shown in Figure 6a. The concentrations moving around values of about 10,000 µg/L are associated with high relative frequencies. The natural distribution exhibits a good agreement with all the experimental data (the log-normal PDF coincides with the sum distribution), while the influenced component is completely absent (the Gaussian PDF is flat and lies on the abscissa axis). Table 2 shows the estimated NBL, together with the mixture weight $A$ and the estimated parameters for $f_1$ and $f_2$ distributions.
by the presence of ammonia nitrogen ($\text{NH}_4^-$) with concentrations exceeding the limit of 0.1 mg/L. The pre-selection method is therefore not compatible with the dataset widely influenced by one of the principle anthropogenic markers (see Section 2.1.4). The application of the component separation (CS) method on the chemical species under consideration is examined below.

3.1. Natural Background Levels Estimation

3.1.1. Arsenic

The outcomes stemming from the application of the software guided procedure for arsenic are shown in Figure 5. The reference threshold value for this element is set at 10 µg/L by the Italian environmental laws [37].

**Table 1. GuEstNBL output: Results of the guided estimation procedure for arsenic dataset.**

| Parameter                  | Natural Component ($f_{nat}$) | Influenced Component ($f_{inf}$) |
|----------------------------|-------------------------------|----------------------------------|
| Mean $\mu$ (µg/L)          | 32.28                         | 197.86                           |
| Standard deviation $\sigma$ (µg/L) | 1.02                         | 16.96                           |
| Mixture weight $a$         | 1.02                          | 0.49                             |
| NBL$_{90}$ (µg/L)          | 119.24                        | 119.24                           |

![Figure 5](image_url)

*Figure 5. GuEstNBL output: (a) Iron empirical frequencies together with the estimated natural and anthropogenic components, and their weighted sum distribution; (b) CDF and 90th percentile of the identified log-normal PDF.*

The distribution used to interpret the experimental data is characterized by a natural component that displays a sharp and narrow peak at about 15 µg/L. The log-normal (natural) component is prevalent for concentrations lower than 150 µg/L, while for larger concentration values the contribution of the Gaussian (anthropogenic) component takes hold and outclasses the former (Figure 5a).
the estimated NBL, namely NBL<sub>90</sub> corresponding to the 90th percentile (Figure 5b), together with the mixture weight <i>a</i> and the estimated standard deviations (σ<sub>nat</sub>, σ<sub>inf</sub>) and means (µ<sub>nat</sub>, µ<sub>inf</sub>) of the two PDFs, are listed in Table 1. The estimated background value is larger than the reference limit and hence, in line with what was proposed by Muller et al. (2006), a threshold value corresponding to the background level must be assigned to the area.

### 3.1.2. Iron

The curve fitting of the observed relative frequencies associated with the iron concentration values derived from the water samples collected in the monitoring wells, is shown in Figure 6a. The concentrations moving around values of about 10,000 µg/L are associated with high relative frequencies. The natural distribution exhibits a good agreement with all the experimental data (the log-normal PDF coincides with the sum distribution), while the influenced component is completely absent (the Gaussian PDF is flat and lies on the abscissa axis). Table 2 shows the estimated NBL, together with the mixture weight <i>a</i> and the estimated parameters for <i>f<sub>nat</sub></i> and <i>f<sub>inf</sub></i> distributions.

**Table 2.** GuEstNBL output: Results of the guided estimation procedure for iron dataset.

| Parameter                  | Natural Component (f<sub>nat</sub>) | Influenced Component (f<sub>inf</sub>) |
|----------------------------|------------------------------------|----------------------------------------|
| Mean µ (µg/L)              | 13,183.62                          | -                                      |
| Standard deviation σ (µg/L)| 0.59                               | -                                      |
| Mixture weight <i>a</i>    | 1.0                                | 1.0                                    |
| NBL<sub>90</sub> (µg/L)    | 27,960.52                          | 27,960.52                              |

### 3.1.3. Manganese

The reference threshold value for manganese is 50 µg/L [37]. Figure 7a shows the results of the CS method for this metal. The lowest monitored concentrations (around 1000 µg/L) are associated with high relative frequencies. The observed frequencies are well interpreted by the log-normal (natural) distribution for concentrations up to 4000 µg/L; then, the Gaussian (anthropogenic) component becomes relevant.

