metals

Article

Analysis of Heavy Metal Pollution in Soil along the Shuimo River by the Grey Relational Method and Factor Analysis

Yidan Zhang [1], Yonglan Wang [1], Honggang Zang [1], Junqin Yao [1,2] and Huiying Ma [1,2,*]

[1] College of Ecology and Environment, Xinjiang University, Urumqi 830017, China; zyd@stu.xju.edu.cn (Y.Z.)
[2] Key Laboratory of Oasis Ecological Education, Xinjiang University, Urumqi 830017, China
* Correspondence: mahuiying@xju.edu.cn

Abstract: Soil samples were collected from the upstream, midstream and downstream areas of the Shuimo River in three layers of 0–20, 20–40 and 40–60 cm, and each group of sample points was located 5 m, 1 km and 2 km away from the river bank, respectively. The analysis was carried out. Based on the investigation and research, six indicators, including As, Pb, Zn, Cu, Ni and Cr, were designated as evaluation factors in combination with the results of the sample collection with low or no detectable values of Cd and Hg. The mean values of the samples measured in the upper, middle and downstream layers were taken, and the degree and source of pollution were evaluated and jointly analyzed using the gray correlation analysis and factor analysis methods. By using the gray correlation analysis, it was found that the evaluation results of the upstream and middle reaches of the soil along the Shuimo River were both level 3, with slight pollution, and the evaluation results of the downstream areas were level 2, with good soil quality. There are two main sources of pollution obtained through the factor analysis; source 1 is mainly heavy metals such as Zn, Cu, Cr, Pb and Ni, while source 2 is mainly heavy metals such as As, Pb and Ni. The amount of pollution sources is inferred from the heavy metal types of each source and the soil environment along the Shuimo River as industrial and human sources of pollution. From the analysis results, the combination of the gray correlation analysis model and factor analysis model is convenient and fast and can accurately quantify the source contribution of various pollution sources. Not only can it reflect the actual situation more objectively and realistically in the evaluation of soil heavy metal pollution and pollution sources, but also the calculation is simple and easily applied with low data requirements.

Keywords: soil along the river; soil heavy metals; analysis of factors; grey correlation analysis

check for updates

Citation: Zhang, Y.; Wang, Y.; Zang, H.; Yao, J.; Ma, H. Analysis of Heavy Metal Pollution in Soil along the Shuimo River by the Grey Relational Method and Factor Analysis. *Metals* **2023**, *13*, 878. https://doi.org/10.3390/met13050878

Academic Editor: Srecko Stopic

Received: 30 March 2023
Revised: 20 April 2023
Accepted: 27 April 2023
Published: 2 May 2023

1. Introduction

Soil is an important part of the ecological environment and is the basis for human existence. With the rapid development of the Chinese economy and the improvement of living standards in recent years, the problem of heavy metal pollution in soil has become more and more prominent [1]. The characteristics of heavy metal pollution include its wide range of sources, refractory degradation, easy enrichment, concealment, etc. that affects the growth of crops and is transferred to the human body through the food chain, thereby adversely affecting human health [2]. The Shuimo River is one of the water systems in Urumqi, and there are more than 200 industrial enterprises, including electric power, coal mines, chemicals, textiles, and construction factories, along its banks. A large number of pollutants emitted through human activities, such as industrial and agricultural production, urban life, transportation, etc., are continuously aggravating the pollution of the watershed, which makes the control mechanism of soil heavy metal pollution in this area unique and complex [3]. Therefore, it is of great significance to study heavy metal pollution in the soil along the Shuimo River [4].

At present, there is a great amount of research and evaluation methods undertaken by scholars at home and abroad on soil heavy metals, and the most widely used mathematical evaluation methods are mathematical statistics, the fuzzy comprehensive evaluation method, pollution index method, analytic hierarchy process, grey relational analysis method, etc. [5–7]. In this paper, the combination of the grey relational analysis model and the factor analysis model is used to evaluate the heavy metal pollution in the soil along the Shuimo River. These methods focus on calculating the correlation degree between the soil samples to be evaluated and the soil environmental quality classification standards, and they can obtain more objective and accurate evaluation results under the condition of reducing the influence of human factors [8]. It is important to directly investigate the sources of pollution through survey research, but it requires a large amount of manpower, material, and financial resources. It is impossible to accurately quantify the contributions of various pollution sources, so the factor analysis model is not only convenient and quick, but it can also accurately quantify the contributions of the various pollution sources.

