The Multi-Parameter Mapping of Groundwater Quality in the Bourgogne-Franche-Comté Region (France) for Spatially Based Monitoring Management

Abderrahim Bousouis, Abdelhak Bouabdli, Meryem Ayach, Laurence Ravung, Vincent Valles, Laurent Barbiero

sustainability

MDPI

Article

The Multi-Parameter Mapping of Groundwater Quality in the Bourgogne-Franche-Comté Region (France) for Spatially Based Monitoring Management

Abderrahim Bousouis [1], Abdelhak Bouabdli [1], Meryem Ayach [2], Laurence Ravung [3], Vincent Valles [4,5] and Laurent Barbiero [6,*]

1 Laboratoire de Géosciences, Faculté des Sciences, Université Ibn Tofaïl, BP 133, Kénitra 14000, Morocco; abderrahim.bousouis@uit.ac.ma (A.B.); bouabdli@uit.ac.ma (A.B.)
2 Geosciences, Water and Environment Laboratory, Faculty of Sciences Rabat, Mohammed V University, Rabat 10000, Morocco; meryem_ayach@um5.ac.ma
3 Agence Régionale de Santé (ARS) Bourgogne-Franche Comté, 8 rue Heim, 90005 Belfort cedex, France; laurence.ravung@ars.sante.fr
4 Mixed Research Unit EMMAH (Environnement Méditerranéen et Modélisation des Agro-Hydrosystèmes), Hydrogeology Laboratory, Avignon University, 84916 Avignon, France; vincent.valles@outlook.fr
5 Soil Science Department (FSTBM), Faculté des Sciences et Techniques, BP 523, Beni Mellal 23000, Morocco
6 Institut de Recherche pour le Développement, Géoscience Environnement Toulouse, CNRS, University of Toulouse, Observatoire Midi-Pyrénées, UMR 5563, 14 Avenue Edouard Belin, 31400 Toulouse, France
* Correspondence: laurent.barbiero@get.omp.eu

Abstract: Groundwater, a vital resource for providing drinking water to populations, must be managed sustainably to ensure its availability and quality. This study aims to assess the groundwater quality in the Bourgogne-Franche-Comté region (~50,000 km^2) of France and identify the processes responsible for its variability. Data were extracted from the Sise-Eaux database, resulting in an initial sparse matrix comprising 8723 samples and over 100 bacteriological and physicochemical parameters. From this, a refined full matrix of 3569 samples and 22 key parameters was selected. The data underwent logarithmic transformation before applying principal component analysis (PCA) to reduce the dimensionality of the dataset. The analysis of the spatial structure, using both raw and directional variograms, revealed a categorization of parameters, grouping major ions according to the regional lithology. Bacteriological criteria (*Escherichia coli* and *Enterococcus*) displayed strong spatial variability over short distances, whereas iron (Fe) and nitrates showed intermediate spatial characteristics between bacteriology and major ions. The PCA allowed the creation of synthetic maps, with the first seven capturing 80% of the information contained in the database, effectively replacing the individual parameter maps. These synthetic maps highlighted the different processes driving the spatial variations in each quality criterion. On a regional scale, the variations in fecal contamination were found to be multifactorial, with significant influences captured by the first four principal components. The 22 parameters can be grouped into six categories based on their spatial and temporal variations, allowing for the redefinition of a resource management and monitoring strategy that is adapted to the identified spatial patterns and processes at the regional scale, while also reducing analytical costs.

Keywords: bacteriological composition; Bourgogne-Franche-Comté region; chemical composition; cluster analysis; groundwater; principal component analysis

check for updates

Citation: Bousouis, A.; Bouabdli, A.; Ayach, M.; Ravung, L.; Valles, V.; Barbiero, L. The Multi-Parameter Mapping of Groundwater Quality in the Bourgogne-Franche-Comté Region (France) for Spatially Based Monitoring Management. *Sustainability* **2024**, *16*, 8503. https://doi.org/10.3390/su16198503

Academic Editor: Andrea G. Capodaglio

Received: 26 August 2024
Revised: 22 September 2024
Accepted: 27 September 2024
Published: 29 September 2024

1. Introduction

Groundwater aquifers, being less susceptible to pollution than surface waters, are crucial for supplying drinking water to populations. Historically, their sustainable management has focused primarily on the quantity of water extracted [1]. However, in recent decades, it has become clear that sustainable management must be approached through the integrated lenses of hydrological, ecological, and socio-economic systems and in a co-evolutionary manner [2,3]. The causes of groundwater resource degradation are diverse, including over-extraction relative to recharge—potentially leading to saline intrusion in coastal aquifers—inadequate protection from anthropogenic activities (urbanization, industrial waste disposal, agricultural intensification), and the additional pressures brought by climate change [4–8]. In this context, the past fifteen years have seen a growing focus on the development of sustainability indices [4,9,10]. Nevertheless, achieving sustainable management fundamentally depends on a deep understanding of groundwater diversity and the mechanisms responsible for this diversity.

Groundwater databases are a valuable source of information, capturing the water–rock interactions, the filtration capacity of surface layers, and land use patterns, as well as meteorological events and climate changes over several decades [11–13]. Groundwater quality observation networks may exist on a local scale, more rarely on a regional or national scale, and are non-existent on a global scale, and these networks are generally active for short periods. In addition, efforts are still required to harmonize monitoring and analytical tools between the organizations concerned, especially in the case of international basins [14]. Another challenge, frequently mentioned, is to make the information available to anyone who is involved or interested [15]. In France, these databases related to water intended for human consumption have been maintained for over 40 years, covering an extensive network of catchment points across the entire country (4% surface water and 96% groundwater) [16]. These databases are managed by the Regional Health Agencies. The monitoring parameters are diverse and have evolved with advancements in analytical techniques, including major and trace ions, bacteriology, hydrocarbons and other organic pollutants, pesticides, nanoparticles, microplastics, and other emerging contaminants that pose potential risks to human health and environmental ecosystems [17]. Data are periodically collected not only for inventory purposes but also for quality monitoring, and they feed into the Sise-Eaux database, which is managed at the national level and administered regionally by the Regional Health Agencies [18].

In recent years, our research group has leveraged the information contained in the Sise-Eaux database by extracting data at the level of administrative regions such as Provence-Alpes-Côte d'Azur [19,20] in the far southeast of France, Occitanie [21], a region spanning two major watersheds with contrasting climates (Mediterranean and Atlantic), Auvergne-Rhône-Alpes [22], and the island of Corsica [23]. These studies aim to understand the key mechanisms driving water diversity, with the ultimate goal of optimizing the protection, monitoring, and quality control efforts carried out by health agencies. The initial findings highlighted that the heterogeneity of natural environments at the regional level obscures some of the key processes involved in water quality formation. This issue was addressed by grouping groundwater into homogeneous subsets after dimensionality reduction in the data space. The presence of extreme values in the dataset amplifies the impact of certain parameters, which was mitigated through logarithmic data transformation. Discriminating between spatial and temporal variances allowed us to identify seasonal mechanisms and long-term trends. A typology of quality parameters was established based on spatial distribution structures and associations between these parameters.

Building on this methodology, we applied it to the whole Bourgogne-Franche-Comté (BFC) region, which is known for its significant spatial diversity. In previous studies, several hypotheses were proposed regarding the pathways of bacterial contamination in groundwater or the role of redox processes in metal release. These hypotheses are tested by incorporating additional parameters such as the turbidity, the total organic carbon, and several trace elements. The objective of this research is to explore the spatial structure of numerous water quality parameters, whether microbiological or physicochemical, and to attempt to categorize these groundwater quality parameters on a regional scale.

2. Materials and Methods

2.1. Bourgogne-Franche-Comté Region

The Bourgogne-Franche-Comté region, located in the center-east of France (Figure 1), is composed of eight administrative departments: Côte-d'Or, Jura, Haute-Saône, Nièvre, Yonne, Saône-et-Loire, Territoire de Belfort, and Doubs. Covering an area of 47,784 km^2, it is the fifth-largest administrative region in France. The region spans an altitudinal range from the Saône plain and the Seine valley (~170 m) to the peaks of the Vosges (1247 m) in the northeast and the Jura (1495 m) in the east. The Morvan Massif (901 m) occupies the center of the region. Bourgogne-Franche-Comté is bordered by the Centre-Val de Loire region to the west, Île-de-France to the northwest, Auvergne-Rhône-Alpes to the south, Grand Est to the north, and Switzerland to the east (Figure 1). The eastern half of the region lies within the Rhône basin, while the western half corresponds to the upper basins of the Loire and Seine, which flow, respectively, towards the Atlantic Ocean and the English Channel.