The estimated NBLs together with the parameters resulting from the curve fitting are collected in Table 3. A natural background concentration (NBL<sub>90</sub>) of 5718.20 µg/L was estimated, which is higher than the regulation limit and, thus, as in the previous cases, becomes the new threshold value (Figure 7b).

**Table 3.** GuEstNBL output: Results of the guided estimation procedure for manganese dataset.

| Parameter                  | Natural Component (f<sub>nat</sub>) | Influenced Component (f<sub>inf</sub>) |
|----------------------------|------------------------------------|----------------------------------------|
| Mean µ (µg/L)              | 2097.65                            | 7671.01                                |
| Standard deviation σ (µg/L)| 0.78                               | 1350.7                                  |
| Mixture weight <i>a</i>    | 0.65                               | 1.0                                    |
| NBL<sub>90</sub> (µg/L)    | 5718.20                            | 27,960.52                              |

**Figure 7.** GuEstNBL output: (a) Manganese empirical frequencies together with the estimated natural and anthropogenic components, and their weighted sum distribution; (b) CDF and 90th percentile of the identified log-normal PDF.
The iron estimated NBL\textsubscript{90} is extremely larger than the reference limit (200 µg/L [37]) and, hence, according to what was done for the arsenic results, a threshold value corresponding to the background level must be assigned to this element (Figure 6b). A natural background value larger than the current regulation limit indicates that the state of the investigated aquifer cannot be regarded as poor as long as the iron concentrations do not exceed this new threshold because of the occurrence of a specific natural signature typical of the examined context.

3.1.3. Manganese

The reference threshold value for manganese is 50 µg/L [37]. Figure 7a shows the results of the CS method for this metal. The lowest monitored concentrations (around 1000 µg/L) are associated with high relative frequencies. The observed frequencies are well interpreted by the log-normal (natural) distribution for concentrations up to 4000 µg/L; then, the Gaussian (anthropogenic) component becomes relevant.

The estimated NBLs together with the parameters resulting from the curve fitting are collected in Table 3. A natural background concentration (NBL\textsubscript{90}) of 5718.20 µg/L was estimated, which is higher than the regulation limit and, thus, as in the previous cases, becomes the new threshold value (Figure 7b).

\textbf{Table 3. GuEstNBL output: Results of the guided estimation procedure for manganese dataset.}

| Parameter                      | Natural Component (f\textsubscript{nat}) | Influenced Component (f\textsubscript{inf}) |
|--------------------------------|----------------------------------------|--------------------------------------------|
| Mean µ (µg/L)                  | 2097.65                                | 7671.01                                    |
| Standard deviation σ (µg/L)    | 0.78                                   | 1350.7                                     |
| Mixture weight a               |                                         | 0.65                                       |
| NBL\textsubscript{90} (µg/L)   |                                         | 5718.20                                    |