2. Materials and Methods

2.1. Background of the Study Area

The Shuimo River is located at the northern foot of Tianshan Mountain, northeast of Urumqi City, originating from the low mountain belt on the eastern and northern slopes of Bogda. It is formed by the collection of groundwater and snow water from Shuimogou, Jianquangou, and Yushugou in Dongshan, and from the Shuimogou and Midong Districts. It flows into Tower Bridge Bay Reservoir, and finally flows into Qinggeda Lake. The main section of the river is situated 24 km from the original enamel factory near Taqiaowan Reservoir, and 36 km from Qinggeda Lake. The total area of the basin is 281.4 km^2 and the area of the water source protection area of the basin is 45.7 km^2 [9]. The terrain of the Shuimo River Basin is generally stepped, high in the south and low in the north. From the southeast to the north, it flows through Hongqiao, Qidaowan, Badaowan, Jiudaowan, and Kaziwan, and through other low mountain foothills, and then into the plain area of Midong District [10]. The basin belongs to a temperate inland arid climate zone, where it is hot in the summer and cold in the winter, has an annual average temperature of 4.2 °C with sufficient sunlight and little precipitation. The annual precipitation is 200–400 mm and the annual average precipitation is 227.9 mm. From north to south, with the increase in altitude, precipitation tends to increase [11].

The Shuimo River is an important water source for industrial production, agricultural irrigation, and residential water in the Shuimogou and Midong Districts of Urumqi City [12]. The upper and middle reaches are important industrial areas in Urumqi. Although large enterprises, such as Qifang and Weihuliang power plants, have been diverted or moved out, many small factories still densely occupy the river banks, and waste and muck are piled along the banks without any cover. The Qidaowan and Midong districts in the middle and lower reaches are important agricultural irrigation areas, mainly producing grain and vegetables, and their irrigation water mainly comes from the Shuimo River [13,14].

2.2. Sample Collection and Processing

Soils along the Shuimo River were sampled at 2 km intervals in the direction of the river, with a total of 30 sampling points set up from the north to Lake Qingda, and 12 sampling points on the opposite side of the river were also sampled at the same time, making a total of 42 soil sampling points. Each sample point was set up in the direction of the vertical structure of the soil, with a total of 3 layers of profiles from 0 to 20 cm, 20 to 40 cm and 40 to 60 cm at a depth of 20 cm, respectively, to prevent contamination of the soil. In some cases, only the top layer or two layers were collected because the soil texture and layers were too hard. In order to highlight the representativeness of the soil samples, the middle part of each layer is usually chosen. A total of 122 soil samples were collected in sealed bags, plus a cloth bag for record keeping. The soil samples were collected from the field and immediately returned to the laboratory on a separate flat sheet, stripped of leaves, weeds, debris and other foreign matter, and placed in a ventilated, sunny and light-protected area to dry naturally. The air-dried samples were crushed by rolling with wooden sticks, passed through a 10 mesh (2 mm) nylon sieve and the rest of the air-dried samples were ground in an agate mortar and then passed through a 100 mesh (0.147 mm) aperture nylon sieve (avoiding contact with any metal products throughout the sample), mixed in quarters and bagged. Samples were collected from 11 types of land use, including industrial areas, residential areas, woodland, agricultural land, vegetable land, roadside bare ground, artificial grassland, mountainous areas, and water outfalls, with sampling densities that illustrate the differences in the effects of different pollution sources on coastal soils. Each sample point was precisely located using GPS. Figure 1 shows the distribution of the sampling points.