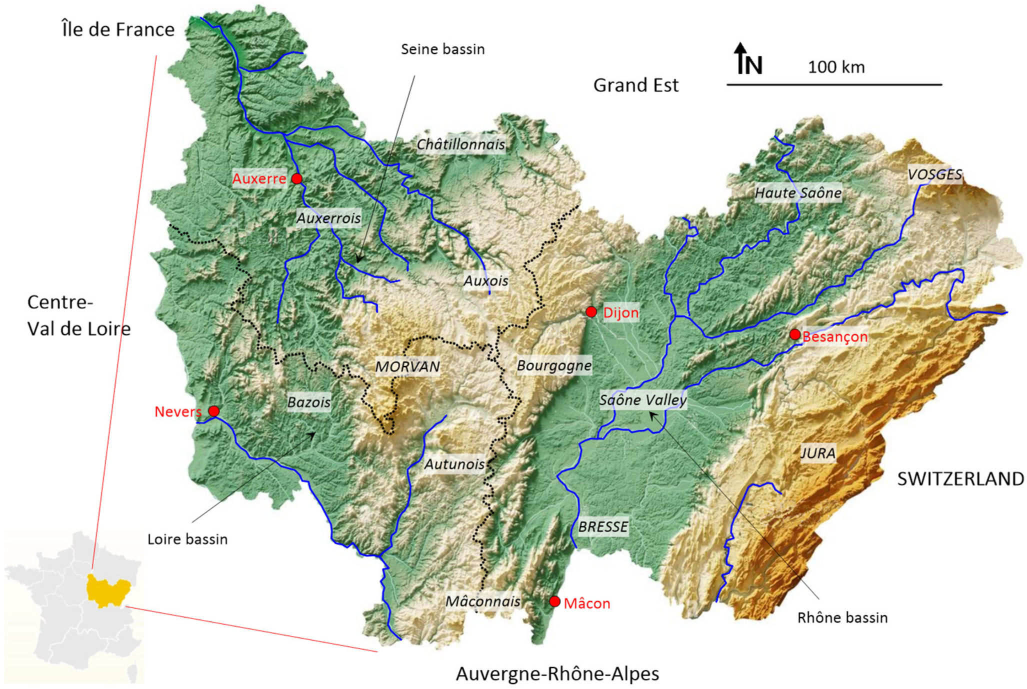
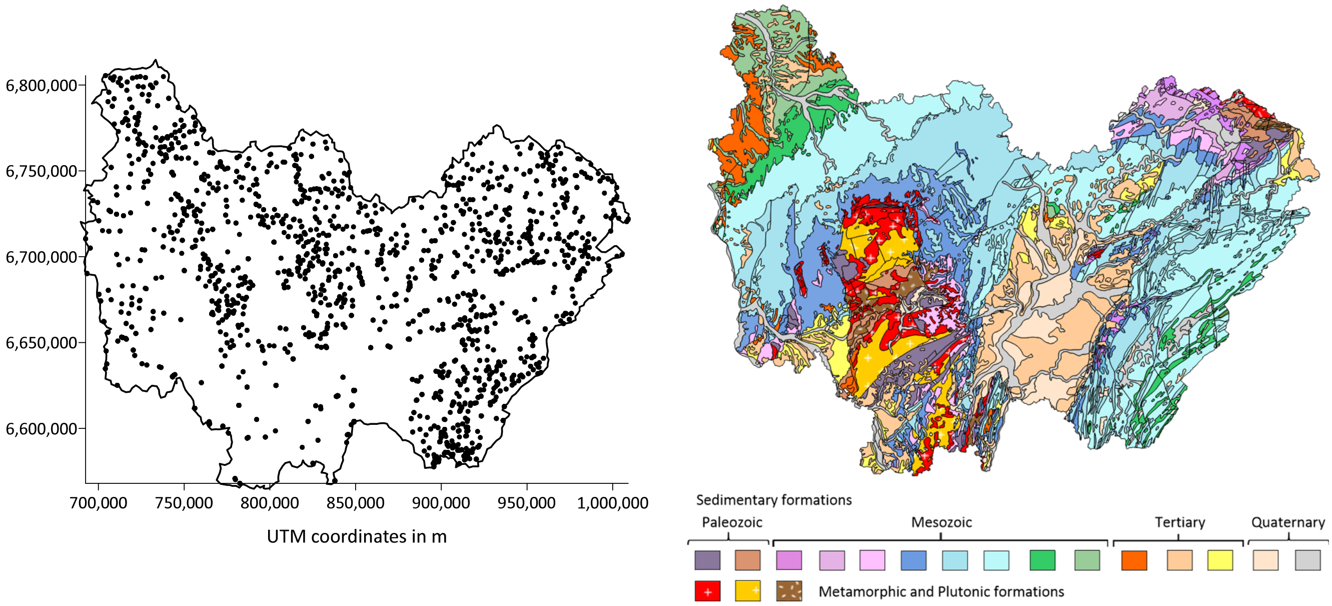


Figure 1. Location and relief of the BFC region.

From a geological perspective (Figure 2), the region is composed of three major structural domains:

- The Hercynian massifs: These are made up of metamorphic and crystalline rocks (schists, gneiss, granites) that form the bedrock of the entire Bourgogne, the Morvan, and the southern Vosges. The erosion of these massifs provided the materials accumulated in the lower areas.
- The Mesozoic Domain: Primarily composed of sedimentary rocks (limestone, chalk, marl) formed in shallow seas over the flattened Hercynian bedrock. These sedimentary deposits are observed on the plateaus of Bourgogne, Haute-Saône, Auxois, Bazois, and the Jura massif, the latter of which was formed 35 million years ago by compression exerted from the Alps towards the west.
- The Cenozoic Domain: Formed after the retreat of the sea and the creation of the Saône graben, this domain consists of materials carried by rivers descending from the emerging areas and is covered by glacial deposits that were laid down during major glaciations in the region.

The aquifers are fractured in the Morvan and karstic in the Jura to the southeast of the study area. On the plains, they are essentially aquifers accompanying watercourses and porous media. In the western part of the region, the climate is characterized as altered oceanic. Toward the east, the Morvan and the plateaus create a zone where the climate becomes more complex. It is typical of mid-altitude mountains, with high rainfall, cold winters, and cool summers. In the south, up to the latitude of Dijon, the plains experience a southern influence, especially along the Bourgogne wine-producing Côte. Further north, continental influences prevail, resulting in hot summers, frequent rainfall accumulating up to 1200 mm near the Vosges and Jura massifs, and cold winters. A low mountain climate dominates the Jura plateaus, where precipitation can reach 1600 mm per year. The high Jura chain is characterized by a mountain climate with significant snowfall, rapidly decreasing temperatures with altitude, mild to cool summers, and frequent thunderstorms.



(**a**) (**b**)

Figure 2. (**a**) Location of the 989 groundwater catchments; (**b**) simplified geological map of the BFC region (adapted from BRGM, https://www.brgm.fr/fr/implantation-regionale/bourgogne-franche-comte, accessed in January 2024).

2.2. The Sise-Eaux Database

The Sise-Eaux database (https://data.eaufrance.fr/concept/sise-eaux, accessed in March 2023) is managed by the Regional Health Agencies (ARS), meaning that the extraction of sanitary analysis data is performed at the regional level. All analyses were conducted by laboratories accredited by the ARS, holding both international accreditation and analytical quality certification [16,18]. Only raw water samples, meaning water not treated for potability, were used in this study. The data extraction covers a 5-year period, from 11 January 2016 to 25 May 2021. The number of parameters analyzed for each water sample varies depending on whether it was a routine check (basic physicochemical parameters, major ions, and fecal bacteriology), a targeted analysis, or a more comprehensive analysis, such as for a first-time sampling (up to a hundred parameters including pesticides, metals, metalloids, trace elements, full bacteriology, etc.). Consequently, extracting these data results in matrices with numerous missing values, creating what are known as "sparse matrices".

For this study, the sparse matrix resulting from regional extraction contains 8723 observations and over 100 parameters. After manually correcting data entry errors and removing missing values, a full matrix with 3569 observations and 22 parameters was selected for analysis. The full matrix was obtained by eliminating incomplete lines, without any data interpolation. These 22 parameters include electrical conductivity at 25 °C (EC); fecal bacteriology with Enterococcus and Escherichia coli (Ent., *E. coli*); major ions, Na, Ca, Mg, Cl, SO_4, and HCO_3; nitrogen forms NO_3 and NH_4; metals and traces, Fe, Mn, B, F, As, Se, Cd, and Ni; total organic carbon (TOC); and turbidity (Turb.). In the following text, a distinction will be made between parameters and major ions, for example, SO_4 representing the parameter of the analyzed ion SO_4^{2-}. The distribution of sampling points in this study is illustrated in Figure 2. The full matrix represents 989 sampling points, meaning that each catchment area was sampled an average of 3.6 times during the extraction period in question. These sampling points cover the whole of the Bourgogne-Franche-Comté region in a fairly homogenous way and bring together the catchments supplying water in quantities that can be used by the communities. Where several aquifers are superimposed, only the most superficial aquifer has been retained.

2.3. Data Treatment

To mitigate the influence of outliers, the data underwent a logarithmic transformation using the formula $y = \log_{10}(x + DL)$, where x represents the value of parameter X (whether physicochemical or bacteriological), and DL is its detection limit. The purpose of this transformation is to bring the distributions of the various parameters closer to a normal distribution, thereby reducing the impact of extreme values that could obscure certain processes responsible for the diversity of water quality in the dataset during analysis [20,24].