4. Discussion

The guided procedure offered by the software has comfortably led to NBL estimation. The three selected elements have important natural components and this can be understood by considering the observed site characteristics and correlation between the concentration values of the involved chemical species. Figure 8 shows a comparison between the median values of arsenic, iron, and manganese concentration recorded in each well during the two-year monitoring period. It can be noticed that iron and manganese concentrations have substantially the same pattern, as it is clearly evident from the two peaks registered for the first element that correspond to the peaks of the second one (monitoring wells MW2 and MW4) even if with significantly lower values. Arsenic basically shows high concentrations in correspondence to iron and manganese high concentrations, and low concentrations where the presence of this two elements is poor. The investigation of the origins of arsenic high concentrations in groundwater is a widely documented topic in the literature [38–45]. All authors highlighted the primary role of peat, which was widely present in the current study area, in the release of arsenic associated with high concentrations of iron and manganese in anoxic conditions. In the past century, the S. Eufemia plain was a large swamp reclaimed at the end of the 20th century. The study area is one of the sectors most affected by swamping. The presence of large marshlands, in which the deposition and decay of the vegetation occurred, is responsible for the creation of the peat deposits. The arsenic release mechanism, which is more accredited in the aforementioned literature, is based on the formation of iron and manganese oxides and hydroxides (naturally and richly present in the Calabrian soils) on which the arsenic is adsorbed. These complexes are stable in an oxidizing environment (the samples taken in wells MW16 and MW17 contain almost zero concentrations of the three elements), while they dissolve in a reducing environment where the bridging oxygen that binds arsenic to iron and manganese oxides and hydroxides fails, causing their simultaneous release into the water.
This can explain what happens in every other well pertaining to the reducing zone of the aquifer and having significant concentrations of all three species (see Section 3 and Figure 4). The review of the dataset, acquired in the two-year monitoring period, did not show significant concentrations of hydrocarbons and pesticides. On the other hand, a variable concentration of Ammonia Nitrogen (NH$_4^-$) and Nitrates (NO$_3^-$) was found throughout the study area. Nitrite (NO$_2^-$) concentration was always found below the detection limit of the adopted analytical method. the trend of nitrogenous compounds highlights concentrations of NH$_4^-$ above the regulation limits in almost all sampling points. the presence of NO$_3^-$ has been recorded in the upstream sector of the aquifer, more oxygenated, and with positive Eh values, while its concentrations become very low in the downstream sectors, scarcely oxygenated, and with negative Eh values. This suggests the generation of oxidizing and reducing processes typical of the nitrogen (N$_2$) cycle, according to which nitrogen compounds are transformed through aerobic (nitrification) and anaerobic (denitrification) processes, starting from the N$_2$ contained in organic substances. Commonly, these processes occur because of the decomposition of dejections and so their occurrence is used as an indicator of organic pollution of anthropogenic nature. However, nitrogen compound transformations can also be triggered by the decomposition of natural organic substances such as the peat found in the study area. the current analyses do not allow to definitively establish whether these compounds pertain to an anthropogenic or a natural pollution, this determination will be the subject of further studies. For this reason, NBL determination was not possible by means of the pre-selection method and the component separation method was used instead. the software therefore provides the possibility of using both approaches, depending on the anthropic impact on the study area, and therefore on the presence, or not, of anthropic markers with relevant concentrations within the observed dataset.

The implemented procedure is hence a versatile instrument but the reliability of its results depend on the compliance of the following general criteria:

- The input data must come from a single aquifer or a portion of it. the study area must have homogeneous characteristics in terms of geological and hydro-geo-chemical properties, since the procedure provides the estimation of a single NBL value assigned to the whole investigated area.
- The investigation scale depends on the spatial extension within which the aquifer homogeneous characteristics are maintained. This means that, depending on the case, the software can be used for local or regional studies. in practical applications, what provides the actual size of the area to be charged with the estimate result is the number and the distribution of the available observation points falling within the zone of homogeneity. the minimum number of observation points,
their distribution, and sampling frequencies are discussed in Section 2.1.3 (Blocks CS5-CS7 (NBL reliability)).

- The operator using the software is an integral part of it. He or she makes important decisions by systematizing the information coming from the study of the conceptual model and from the results gradually emerging from the guided procedure. This implies that not everyone is able to obtain reliable results.

5. Conclusions

The collaboration between the Department of Environmental Engineering of the University of Calabria and the Department of Environmental Policies of the Calabria Region resulted in the implementation of a useful procedure that relied on the joint application of a web-based geographic information system (GIS) and an ad hoc developed software (GuEstNBL). Thus, the onerous and complex issue of the Natural Background Levels estimation was more easily addressed. The environmental agencies responsible for the protection, management, and remediation of the territories can take advantage of such a tool, since it is equipped with a graphical wizard that guides the operators to implement the component separation and pre-selection methods to perform all computations required therein. The procedure can be further improved by adding additional statistical and graphical tests, and, in the present state, is going to be the official evaluation tool of the Calabria Region to identify groundwater thresholds.