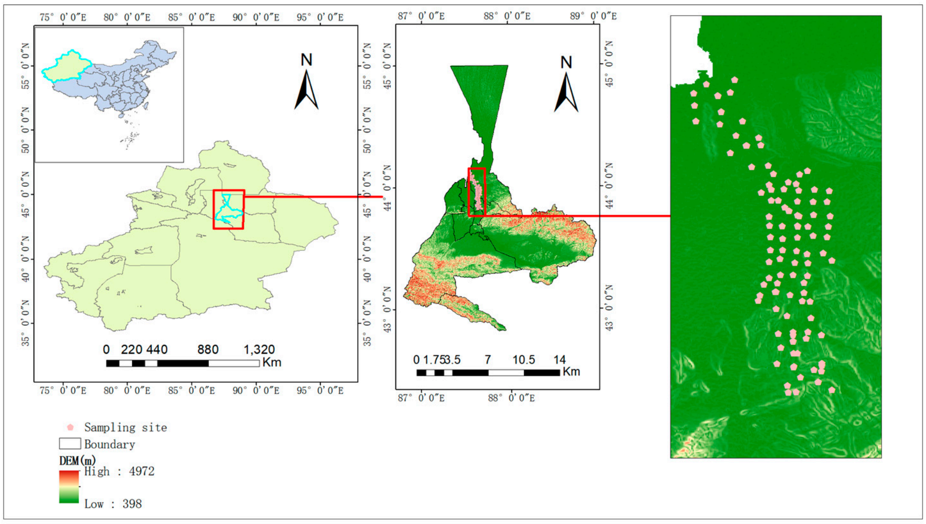

Figure 1. Distribution of sampling points.

The soil samples were tested using the following two methods: rapid detection and national standard detection. A portable X-ray fluorescence spectrometer (Thermo Scientific Niton_XL3t_GOLDD analyzer, Thermo Fisher, Wortham, MA, USA) was used for the rapid detection of all samples, and 30% of the samples were tested using the national standard method.

The national standard method test is divided into the following two stages: pretreatment and determination. Firstly, the acid solution method was used for pretreatment. We weighed 0.5000 g of the soil sample into a Teflon crucible, moistened it with 2–3 drops of high purity water and filled it with $HF \rightarrow HNO_3 \rightarrow HClO_4$ in strict order; when the sample was slightly smouldering, it was removed and cooled down; then, $HF \rightarrow HClO_4$ was added and digested to a paste; the paste residue was dissolved in HNO_3 at low temperature and washed into a colourimetric tube [15]. After pretreatment, the As content was determined by microwave digestion/atomic fluorescence (Aurora Lumina Model 3400 Atomic Fluorescence Spectrometer, Aurora Instruments, Vancouver, BC, Canada) and the heavy metals Zn, Cu, Ni and Cr were determined by flame atomic absorption spectrophotometry (Perkin Elmer AA900 Atomic Absorption Spectrometer, Waltham, MA, USA). The results of the rapid tests compared with those of the national standard method showed an accuracy of more than 95%. The results showed that the X-ray fluorescence spectrometer was more accurate and convenient, so all data measured by this instrument were used for statistical analysis to reduce errors.

2.3. Quality Assurance and Control

The testing standard refers to HJ/T166-2004 "Technical Specification for Soil Environmental Monitoring and Modern Analytical Methods for Soil Elements in Chinese National Environmental Monitoring Stations". Compared with the national standard method, the accuracy of the rapid detection method used is above 95%.

2.4. Grey Relational Analysis Method

2.4.1. Principle

The grey correlation analysis method measures the degree of correlation between curves, and its basic idea is to determine the degree of the relationship between the curves based on the similarity of the geometric curves in the sequence [16]. This method compares the geometric relationship of the statistical data from the time series of a system through the quantitative analysis of the development and change trends, and obtains the grey correlation degree between the reference series and each comparison series. If the sample data indicate that there is a greater difference in the changes between two factors, the degree of correlation between them is smaller; otherwise, the degree of correlation is greater [17–19]. When grading and evaluating soil heavy metal pollution, the values obtained of the evaluated factors of the objects in question are used as a reference series, and the grading standard of the soil heavy metal environmental quality is used as a comparison series. Multiple correlation degrees are calculated, and the grade corresponding to the comparison sequence with the highest correlation degree with the reference sequence is the quality grade of the evaluated soil sample [20–22]. Compared with the traditional multi-factor analysis method, the grey relational analysis method has lower data requirements and fewer calculations, which is convenient for dissemination and application [23].

2.4.2. Calculation Procedure

(1) Determine the soil classification standard sequence and the measured parameter sequence of the soil samples at each monitoring point. Then, divide the soil grading standard into h grades, and there are p evaluation factors, then compare sequence $u_i(k)$ of the soil quality standards at all levels, and the parameter sequence $v_j(k)$ composed of the measured values of all soil samples is obtained.

$$u_i(k) = \{u_i(1), u_i(2), \ldots, u_i(p)\}\ (i = 1, 2, \ldots, h);$$

$$v_j(k) = \{v_j(1), v_j(2), \ldots, v_j(p)\}\ (j = 1, 2, \ldots, t);$$

In the formula, $u_i(k)$ is the evaluation factor value of the k-th item in the i-level of the soil heavy metal environmental quality standard; $v_j(k)$ is the measured value of the k-th item evaluation factor in the j-th soil sample.