A principal component analysis (PCA) was conducted on the full matrix of logarithmically transformed data to reduce the dimensionality of the data hyperspace [19] and identify and rank the sources of water quality variability and the mechanisms associated with them [25]. The PCA is based on the correlation matrix, utilizing standardized variables, which allows for the comparison of parameters with different units, or even those without units, like pH. This analysis was performed by diagonalizing the correlation matrix, making the principal axes orthogonal, and thus representing independent processes.

The spatial structure of the parameters was examined through the construction and analysis of variograms, based on the sparse matrix, to leverage the maximum amount of data for mapping each parameter using kriging. Variograms depict the evolution of semivariance between pairs of points as a function of the distance between them [26]. In the context of the Sise-Eaux database, samples were collected at various locations (different catchments) and on different dates, with each catchment being sampled multiple times during the extraction period. The data therefore contain both spatial and temporal variability. To differentiate these two aspects, variograms were calculated in two ways [22]:

1. By using the entire dataset, thereby preserving both spatial and temporal variability (referred to as VarTot).
2. By averaging the values of each parameter at each sampling point (referred to as VarMean). This method helps to minimize the temporal variance component. However, the variability studied in this manner is not exclusively spatial, as the sampling dates were not synchronized across all catchments during the extraction period.

The experimental variograms, obtained under uniform conditions for all parameters (same number of points and the same number of distance classes), were fitted with models incorporating a nugget effect followed by a spherical structure. These models were then used to generate kriged maps. The characteristics of the variograms (range, sill, nugget effect) were compared across different parameters. To detect potential anisotropy in the distribution of parameter values, directional variograms were calculated with a 15° rotational increment.

An unsupervised Hierarchical Ascendant Classification (HAC) [27,28] was performed on the bacteriological and physicochemical parameters (using the full matrix), based on the principal axes derived from the PCA. The goal was to better understand the diversity and similarities in the distribution of parameters at the regional scale and to establish a typology of these parameters. All calculations were performed with the XLstat software 2019.2.2 (Addinsoft).

3. Results

3.1. Maps and Variograms

The descriptive statistics of the 22 parameters in the dataset are summarized in Table 1. The greatest variations are observed for the bacteriological parameters and turbidity.

Table 1. Descriptive statistics for the 22 parameters.

Parameter * (3569 Data)	Unit	Min	Max	Mean	Standard Deviation
EC	mS cm^{-1}	1.279	3.447	2.593	0.290
E. coli	n/100 mL	0.000	4.398	0.701	0.748
Enter.	n/100 mL	0.000	4.540	0.582	0.616
NH4	mg L^{-1}	−1.699	0.072	−1.392	0.281
As	µg L^{-1}	0.000	1.699	0.481	0.239
Na	mg L^{-1}	−0.503	2.617	0.609	0.351
Ca	mg L^{-1}	0.004	2.285	1.811	0.418
Mg	mg L^{-1}	−0.959	1.732	0.471	0.311
Cl	mg L^{-1}	−0.387	2.857	0.856	0.357
SO_4	mg L^{-1}	−0.398	2.205	0.961	0.369
HCO_3	mg L^{-1}	0.176	2.729	2.291	0.400
NO_3	mg L^{-1}	−2.000	1.979	0.921	0.543
Fe	µg L^{-1}	0.301	3.776	1.104	0.449
Mn	µg L^{-1}	0.000	3.020	0.926	0.536
B	mg L^{-1}	−3.222	1.206	−1.905	0.324
F	mg L^{-1}	−1.959	0.628	−1.120	0.256
NO_2	mg L^{-1}	−3.000	−0.107	−1.831	0.404
TOC	mg L^{-1}	−1.222	2.301	−0.064	0.355
Turbidity	NFU	−1.523	2.041	−0.215	0.575
Se	µg L^{-1}	0.000	0.903	0.431	0.217
Cd	µg L^{-1}	0.000	1.041	0.212	0.139
Ni	µg L^{-1}	0.000	2.196	0.565	0.278

* The values correspond to the calculation $y = \log_{10}(x + DL)$.

The maps of several major parameters (EC, As, *E. coli*, and Enter.) are shown in Figure 3. These maps, derived from several thousand catchment pairs, revealed contrasting regions with different distributions depending on the parameter.

Several representative examples of the spatiotemporal (VarTot, black curves) variograms obtained are shown in Figure 4. Notable differences were observed depending on the parameter. For electrical conductivity and major ions, the range was high, about 60 to 80 km. The nugget effect relative to the sill was low, ranging from approximately one-fifth to one-tenth of the maximum semivariance. The "spatial" variograms calculated from the mean values of each parameter at each sampling point (VarMean, red curves) allow us to estimate the impact of temporal variance by comparison. The differences were slight for major ions and EC, indicating low temporal variance. Directional variogram (not shown) analysis revealed slight anisotropy between the north–south and east–west directions. In contrast, the bacteriological parameters (*E. coli* and Enter.) exhibited a significant nugget effect, around two-thirds of the sill, and a reduced range of less than 5 km. These characteristics indicated a highly localized source of variability that differed markedly from the spatial structure of the major ions. The gap between VarTot and VarMean was significant (Figure 4g,h) where the sill values decreased considerably between the spatiotemporal and the spatial approach (VarMean), indicating high temporal variance. The same was true for water turbidity (not shown). The gap between VarTot and VarMean was also significant for metals, where the sill values decreased considerably between the spatiotemporal and the spatial approach (VarMean) (Figure 4e,f). The temporal variance was intermediate for nitrates (Figure 4c), organic carbon, and trace elements such as arsenic (Figure 4d).

The nugget effect relative to the sill was approximately one-third for nitrates, arsenic, and metals.

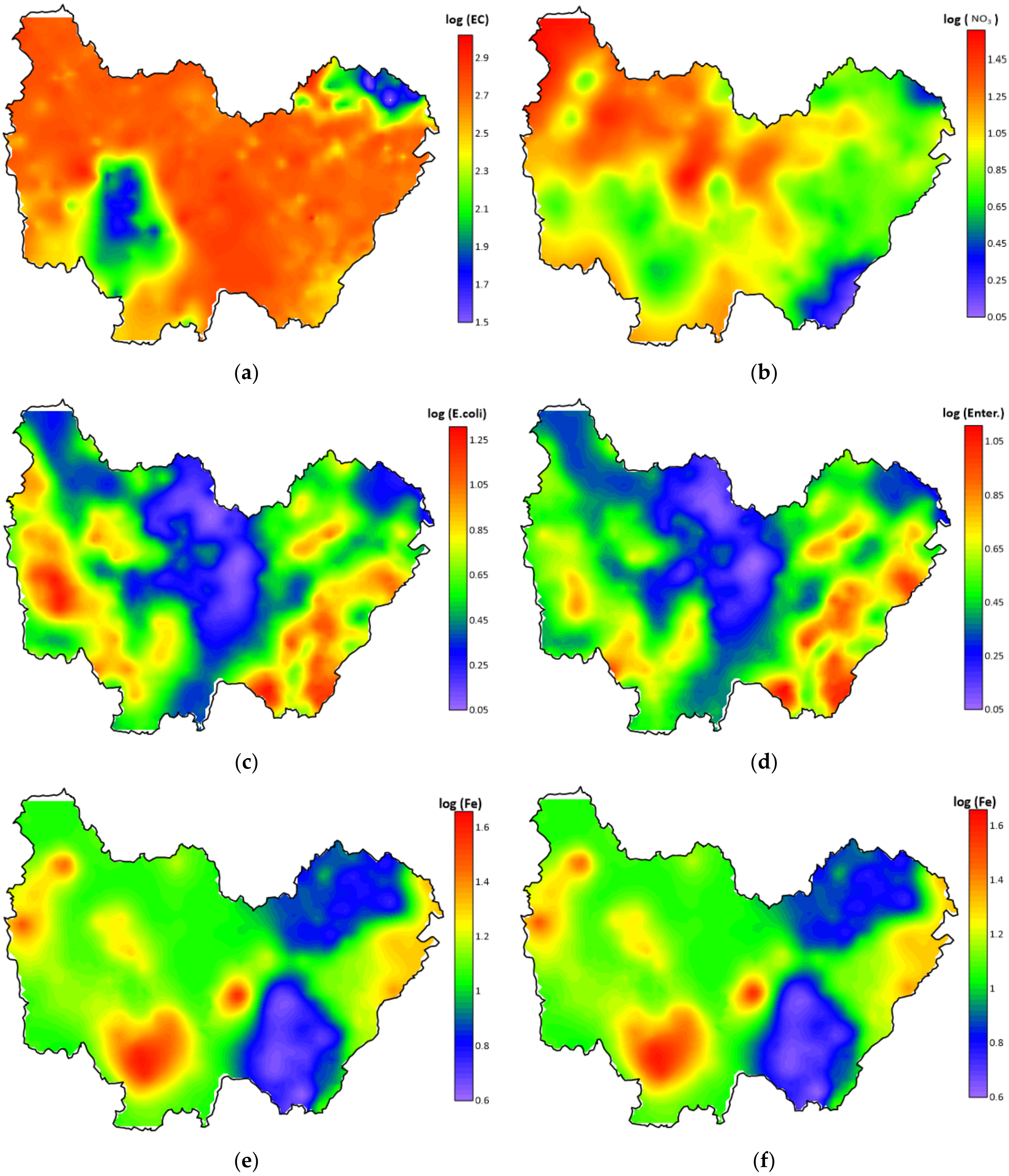
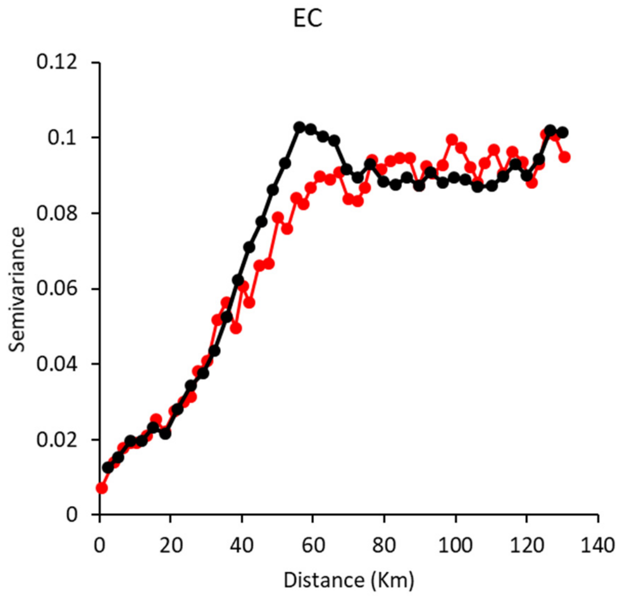
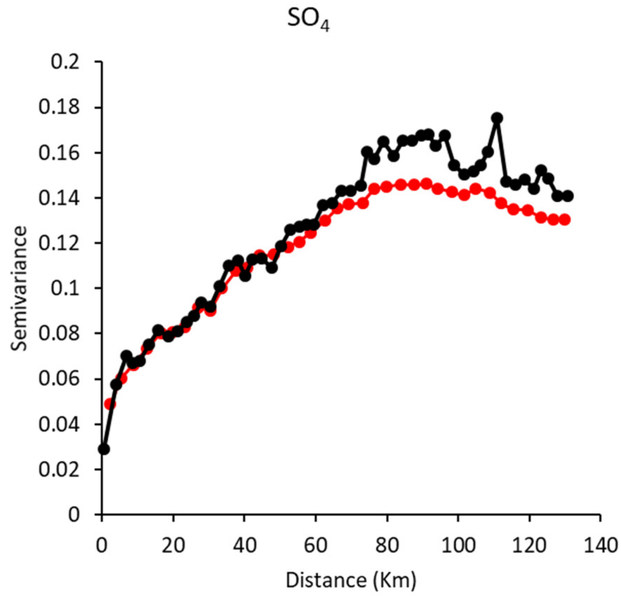