In the present study, the developed calculation program was used for the NBL estimation of arsenic, iron, and manganese in a potentially contaminated industrial site located in Lamezia Terme, Italy. The outcomes of the software elaborations, together with the cross-reading of geological, hydrogeological, and hydro-geochemical data, have shown that the presence of the anomalous concentrations detected in the shallow aquifer may be due to natural phenomena linked to the degradation of peat deposits that drive the triggering of the redox reaction that generates a dissolution of the pollutants into the groundwater. The three investigated chemical species have in fact important natural components that result in natural background concentrations higher than the regulation limit. Thus, they can become the new threshold values. NBL estimation was obtained through the component separation method since the dataset was almost only affected by the presence of ammonia nitrogen ($\text{NH}_4^-$) with concentrations exceeding the limit of 0.1 mg/L. The pre-selection method was therefore incompatible with the dataset widely influenced by one of the principle anthropogenic markers.

The results of this study showed that groundwater systems are complex and require an accurate characterization to distinguish effective anthropogenic contamination from natural conditions. This is extremely important to define clean-up goals consistent with specific natural features of an analyzed water body, since the costs for aquifer management practices and remediation actions are dramatically high. The realized calculation framework has a great exploitation perspective since it can be used worldwide.

Author Contributions: Conceptualization, F.C. and M.D.B.; methodology, F.C. and M.D.B.; software, F.C. and M.D.B.; validation, F.C., M.D.B., A.C., E.C., and C.M.; formal analysis, F.C. and M.D.B.; investigation, A.C. and E.C.; resources, F.C., M.D.B., A.C., E.C., and C.M.; data curation, F.C., M.D.B., A.C., E.C., and C.M.; writing—original draft preparation, F.C.; writing—review and editing, F.C.; visualization, F.C. and M.D.B.; supervision, O.R., I.T., and S.S.; project administration, O.R., I.T., and S.S.; funding acquisition, O.R., I.T., and S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Regione Calabria.

Acknowledgments: This research has been supported by the “AIM: Attraction and International Mobility”, PON R&I 2014-2020 Calabria

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.
References

1. Hinsby, K.; Conde do Melo, M.T. Application and evaluation of a proposed methodology for derivation of groundwater threshold values—a case study summary report. Report to the EU Project ‘BRIDGE’. 2006. Deliverable D22. Available online: http://nfp-at.eionet.europa.eu/Public/irc/eionet-circle/bridge/library?/deliverables/d22_final_reppdf_/EN_1.0_&a=s (accessed on 29 September 2020).

2. 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]

3. European Commission. Guidance on Groundwater Status and Trend Assessment, Guidance Document #18; Technical Report; European Communities: Luxembourg, 2009; ISBN 978-92-79-11374-1.

4. Panno, S.V.; Kelly, W.R.; Martinsek, A.T.; Hackley, K.C. Estimating background and threshold nitrate concentrations using probability graphs. GroundWater 2006, 44, 697–709. [CrossRef] [PubMed]

5. 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.

6. 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]

7. BRIDGE. Background criteria for the identification of groundwater thresholds. 2007. Available online: http://nfp-at.eionet.europa.eu/irc/eionet-circle/bridge/info/data/en/index.htm (accessed on 27 August 2020).

8. Edmunds, W.M.; Shand, P. Natural Groundwater Chemistry; Blackwell Publishing Ltd.: Oxford, UK, 2008; p. 469.

9. Hinsby, K.; Conde do Melo, M.T.; 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]

10. 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]

11. Kim, K.H.; Yun, S.T.; Kim, H.K.; Kim, J.W. Determination of natural backgrounds and thresholds of nitrate in South Korean groundwater using model-based statistical approaches. J. Geochem. Explor. 2015, 148, 196–205. [CrossRef]

12. 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]

13. Stan, C.O.; Buliga, I. Natural background levels and threshold values for phreatic waters from the Quaternary deposits of the Bahlui River basin. Aui Geol. 2013, 59, 51–60.