(2) Normalized processing

Due to the different dimensions of the original data of each evaluation index, the order of magnitude difference is also very different. To eliminate the dimension of the original data and merge the order of magnitude to make it comparable, the original data should be preprocessed first. The following two formulas can be used for normalization:

$$u_i'(k) = \frac{u_i(k)}{\frac{1}{h}\sum_{i=1}^{p} u_i(k)}(i = 1, 2, \cdots, h; k = 1, 2, \ldots, p)$$

In the formula, $u_i'(k)$ is the dimensionless value of the k-th evaluation factor of the i-level soil heavy metal environmental quality standard. u_i is the standard value of each evaluation factor in the soil heavy metal environmental quality classification standard, where subscript i is the soil heavy metal environmental quality level.

$$v_j'(k) = \frac{v_j(k)}{\frac{1}{h}\sum_{i=1}^{p} u_i(k)}(i = 1, 2, \cdots, h; k = 1, 2, \cdots, p; j = 1, 2, \cdots, t)$$

In the formula, $v_j'(k)$ is the dimensionless value of the k-th evaluation factor of the j-th evaluated soil sample; $v_j(k)$ is the measured value of each evaluation factor of the evaluated soil sample.

(3) Calculate the difference sequence $(\Delta_{ji}(k))$ of the two-level minimum difference (Δ_{jmin}) and the two-level maximum difference (Δ_{jmax}), respectively, and the calculation formulas are as follows:

$$\Delta_{ji}(k) = \left|v'_j(k) - u'_i(k)\right|$$

$$\Delta_{jmin} = \begin{matrix} min \\ i \end{matrix} \begin{matrix} min \\ k \end{matrix} \left|v'_j(k) - u'_i(k)\right|$$

$$\Delta_{jmax} = \begin{matrix} max \\ i \end{matrix} \begin{matrix} max \\ k \end{matrix} \left|v'_j(k) - u'_i(k)\right|$$

(4) Calculate the correlation coefficient $\xi_{ji}(k)$.

$$\xi_{ji}(k) = \frac{\Delta_{jmin} + 0.5\Delta_{jmax}}{\Delta_{ji}(k) + 0.5\Delta_{jmax}}$$

(5) Calculate the degree of correlation γ_{ji}.

$$\gamma_{ji} = \frac{1}{p}\sum_{k=1}^{p} \xi_{ji}(k)$$

Find the largest degree among the h correlation degrees, and that is the quality level of the soil sample.

2.5. Factor Analysis Model

2.5.1. Principle

Factor analysis is a multivariate statistical method [24]. It uses a linear function of a small number of common factors and a combination of specific factors to express each variable that was originally observed. Starting with the study of the internal dependence of the correlation matrix, it is a multivariate statistical analysis method that summarizes some complex variables into a few comprehensive factors [25]. When the cumulative variance and contribution rate of these common factors reaches 85% or more than 95%, it means that these common factors reflect most of the information of the research problem, and the factors are not related to each other and do not overlap [26]. The advantage of the factor analysis is mainly that it seeks the basic structure and reduces the variable dimension. When classifying evaluation indicators or samples, the original numerous variables can be finally reflected by a few independent public factors. It is good at expressing many variables with complex relationships with several common factors that have fewer correlations, classifying variables with greater correlations into one category to represent certain influencing factors, and then naming them reasonably [27–29]. In order to identify the main pollution sources, a factor analysis was performed on the normalized data of the six elements. The results of the factor analysis and the possible source types are listed separately for the different pollution types. The larger the eigenvalue, the more important the corresponding factor.

2.5.2. Calculation Procedure

The factor analysis begins with the correlation matrix of the variables and then decomposes an m-dimensional random vector X into less than m, with the representative common factors and a special m-dimensional vector. The number of common factors can obtain the best number, so that they transform the study of m-dimensional random vectors into the study of less common factors. Assuming n samples, n indicators constitute the sample space X.

$$X = (X_{ij})n \times m \quad (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m)$$

The process of the factor analysis generally goes through the following steps:

(1) Standardization of the raw data: the normalized formula is $X'_{ij} = \left(X_{ij} - X_j\right)/\delta_j$, where X_j is the j-th index value of the i-th sample, and X_j and δ_j are the values of the j index's mean and standard deviation, respectiely. The purpose of standardization is to eliminate the influence of the dimensions of the different variables, and the standardized transformation will not change the correlation coefficient of the variables.
(2) Calculate the correlation coefficient matrix of the standardized data, and obtain the eigenvalues and eigenvectors of the correlation coefficient matrix.
(3) Carry out the orthogonal transformation and use the variance maximum method. Its purpose is to make the factor loadings polarized, yet the rotated factors remain orthogonal.
(4) Determine the number of factors, calculate the factor scores, and conduct the statistical analysis.