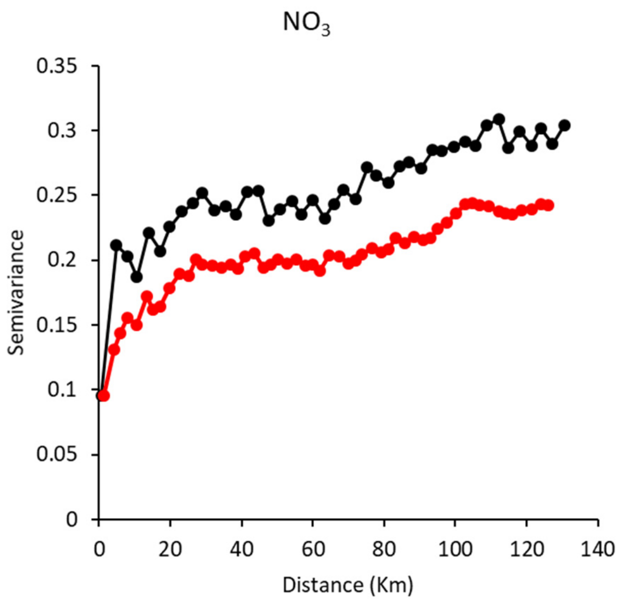
Figure 3. Distribution maps for parameters (**a**) EC (proportional to the sum of major ions), (**b**) NO_3 (sensitive to agricultural pollution), (**c**) *E. coli* and (**d**) Enter. (representing fecal contamination), (**e**) Fe (trace metal element), and (**f**) arsenic. All the maps were drawn using the variogram calculated based on all the data (VarTot). Units are μS cm^{-1} for EC, mg L^{-1} for major ions, nitrates, and traces, and number of cells per 100 mL for bacteria.



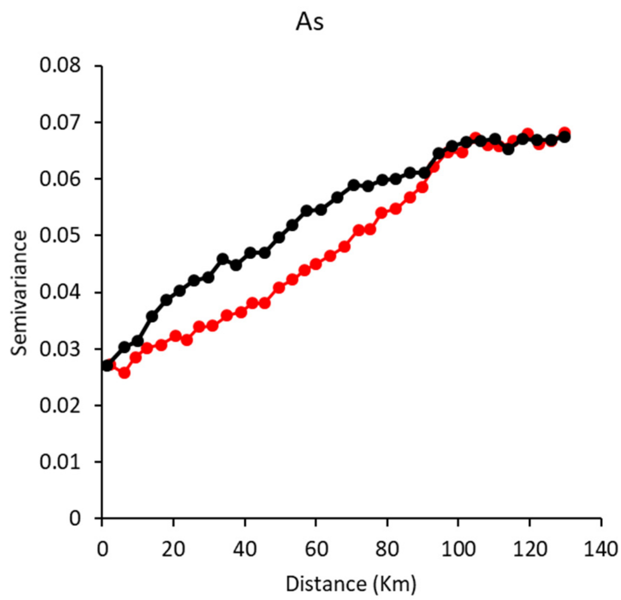
(**a**)



(**b**)



(**c**)



(**d**)