14. Sun, L.; Gui, H.; Peng, W.; Lin, M. Heavy Metals in Deep Seated Groundwater in Northern Anhui Province, China: Quality and Background. Nat. Environ. Pollut. Technol. 2013, 12, 533–536.

15. Molinari, A.; Chidichimo, F.; Straface, S.; Guadagnini, A. Assessment of natural background levels in potentially contaminated coastal aquifers. Sci. Total Environ. 2014, 476–477, 38–48. [CrossRef]

16. Preziosi, E.; Parrone, D.; Del Bon, A.; Ghergo, S. Natural background level assessment in groundwater: Probability plot versus pre-selection method. J. Geochem. Explor. 2014, 143, 43–53. [CrossRef]

17. Chidichimo, F.; De Biase, M.; Straface, S. Groundwater pollution assessment in landfill areas: Is it only about the leachate? Waste Manag. 2020, 102, 655–666. [CrossRef]

18. ISPRA—Istituto Superiore per la Protezione e la Ricerca Ambientale. Linee Guida per la Determinazione dei Valori di Fondo per i Suoli e per le Acque Sotterranee; Manuali e Linee Guida 174/2018; ISPRA: Rome, Italy, 2018; ISBN 978-88-448-0880-8.

19. Tiedeken, E.J.; Tahar, A.; McHugh, B.; Rowan, N.J. Monitoring, sources, receptors, and control measures for three European Union watch list substances of emerging concern in receiving waters—a 20 year systematic review. Sci. Total Environ. 2017, 574, 1140–1163. [CrossRef]

20. Xiao, F. Emerging poly- and perfluoroalkyl substances in the aquatic environment: a review of current literature. Water Res. 2017, 124, 482–495. [CrossRef]

21. Tursi, A.; Chatzisymine, E.; Chidichimo, F.; Beneduci, A.; Chidichimo, G. Removal of Endocrine Disrupting Chemicals from Water: Adsorption of Bisphenol-A by Biobased Hydrophobic Functionalized Cellulose. Int. J. Environ. Res. Public Health 2018, 15, 2419. [CrossRef]
22. Tursi, A.; De Vietro, N.; Beneduci, A.; Milella, A.; Chidichimo, F.; Fracassi, F.; Chidichimo, G. Low pressure plasma functionalized cellulose fiber for the remediation of petroleum hydrocarbons polluted water. *J. Hazard. Mater.* 2019, 373, 773–782. [CrossRef]

23. De Vietro, N.; Tursi, A.; Beneduci, A.; Chidichimo, F.; Milella, A.; Fracassi, F.; Chatzisymeon, E.; Chidichimo, G. Photocatalytic inactivation of Escherichia coli bacteria in water using low pressure plasma deposited TiO2 cellulose fabric. *Photoch. Photobiol. Sci.* 2019. [CrossRef]

24. Barnett, V.; Lewis, T. *Outliers in Statistical Data*, 3rd ed.; Wiley: Chichester, UK, 1994; 584p.

25. Huber, P.J. *Robust Statistics*; John Wiley and Sons: Hoboken, NY, USA, 1981.

26. Walsh, J.E. Large sample nonparametric rejection of outlying observations. *Ann. I. Stat. Math.* 1958, 10, 223–232. [CrossRef]

27. Dixon, W.J. Processing Data for Outliers. *Biometrics* 1953, 9, 74–89. [CrossRef]

28. Rosner, B. On the Detection of many outliers. *Technometrics* 1975, 17, 221–227. [CrossRef]

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

30. Kendall, M.G. *Rank Correlation Methods*; Griffin: London, UK, 1975.

31. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* 1965, 52, 591–611. [CrossRef]