When the sum of the squares of the simple correlation coefficients among all variables is much greater than the sum of the squares of the partial correlation coefficients, the value of KMO (Kaiser–Meyer–Olkin) is close to 1, and the closer the value is to 1, the stronger the correlation between the variables, and the more suitable the original variable is as a factor analysis; when the sum of squares of the simple correlation coefficients among all variables is close to 0, the KMO value is close to 0. The closer the KMO value is to 0, the weaker the correlation between variables, and the original variables are less suitable for the factor analysis [30–32]. Kaiser gave the commonly used KMO metrics, which are as follows: above 0.9 means very suitable; 0.8 means suitable; 0.7 means general; 0.6 means not suitable; below 0.5 means extremely unsuitable [33].

3. Results and Analysis

3.1. Example Analysis of Grey Correlation Method

3.1.1. Determine the Evaluation Factors

When evaluating soil environmental quality for heavy metal pollution, the parameters that are most harmful to the environment, biology, economy, and society are generally selected as the evaluation factors [34]. On the basis of investigation and research, combined with the sample collection results of its Cd and Hg measured value is low or not measured will As, Pb, Zn, Cu, Ni, Cr and other six indicators as evaluation factors. The results of the 122 soil samples tested were calculated and evaluated by averaging the results according to nine zones in the upper, middle and lower reaches of the surface, middle and lower layers. The average values of soil samples in the nine districts were numbered S_1–S_9, in which S_1, S_2, and S_3 represent the upstream surface, middle, and lower layers. S_4, S_5, and S_6 represent the midstream surface, middle, and lower layers, and S_7, S_8, and S_9 represent the downstream surface, middle, and lower layers. Based on the National Environmental Quality Analysis Standard for Heavy Metals in Soil (GB15618-2018) and the Soil Environmental Quality Construction Land Soil Pollution Risk Control Standard (Trial) (GB36600-2018), combined with the local soil environmental background value in Xinjiang, the soil environmental quality is divided into four grades, as shown in Table 1. Figure 2 shows the measurement results of the soil heavy metal content from the sampling points along the Shuimo River.

Table 1. National evaluation standards of heavy metal concentrations in the soil environment (unit: mg.kg^{-1}).

Grading	Pb	Zn	As	Cu	Ni	Cr
I	35	100	15	35	40	90
II	250	200	25	50	60	150
III	350	250	40	75	80	200
IV	500	300	50	100	200	250

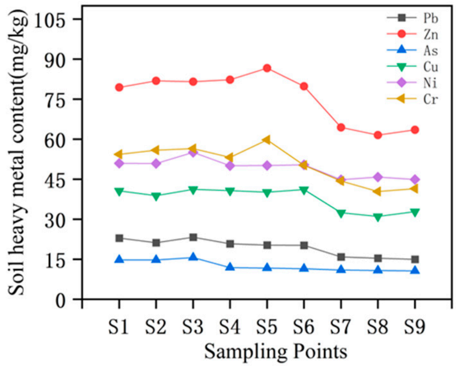

Figure 2. Identification of the content of heavy metal in soil (unit: mg.kg^{-1}).

It can be observed from Figure 1 that the soil heavy metal content at each sampling point showed an increasing trend first, followed by a decreasing trend. The maximum values of Zn and Cr appear at sampling point S5, while the maximum values of Ni, Cu, Pb, and As all appear at S3. The minimum values of Ni appear at sampling point S7, and the minimum values of Zn, Cr, and Cu all appear at sampling point S8, whereas the minimum values of Pb and As occur at S9. It can be inferred that the heavy metal content in the downstream soil is relatively low, and the soil quality is relatively good.