Figure 4. *Cont.*

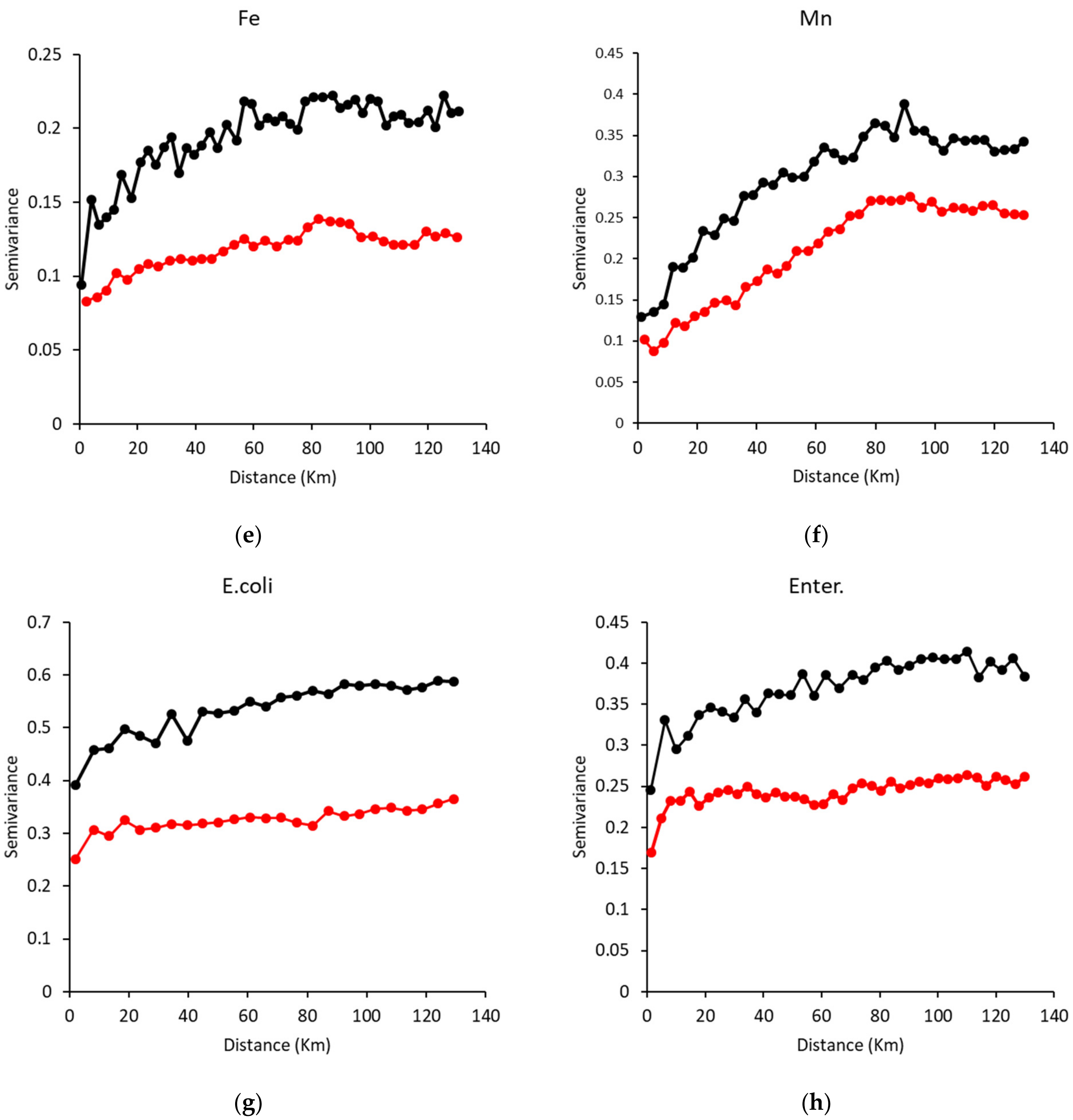


Figure 4. Examples of variograms obtained for (**a**,**b**) EC and major ions, (**c**) nitrates, (**d**) As, (**e**) Fe, (**f**) Mn, and (**g**,**h**) bacteriological parameters. For all the graphs, the black lines and symbols represent the variogram calculated using all the data (VarTot), and the red lines and symbols represent the variogram calculated using the mean of the parameter at each measurement point (VarMean).

3.2. Principal Component Analysis

The distribution of inertia across the principal components (PCs) is shown in Figure 5, and the contributions of different parameters to the primary PCs are detailed in Table 2. The inertia of the principal components revealed complex information, requiring the consideration of the first seven PCs to capture 80% of the information contained in the full matrix. Moreover, the first six principal components had eigenvalues greater than one, indicating that they accounted for more information than any single initial parameter, thus demonstrating significant dimensional reduction. The analysis of the first four principal components (Table 2), which explained between 22% and 9.1% of the variance and cumulatively accounted for 65% of the total information across all samples and parameters, is presented below. For the subsequent PCs, which carried less significant information, the analysis becomes more challenging [25].

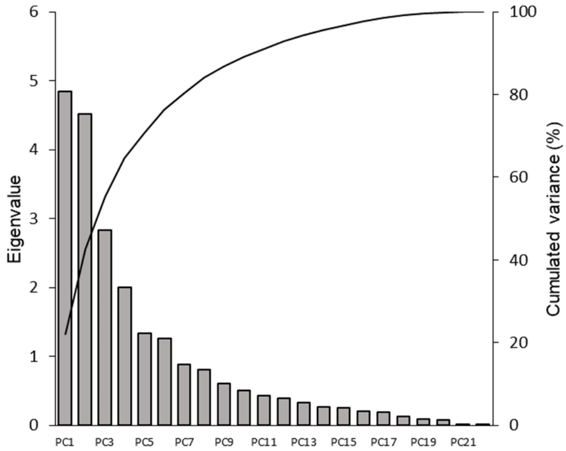


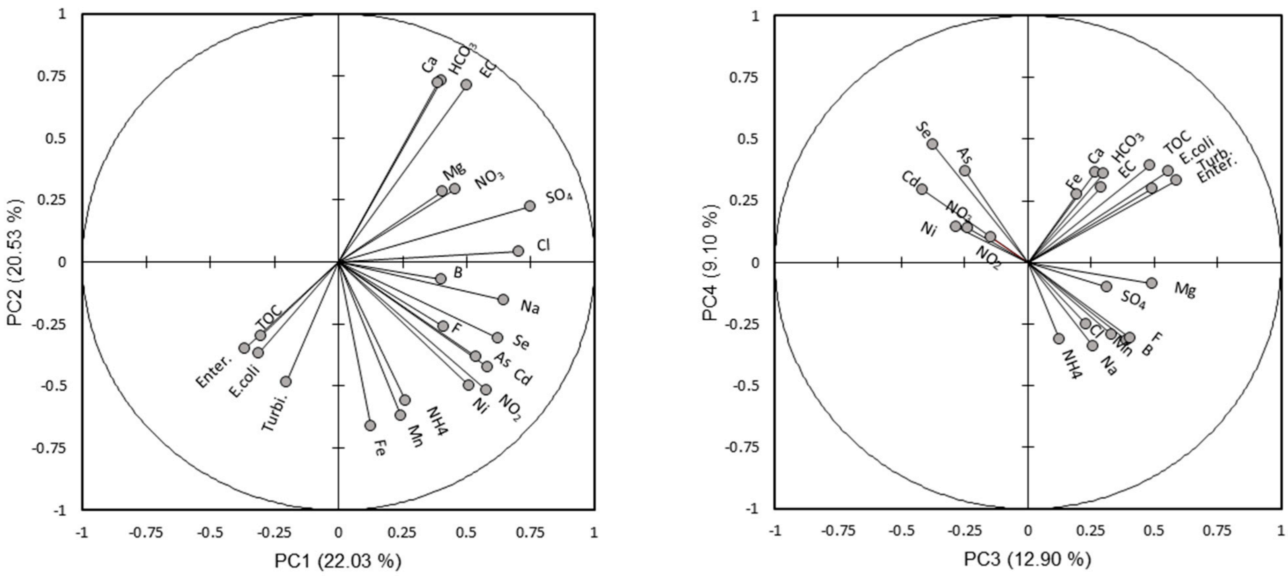
Figure 5. Inertia of the factorial axes of the PCA conducted on the log-transformed data.

Table 2. Contribution of the parameters (log-transformed data) to the first four factorial axes. The bold values represent the strongest positive and negative contributions.

Log-Transformed Parameter	PC1	PC2	PC3	PC4
EC	**0.500**	**0.712**	0.288	0.302
E. coli	**−0.314**	**−0.366**	**0.556**	**0.370**
Enter.	**−0.368**	**−0.352**	**0.585**	**0.334**
NH4	0.261	**−0.558**	0.124	**−0.311**
As	**0.536**	−0.381	−0.246	**0.370**
Na	**0.645**	−0.152	0.257	**−0.342**
Ca	0.402	**0.730**	0.267	**0.367**
Mg	0.404	0.284	**0.493**	−0.088
Cl	**0.703**	0.041	0.229	−0.252
SO_4	**0.747**	0.224	0.311	−0.102
HCO_3	0.385	**0.721**	0.296	**0.360**
NO_3	0.456	0.294	−0.236	0.138
Fe	0.127	**−0.659**	0.195	0.275
Mn	0.241	**−0.619**	0.331	**−0.293**
B	0.402	−0.069	0.388	**−0.317**
F	0.412	−0.260	**0.403**	**−0.307**
NO_2	**0.579**	**−0.518**	−0.146	0.102
TOC	−0.301	**−0.299**	**0.484**	0.392
Turb.	−0.205	**−0.483**	**0.493**	0.301
Se	**0.622**	**−0.306**	**−0.377**	**0.478**
Cd	**0.580**	**−0.425**	**−0.416**	0.294
Ni	**0.511**	**−0.496**	−0.285	0.144

The first principal component (Figure 6a), accounting for 22% of the total variance, was positively associated with mineralized waters that had a chloride–sulfate–sodium chemical profile. These waters were characterized by the presence of various forms of nitrogen, as well as high concentrations of heavy metals and arsenic. Areas with high values (Figure 7) were primarily found in the Bresse plain, Mâconnais, the Bourgogne plateaus, Auxois, and the southern part of Auxerrois, regions dominated by marls and limestones. In contrast, more diluted waters, marked by fecal contamination, were found in the Morvan massif, the Vosges, and the heights of the Jura. This axis reflected the contrast between mineralized groundwater, which is tapped during low-flow periods, and water that is diluted by surface runoff and contaminated by bacteria associated with suspended particles containing organic matter (TOC and turbidity were associated with fecal bacteria) in the sectors mentioned above. Deep waters were marked by reduced forms of nitrogen (NO_2 and, occasionally, NH4 parameters), while more oxidizing surface waters were marked by NO_3. There are redox processes affecting nitrates depending on the depth and residence time. The alteration of sulfide minerals contained in the local lithology (notably pyrites and arsenopyrites, etc.) releases SO_4^{2-}, Se, and As. The oxidation of these sulfide minerals consumes oxidizing agents (NO_3^-, dissolved O_2) and explains the reducing nature of the groundwater. Water of deep origin also contains Mn and Fe, confirming the reducing nature of these environments.

The second principal component, explaining 20.5% of the variance, contrasted calcium-carbonate-mineralized waters with diluted waters, which were also affected by fecal contamination and heavy metals. Notably, the diluted waters were associated with reduced forms of nitrogen (ammonium and nitrites) and metals (Fe, Mn) and metalloids such as selenium and arsenic, typical of metalliferous deposits. This redox gradient suggested that the diluted waters are likely subsurface waters, characterized by higher turbidity and total organic carbon (TOC) levels, similar to runoff waters during rainy events. The most negative values on this axis, indicating higher fecal and metal contamination, corresponded to the Jura massif, while the most positive values (Figure 7) were found in the alluvium of the Saône and Bresse plains and the elevated areas around the Morvan (Mâconnais, Autunois, Bourgogne plateaus, Bazois).