32. D’Agostino, R.B.; Stephens, M.A. *Goodness-of-Fit Techniques*; Marcel Dekker Inc.: New York, NY, USA, 1986.

33. Lilliefors, H.W. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *J. Am. Stat. Assoc.* 1967, 62, 399–404. [CrossRef]

34. Box, G.E.P.; Cox, D.R. An analysis of transformations. *J. R. Stat. Soc. B* 1964, 26, 211–252. [CrossRef]

35. ISO 16269-6, Statistical Interpretation of Data, Part 6: Determination of Statistical Tolerance Intervals, Technical Committee ISO/TC 69, Applications of Statistical Methods. Available online: https://www.iso.org/obp/ui/#iso:std:iso:16269:-6:ed-2:v1:en (accessed on 27 August 2020).

36. Puls, R.W.; Barcelona, M.J. *Ground Water Issue: Low-Flow (Minimal Drawdown) Groundwater Sampling Procedures*; Technical Report EPA/540/S-95/504; US EPA: Washington, DC, USA, 1996.

37. D.Lgs. 152/06—Decreto Legislativo n. 152 del 3 aprile 2006. Norme in materia ambientale. Pubblicato nella Gazzetta Ufficiale n. 88 del 14 Aprile 2006—Supplemento Ordinario n. 96. Available online: https://www.gazzettaufficiale.it/detttaglio/codici/materiaAmbientale (accessed on 29 September 2020).

38. Mc Arthur, J.M.; Ravenscroft, P.; Safiulla, S.; Thirlwall, M.F. Arsenic in groundwater: Testing pollution mechanisms for sedimentary acquirers in Bangladesh. *Water Resour. Res.* 2001, 37, 109–117. [CrossRef]

39. Mc Arthur, J.M.; Banerjee, D.M.; Hudson Edwards, K.A.; Mishra, R.; Purohit, R.; Ravenscroft, P.; Cronin, A.; Howarth, R.J.; Chatterjee, A.; Talukder, T.; et al. 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. Geochem.* 2004, 19, 1255–1293. [CrossRef]

40. Peters, S.C.; Burkert, L. the occurrence and geochemistry of arsenic in groundwater of the Newark basin of Pennsylvania. *Appl. Geochem.* 2008, 23, 85–98. [CrossRef]

41. Saunders, J.A.; Lee, M.K.; Shamsudduha, M.; Dhakal, P.; Uddin, A.; Chowdury, M.T.; Ahmed, K.M. Geochemistry and mineralogy of arsenic in (natural) anaerobic groundwaters. *Appl. Geochem.* 2008, 23, 3205–3214. [CrossRef]

42. Hung Hsu, C.; Tang Han, S.; Hsuan Kao, Y.; Wuing Liu, C. Redox characteristics and zonation of arsenic affected multy layer aquifers in the Choushui River alluvional fan, Taiwan. *J. Hydrol.* 2010, 391, 351–366.

43. Lawati, W.M.; Rizoulis, A.; Eiche, E.; Boothman, C.; Polya, D.A.; Lloyd, J.R.; Berg, M.; Vasquez-Aguilar, P.; Van Dongen, B.E. Characterization of organic matter and microbial communities in contrasting arsenic rich Holocene and arsenic-poor Pleistocene aquifers, Red River Delta, Vietnam. *Appl. Geochem.* 2012, 27, 315–325. [CrossRef]
44. Rotiroti, M.; Sacchi, L.; Fumagalli, L.; Bonomi, T. Origin of Arsenic in groundwater from multilayer aquifer in Cremona (Northern Italy). *Environ. Sci. Technol.* 2014, 14, 5395–5403. [CrossRef]

45. Carraro, A.; Fabbri, P.; Giaretta, A.; Peruzzo, L.; Tateo, F.; Tellini, F. Effect of redox conditions on the control of arsenic mobility in shallow alluvional aquifers on the Venezia Plain (Italy). *Sci. Total Environ.* 2015, 532, 581–594. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).