3.1.2. Evaluation Process and Results

Taking S1 as an example, the heavy metal pollution in the soil along the Shuimo River is evaluated according to the steps of the grey correlation analysis method. The specific evaluation process is as follows:

(1) List the matrix composed of the measured values of the heavy metal content in the soil samples and the graded standard values to obtain the following:

$$\begin{bmatrix} v_1(k) \\ u_i(k) \end{bmatrix} = \begin{bmatrix} 22.92 & 79.42 & 14.74 & 40.63 & 50.95 & 54.36 \\ 35 & 100 & 15 & 35 & 40 & 90 \\ 250 & 200 & 25 & 50 & 60 & 150 \\ 350 & 250 & 40 & 75 & 80 & 200 \\ 500 & 300 & 50 & 100 & 200 & 250 \end{bmatrix}$$

(2) Following the normalization processing, we have the following:

$$\begin{bmatrix} v_1'(k) \\ u_i'(k) \end{bmatrix} = \begin{bmatrix} 1.18 & 1.05 & 1.18 & 1.08 & 1.03 & 1.07 \\ 0.12 & 0.47 & 0.46 & 0.54 & 0.42 & 0.52 \\ 0.88 & 0.94 & 0.77 & 0.77 & 0.63 & 0.87 \\ 1.23 & 1.18 & 1.23 & 1.15 & 0.84 & 1.16 \\ 1.76 & 1.41 & 1.54 & 1.54 & 2.11 & 1.45 \end{bmatrix}$$

(3) Calculate the difference sequence $(\Delta_{ji}(k))$, the two-level minimum difference (Δ_{jmin}), and the two-level maximum difference (Δ_{jmax}), using the above formula.

$$\Delta_{1i}(k) = \begin{bmatrix} 1.18 & 1.05 & 1.18 & 1.08 & 1.03 & 1.07 \\ 0.12 & 0.47 & 0.46 & 0.54 & 0.42 & 0.52 \\ 0.88 & 0.94 & 0.77 & 0.77 & 0.63 & 0.87 \\ 1.23 & 1.18 & 1.23 & 1.15 & 0.84 & 1.16 \\ 1.76 & 1.41 & 1.54 & 1.54 & 2.11 & 1.45 \end{bmatrix}$$

$$\Delta_{1max} = 1.071$$

$$\Delta_{1min} = 0.055$$

(4) Calculate the correlation coefficient and correlation degree.

$$\xi_{ji}(k) = \frac{0.055 + 0.5 \times 1.071}{\Delta_{1i}(k) + 0.5 \times 1.071}$$

$$\xi_{1i}(k) = \begin{bmatrix} 0.37 & 0.53 & 0.470.55 & 0.51 & 0.54 \\ 0.71 & 0.92 & 0.630.70 & 0.63 & 0.80 \\ 1.00 & 0.89 & 1.000.97 & 0.81 & 0.95 \\ 0.53 & 0.66 & 0.660.59 & 0.37 & 0.65 \end{bmatrix}$$

$$\gamma_{1i} = \{\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{14}\} = \{0.50, 0.73, 0.94, 0.58\}$$

Among them, the maximum correlation degree is γ_13 = 0.936 and the quality level of S1 is Level III;.

Repeat the above steps in turn to find the correlation degree between all soil sample data and the soil heavy metal environmental quality classification standard series. The detailed results are shown in Figure 3. In Figure 3a–i show the correlation between the soil sample data from sampling sites S1–S9 and the series of environmental quality criteria for soil heavy metals, respectively.