(**a**) (**b**)

Figure 6. Distribution of the parameters in the factorial plans (**a**) PC1–PC2 and (**b**) PC3–PC4.

Fecal contamination played a major role in the third factorial axis (12.9% of the variance), not associated with diluted waters as in the first two PCs but with mineralized waters, with high turbidity and TOC, low in nitrates and metals except for Fe and Mn, which are likely surface markers in the form of metallic colloids (Figure 6b). Numerous clusters were observed in the alluvial plains, south of the Vosges massif, on the heights of the Doubs and Saône valleys, on the Jura massif, and to the west of the region (between Auxerre and Nevers, Figure 7).

Similar to the third axis, the fourth factorial axis (9.1% of the variance) associated fecal contamination with high turbidity and TOC, positively correlated with the mineral content (EC) and a calcium carbonate chemical profile. This axis only concerned a few isolated clusters, north of the Bazois and in the upper Saône valley (Figure 7). These last two factorial axes thus reflected a lithological opposition.

The contribution of fecal contamination parameters to the four main factorial axes, collectively accounting for nearly 65% of the total variance, highlighted the complexity and multifactorial nature of the relationship between fecal contamination and groundwater characteristics.

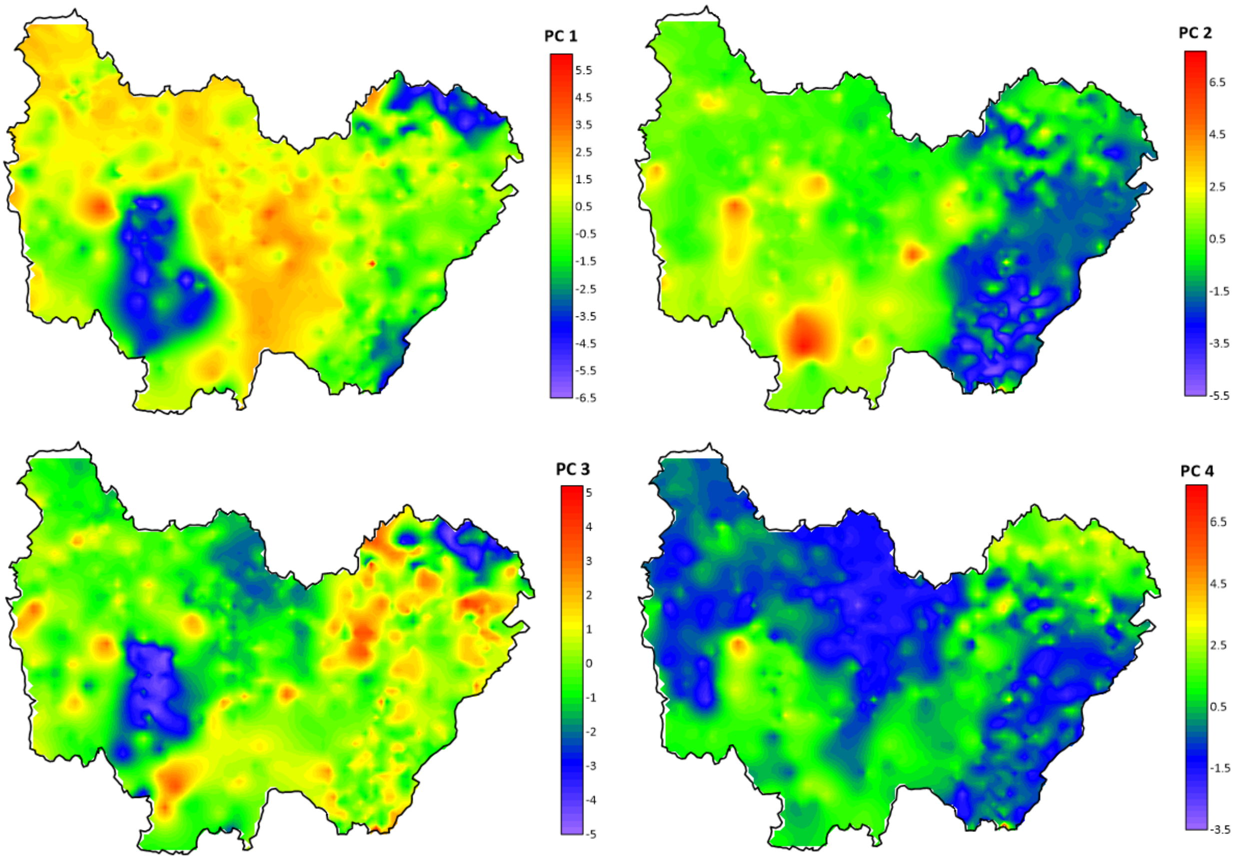


Figure 7. Distribution of the first four factorial axes across the Bourgogne-Franche-Comté region.

The result of the clustering performed on the 22 parameters, based on the first 10 principal components, which account for 90% of the information carried by the dataset, is presented in Figure 8. The positioning of the phenom line allowed for the distinction of six groups. The first group consisted of metals and metalloids, except for manganese. The second group included major ions and fluorides, except for calcium ions and carbonates. The third group comprised different forms of nitrogen and manganese. Boron was an isolated parameter. The fifth group linked electrical conductivity with the major ions Ca and HCO_3, and the last group was related to fecal contamination, always associated with the turbidity and TOC.

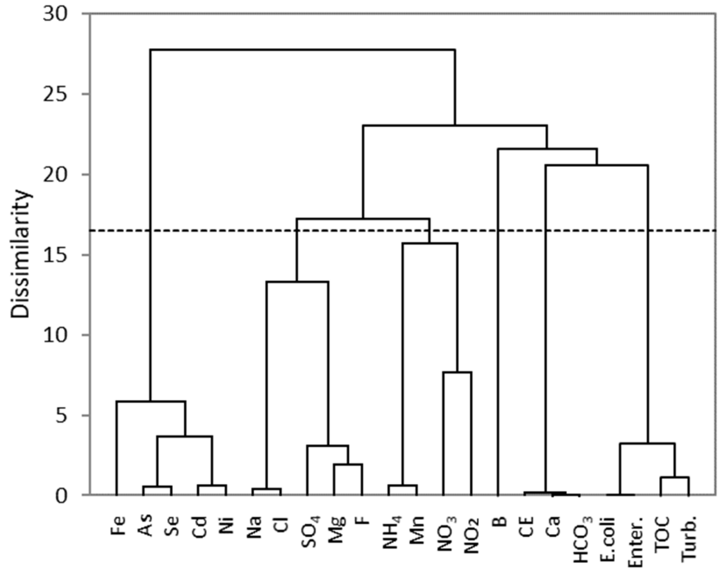


Figure 8. Clustering of the parameters into six groups.

4. Discussion

4.1. Diversity of Natural Environments

The multi-parameter approach, rather than the parameter-by-parameter procedure typically followed by the Regional Health Agencies, is based on correlations between parameters and highlights common sources of variation. In this sense, the results reflect the ongoing processes within the region's groundwater. The diversity of natural environments is reflected in the results of the principal component analysis (PCA), with at least seven principal components required to explain 80% of the total variance in the dataset. This result indicates not only a diversity of lithology (Figure 2) but also variations in altitude (Figure 1), types of groundwater resources, climate, and human activity. It also reflects the diversity of mechanisms involved in acquiring the chemical and bacteriological composition of the water. The hierarchical classification (Figure 8) is based on the PCA results, meaning it takes into account essential information after eliminating statistical noise.

4.2. Local or Regional Determinants of Water Quality

The cartographic representation based on the logarithmic transformation of electrical conductivity (Figure 3a) reveals a structure that coincides with the major structural units. The least mineralized waters are observed in the crystalline bedrock of the Morvan and Vosges regions, as well as in the schists, sandstones, and arkoses of the Permian basin in Autunois, while the most mineralized waters are found in the alluvial valleys and in sedimentary formations from the Jurassic, Cretaceous, and Miocene periods (Figure 2). Besides the geographic and lithological coincidence between major ions/conductivity and the major structural units, variograms of these parameters show a range of about 60 to 80 km and a low nugget effect, which also corresponds to the size of the major structural units in the region. This regional determinism affects the parameters of groups 2 and 5 in the classification (Figure 8). In detail, the major structural units, from east to west, include the Vosges (about 50 km wide), the folded covers of the Jura (about 50 to 60 km), the Saône plain (about 80 km) and its tributaries, the Morvan massif and its surroundings (about 80 km), and then the reliefs of the western region cut by the valleys of the Loire and Seine basins (Auxerrois and Bazois, about 80 to 100 km). Although the boundaries between these major units show a certain north–south orientation, this orientation is less pronounced than in the neighboring Auvergne-Rhône-Alpes region [22], where a very clear preferential direction had been detected on directional variograms.

The distribution of total mineralization is consistent with results generally observed in other regions of France [19–23], indicating that it depends primarily on the solubility and alterability of the rocks, increasing along flow paths due to a longer interaction between the water and rock. It is also noteworthy that there is continuity to the south in the chemical profile between the folded covers of the Jura and the Pre-Alps, the Saône plain and the Rhone corridor, and the Morvan and the Massif Central [22]. Notably, the most mineralized waters in the region are found in the Saône plain and Bresse (Figure 1), with a chemical profile rich in chlorides, sulfates, and sodium, a chemical signature that aligns well with the waters of the Rhone corridor extending southward [29].

In contrast, for bacteriological parameters, the significant proportion of the nugget effect in the total semivariance and the short range of the variograms indicate that the determinism of fecal contamination is rather local in nature. It is worth noting that the analysis, which includes turbidity and TOC parameters—parameters that were not considered in previous studies on other French regions—confirms the close relationship (Figure 6) between turbidity and the transport of bacteria (*E. coli* and Enter.), also confirming that part of the turbidity is of organic origin [30–32]. The extraction of the Sise-Eaux database for the region highlights the complexity of the determinism of fecal contamination, which contributes to the first four principal components. These principal components are orthogonal to each other and theoretically represent independent processes. Such a result has already been observed in other French regions [20,33]. The general conclusion drawn from this observation is that fecal contamination occurs in different contexts. This is also true here, where the first two principal components reflect chemical facies contrasts in the waters, with rainwater dilution leading to higher turbidity. Comparing the spatiotemporal variograms (VarTot) and spatial variograms (VarMean) allows for further analysis.

4.3. Spatial vs. Temporal Variability

Let us recall here that the database includes both spatial and temporal variance. As the parameter typology (Figure 8) is based on PCA axes, it incorporates both spatial and temporal variability within the dataset. During the spatial analysis, the weight of the temporal variance was reduced by considering the averages of each parameter at each sampling point (VarMean). This procedure allows for a reduction in temporal variability, but some remains, as not all samples were taken synchronously at all catchments. Therefore, a residual temporal component remains in the spatial variograms. Only the proportion of the nugget effect in the total semivariance should be considered as strictly spatial information, reflecting the significance of the temporal variance. The two approaches, statistical by PCA and AHC and geostatistical by the analysis of the structure of the variograms, are independent but converge in terms of interpretation.

For the parameters in groups 2 and 5 (major ions, Figure 8), the difference between the two variograms is small, which was interpreted as a low portion of temporal variance. However, this variance is underestimated due to the samples being taken on different dates. To achieve an absolute separation between spatial and temporal variance, only samples taken on the same date would have to be considered, which would have significantly reduced the number of sampling points and prevented the application of geostatistical techniques for analysis. The contrast in the mineral content in the principal components (notably the first two) reflects dilution based on rainfall and seasons, confirming the presence of a seasonal (and thus temporal) component in the concentrations of major ions.

For microbiological parameters, the turbidity, and the TOC (group 6 parameters, Figure 8), the significant difference between the spatial (VarMean) and spatiotemporal (VarTot) variograms indicates the major temporal nature of fecal contamination. The commonly advanced explanation is the impact of late-summer thunderstorms, which are accompanied by runoff and, depending on the level of protection of the sampling points, the potential for water contamination by this runoff [19,22,23].

The difference between the spatial and spatiotemporal variograms increases with distance, with the curves diverging (Figure 4g,h). This suggests that the vulnerability of sampling points to bacterial contamination has a spatial dimension. This fact can reflect several aspects:

- Different soil types, including the presence of flocculating ions that limit water turbidity and bacterial transport, or different textures that provide more (fine texture) or less (coarse texture) protection to the sampling points from contaminated surface water intrusions.
- Unevenly distributed contamination pressure across the territory. Livestock farming, a major contributor to contamination, is particularly concentrated in mountainous areas and on high ground and is much less present in the plains.

This highlights that even processes with a strong temporal component, such as bacterial contamination, have a spatial dimension that can be observed when processing the data.

4.4. Consequences for Sustainable Management

In France, about 95% of non-compliance cases related to the public water supply are due to fecal contamination issues. This is a major concern for the Health Agencies responsible for monitoring and surveillance, and understanding the temporal and spatial nature of contamination is essential. On this regional scale and based on thousands of data points, the relationship with the turbidity and TOC, the main vectors of contamination, is confirmed, as well as the role of momentary phenomena (such as storms) despite a spatial structure in the vulnerability of the catchment points. Beyond the Bourgogne-Franche-Comté region discussed here, extracting such databases—assuming they exist—highlights several solutions for the targeted multi-parameter monitoring of the resource:

- Developing parameter maps, a procedure traditionally followed until now.
- Performing dimensional reduction and tracking the principal components, which provides a synthetic view of the independent sources of quality variation. In the case discussed here, the synthetic approach to fecal contamination through PCA shows that the two main contamination mechanisms are linked to the inflow of surface water, particularly during rainy episodes that generate runoff, in two different lithological zones (PCs 1 and 2). The two mechanisms represented by the subsequent PCs (3 and 4) are related to the regional contrast between two distinct lithological areas. The comprehensive information enables targeting areas for protection with mechanisms specific to each.
- After dimensional reduction and grouping parameters according to their behavior, selecting a representative from each group to monitor—a procedure that is more familiar to the agents responsible for monitoring.

These proposals should enable the more targeted monitoring of local issues related to specific lithological environments, land uses, etc., while also streamlining the monitoring system, achieving cost savings in analysis, and simplifying the monitoring and experimental protocol to avoid the duplication of information, which can be costly in terms of both analytical and logistical budgets.

5. Future Directions

Knowledge of groundwater diversity is a key aspect of good resource management. This study, based on an existing database (Sise-Eaux), is a contribution to that knowledge. As mentioned above, databases on groundwater quality on a national or even regional scale are rare and often concern data collected over short to medium periods. The existence of such a database, collected over several decades throughout France, is a godsend for understanding the diversity of groundwater but also for developing a method for analyzing the information it contains, aiming at the targeted monitoring and sustainable management of the resource. This work is a further step in the analysis of such a source of information. Having shown, in previous articles, the role of dimensional reduction, data conditioning, and cross-referencing with other independent databases, such as the compartmentalization of groundwater bodies, it now makes a start on separating the spatial and temporal variances within the datasets, which makes it possible to refine their interpretation. However, this effort needs to be deepened by multi-variate analyses that specifically preserve only the temporal or spatial variance of the datasets. The comparison of the results should make it possible to further refine and better distinguish and identify the processes responsible for water quality.