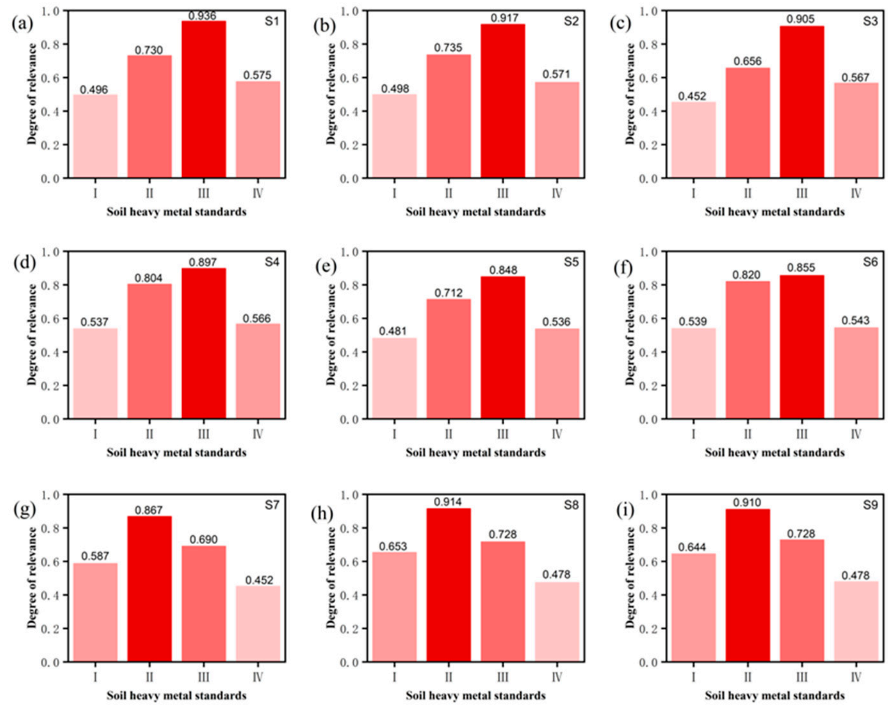

Figure 3. Correlation degree between the soil sample data and the soil heavy metal environmental quality classification standard along the Shuimo River. (**a**–**i**) are the correlations of the sampling points S1–S9.

Figure 3 shows the correlation degree between the data of S1–S9 sampling points and the four grading standard series of soil heavy metals. The higher the column and the darker the color, the greater the correlation degree. It can be clearly observed from the figure that S1–S6 is evaluated as Grade 3 according to the soil heavy metal standard, and S7–S9 is evaluated as Grade 2 according to the soil heavy metal standard. The detailed maximum correlation degree and evaluation results are shown in Table 2.

Table 2. Evaluation results of the soil environmental quality along the Shuimo River.

Soil Number	Maximum Relevance	Level
S1	0.936	III
S2	0.917	III
S3	0.905	III
S4	0.897	III
S5	0.848	III
S6	0.855	III
S7	0.867	II
S8	0.914	II
S9	0.910	II

3.2. Example of Factor Analysis

The statistic of the Bartlett sphericity test is obtained from the determinant of the correlation coefficient matrix. If the value is large and the corresponding associated probability value is less than the significance level in the user's mind, the null hypothesis should be rejected. It is considered that the correlation coefficient matrix cannot be a unit matrix, that is, there is a correlation between the original variables, which is suitable for the principal component analysis; on the contrary, it is considered unsuitable for the factor analysis [35,36]. If the significance is <0.05 for the Bartlett sphericity test, the data are suitable for the factor analysis method.

The value of KMO is 0.754, obtained by SPSS 2021 processing of the soil heavy metal data of the upper, middle, and lower reaches of the soil along the Shuimo River, and the sig (significance) of the Bartlett sphericity test is $0 < 0.05$, which is suitable for the factor analysis.

In Figure 4, the darker the color and the larger the radius of the circle, the stronger the correlation. It can be observed from the figure that the correlation coefficients of Zn and Cr, Pb and Ni, Cu and Zn, and As and Pb are the largest.

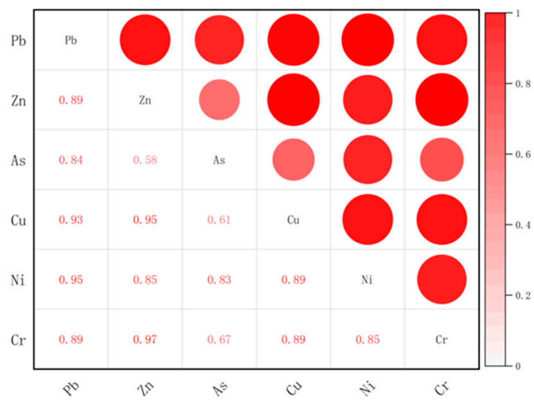

Figure 4. Variable correlation matrix.

On the premise that the cumulative variance is 96.325% (>90%), the analysis results in two main factors. It can be observed that the two main factors provide 96.325% of the information of the source data, satisfying the principle of the factor analysis. From Figure 5, it can be observed that the total cumulative contribution rate has not changed before and after rotation, that is, the total amount of information has not been lost. It can be observed from Figure 5 that after rotation, the variance contribution rate of Main Factor 1 is about 59.493%, and the variance contribution rate of Main Factor 2 is 36.831%. This can be explained by Factor 1 and Factor 2 as possibly the most important pollution sources of heavy metal pollution in the soil along the Shuimo River, and these play an important role in the heavy metal pollution along the Shuimo River. The main purpose of the factor analysis is to place the variables with similar factor loads under a common factor, and the maximum rotation of the orthogonal variance means that each main factor only has a correlation with the least number of variables. A sufficient amount of small factor loadings is required to allow for a more reasonable interpretation of the factor's significance [37]. The output results are shown in Figures 6 and 7.