Author Contributions: Conceptualization, A.B. (Abdelhak Bouabdli) and V.V.; methodology, V.V., L.B. and A.B. (Abderrahim Bousouis); software, A.B. (Abderrahim Bousouis); validation, A.B. (Abderrahim Bousouis), M.A. and L.B.; formal analysis, A.B. (Abderrahim Bousouis) and L.R.; investigation, A.B. (Abderrahim Bousouis); resources, L.R.; data curation, A.B. (Abderrahim Bousouis) and L.R.; writing—original draft preparation, A.B. (Abderrahim Bousouis) and V.V.; writing—review and editing, L.B.; visualization, A.B. (Abderrahim Bousouis) and L.R.; supervision, A.B. (Abdelhak Bouabdli) and V.V.; project administration, A.B. (Abdelhak Bouabdli); funding acquisition, L.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets presented in this article are not easily accessible for reasons of sensitivity to possible malicious acts. Requests for access to the datasets should be addressed to the Health Agency ARS of the Bourgogne-Franche-Comté region.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gleeson, T.; Cuthbert, M.; Ferguson, G.; Perrone, D. Global Groundwater Sustainability, Resources, and Systems in the Anthropocene. *Annu. Rev. Earth Planet. Sci.* **2020**, *48*, 431–463. [CrossRef]
2. Elshall, A.S.; Arik, A.D.; El-Kadi, A.I.; Pierce, S.; Ye, M.; Burnett, K.M.; Wada, C.A.; Bremer, L.L.; Chun, G. Groundwater sustainability: A review of the interactions between science and policy. *Environ. Res. Lett.* **2020**, *15*, 93004. [CrossRef]
3. Deng, Z.; Ma, Q.; Zhang, J.; Feng, Q.; Niu, Z.; Zhu, G.; Jin, X.; Chen, M.; Chen, H. A New Socio-Hydrology System Based on System Dynamics and a SWAT-MODFLOW Coupling Model for Solving Water Resource Management in Nanchang City, China. *Sustainability* **2023**, *15*, 16079. [CrossRef]
4. Pandey, V.P.; Shrestha, S.; Chapagain, S.K.; Kazama, F. A framework for measuring groundwater sustainability. *Environ. Sci. Policy* **2011**, *14*, 396–407. [CrossRef]
5. Li, S.; Wang, Z.; Lai, C.; Lin, G. Quantitative assessment of the relative impacts of climate change and human activity on flood susceptibility based on a cloud model. *J. Hydrol.* **2020**, *588*, 125051. [CrossRef]
6. Ahmed, W.; Ahmed, S.; Punthakey, J.F.; Dars, G.H.; Ejaz, M.S.; Qureshi, A.L.; Mitchell, M. Statistical Analysis of Climate Trends and Impacts on Groundwater Sustainability in the Lower Indus Basin. *Sustainability* **2024**, *16*, 441. [CrossRef]
7. Chen, P.; Ma, R.; Shi, J.; Si, L.; Zhao, L.; Wu, J. Ecological Risks Arising in the Regional Water Resources in Inner Mongolia Due to a Large-Scale Afforestation Project. *Sustainability* **2023**, *15*, 16091. [CrossRef]
8. Beaudeau, P.; Pascal, M.; Mouly, D.; Galey, C.; Thomas, O. Health risks associated with drinking water in a context of climate change in France: A review of surveillance requirements. *J. Water Clim. Chang.* **2011**, *2*, 230–246. [CrossRef]
9. Mays, L.W. Groundwater Resources Sustainability: Past, Present, and Future. *Water Resour. Manag.* **2013**, *27*, 4409–4424. [CrossRef]
10. Kemper, K.E. Groundwater—From development to management. *Hydrogeol. J.* **2004**, *12*, 3–5. [CrossRef]
11. Yuan, H.; Yang, S.; Wang, B. Hydrochemistry characteristics of groundwater with the influence of spatial variability and water flow in Hetao Irrigation District, China. *Environ. Sci. Pollut. Res.* **2022**, *29*, 71150–71164. [CrossRef] [PubMed]
12. Gao, Y.; Chen, J.; Qian, H.; Wang, H.; Ren, W.; Qu, W. Hydrogeochemical characteristics and processes of groundwater in an over 2260 year irrigation district: A comparison between irrigated and nonirrigated areas. *J. Hydrol.* **2022**, *606*, 127437. [CrossRef]
13. Greene, R.; Timms, W.; Rengasamy, P.; Arshad, M.; Cresswell, R. Soil and Aquifer Salinization: Toward an Integrated Approach for Salinity Management of Groundwater. In *Integrated Groundwater Management: Concepts, Approaches and Challenges*; Jakeman, A.J., Barreteau, O., Hunt, R.J., Rinaudo, J.-D., Ross, A., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 377–412, ISBN 978-3-319-23576-9.
14. Quevauviller, P.; Batelaan, O.; Hunt, R.J. Groundwater Regulation and Integrated Water Planning. In *Integrated Groundwater Management: Concepts, Approaches and Challenges*; Jakeman, A.J., Barreteau, O., Hunt, R.J., Rinaudo, J.-D., Ross, A., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 197–227, ISBN 978-3-319-23576-9.
15. Eamus, D.; Fu, B.; Springer, A.E.; Stevens, L.E. Groundwater Dependent Ecosystems: Classification, Identification Techniques and Threats. In *Integrated Groundwater Management: Concepts, Approaches and Challenges*; Jakeman, A.J., Barreteau, O., Hunt, R.J., Rinaudo, J.-D., Ross, A., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 313–346, ISBN 978-3-319-23576-9.
16. Gran-Aymeric, L. Un portail national sur la qualite des eaux destinees a la consommation humaine. *Tech. Sci. Méthodes* **2010**, *12*, 45–48. [CrossRef]

17. Gallay, A.; De Valk, H.; Cournot, M.; Ladeuil, B.; Hemery, C.; Castor, C.; Bon, F.; Mégraud, F.; Le Cann, P.; Desenclos, J.C. A large multi-pathogen waterborne community outbreak linked to faecal contamination of a groundwater system, France, 2000. *Clin. Microbiol. Infect.* **2006**, *12*, 561–570. [CrossRef] [PubMed]
18. Chery, L.; Laurent, A.; Vincent, B.; Tracol, R. *Echanges SISE-Eaux/ADES: Identification des Protocoles Compatibles Avec les Scénarios D'échange SANDRE*; Onema: Vincennes, France; BRGM: Orléans, France, 2011.
19. Tiouiouine, A.; Yameogo, S.; Valles, V.; Barbiero, L.; Dassonville, F.; Moulin, M.; Bouramtane, T.; Bahaj, T.; Morarech, M.; Kacimi, I. Dimension reduction and analysis of a 10-year physicochemical and biological water database applied to water resources intended for human consumption in the provence-alpes-cote d'azur region, France. *Water* **2020**, *12*, 525. [CrossRef]
20. Mohsine, I.; Kacimi, I.; Abraham, S.; Valles, V.; Barbiero, L.; Dassonville, F.; Bahaj, T.; Kassou, N.; Touiouine, A.; Jabrane, M.; et al. Exploring Multiscale Variability in Groundwater Quality: A Comparative Analysis of Spatial and Temporal Patterns via Clustering. *Water* **2023**, *15*, 1603. [CrossRef]
21. Jabrane, M.; Touiouine, A.; Valles, V.; Bouabdli, A.; Chakiri, S.; Mohsine, I.; El Jarjini, Y.; Morarech, M.; Duran, Y.; Barbiero, L. Search for a Relevant Scale to Optimize the Quality Monitoring of Groundwater Bodies in the Occitanie Region (France). *Hydrology* **2023**, *10*, 89. [CrossRef]
22. Ayach, M.; Lazar, H.; Bousouis, A.; Touiouine, A.; Kacimi, I.; Valles, V.; Barbiero, L. Multi-Parameter Analysis of Groundwater Resources Quality in the Auvergne-Rhône-Alpes Region (France) Using a Large Database. *Resources* **2023**, *12*, 143. [CrossRef]
23. Lazar, H.; Ayach, M.; Barry, A.; Mohsine, I.; Touiouine, A.; Huneau, F.; Mori, C.; Garel, E.; Kacimi, I.; Valles, V.; et al. Groundwater bodies in Corsica: A critical approach to GWBs subdivision based on multivariate water quality criteria. *Hydrology* **2023**, *10*, 213. [CrossRef]
24. Jabrane, M.; Touiouine, A.; Bouabdli, A.; Chakiri, S.; Mohsine, I.; Valles, V.; Barbiero, L. Data Conditioning Modes for the Study of Groundwater Resource Quality Using a Large Physico-Chemical and Bacteriological Database, Occitanie Region, France. *Water* **2023**, *15*, 84. [CrossRef]
25. Helena, B.; Pardo, R.; Vega, M.; Barrado, E.; Fernandez, J.M.; Fernandez, L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Res.* **2000**, *34*, 807–816. [CrossRef]
26. Cressie, N. The origins of kriging. *Math. Geosci.* **1990**, *22*, 239–252. [CrossRef]
27. Day, W.H.E.; Edelsbrunner, H. Efficient algorithms for agglomerative hierarchical clustering methods. *J. Classif.* **1984**, *1*, 7–24. [CrossRef]
28. Bouguettaya, A.; Yu, Q.; Liu, X.; Zhou, X.; Song, A. Efficient agglomerative hierarchical clustering. *Expert Syst. Appl.* **2015**, *42*, 2785–2797. [CrossRef]
29. Ayach, M.; Lazar, H.; Lamat, C.; Bousouis, A.; Touzani, M.; El Jarjini, Y.; Kacimi, I.; Valles, V.; Barbiero, L.; Morarech, M. Groundwaters in the Auvergne-Rhône-Alpes Region, France: Grouping Homogeneous Groundwater Bodies for Optimized Monitoring and Protection. *Water* **2024**, *16*, 869. [CrossRef]
30. Boithias, L.; Choisy, M.; Souliyaseng, N.; Jourdren, M.; Quet, F.; Buisson, Y.; Thammahacksa, C.; Silvera, N.; Latsachack, K.; Sengtaheuanghoung, O.; et al. Hydrological Regime and Water Shortage as Drivers of the Seasonal Incidence of Diarrheal Diseases in a Tropical Montane Environment. *PLoS Neglected Trop. Dis.* **2016**, *10*, e0005195. [CrossRef]
31. Abbas, A.; Baek, S.; Silvera, N.; Soulileuth, B.; Pachepsky, Y.; Ribolzi, O.; Boithias, L.; Cho, K.H. In-stream Escherichia coli modeling using high-temporal-resolution data with deep learning and process-based models. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 6185–6202. [CrossRef]
32. Pachepsky, Y.A.; Shelton, D.R. Escherichia Coli and Fecal Coliforms in Freshwater and Estuarine Sediments. *Crit. Rev. Environ. Sci. Technol.* **2011**, *41*, 1067–1110. [CrossRef]
33. Mohsine, I.; Kacimi, I.; Valles, V.; Leblanc, M.; El Mahrad, B.; Dassonville, F.; Kassou, N.; Bouramtane, T.; Abraham, S.; Touiouine, A.; et al. Differentiation of multi-parametric groups of groundwater bodies through Discriminant Analysis and Machine Learning. *Hydrology* **2023**, *10*, 230. [CrossRef]