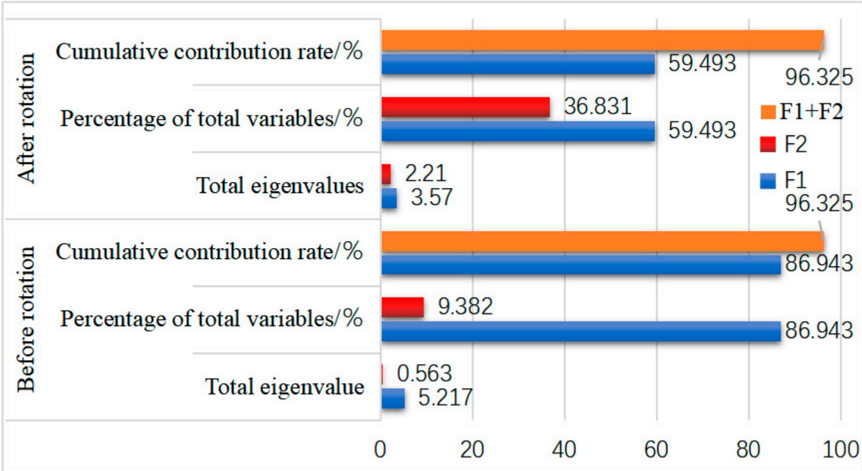

Figure 5. Eigenvalues and cumulative contribution rates.

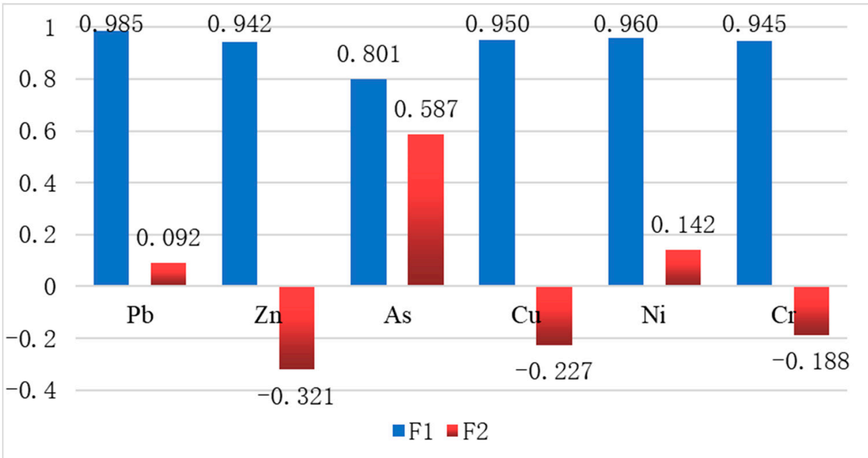

Figure 6. Pre-rotation factor load matrix.

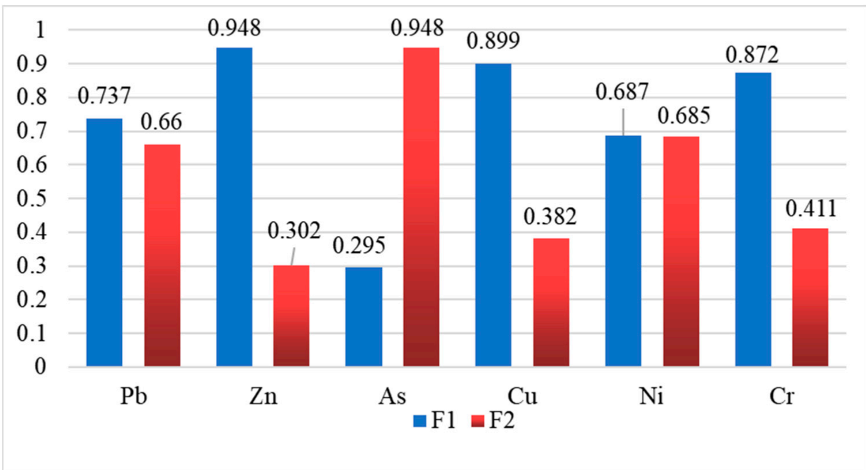

Figure 7. Variance in the maximum orthogonal post-rotation factor load matrix.

It can be observed from the output results that the variable results of the factor loading before and after rotation are basically consistent. The greater the absolute value (loading) of the link coefficient between a variable and a certain factor, the closer the relationship between the factor and the variable. The orthogonal sub-solution shows that Zn, Cu, and Cr pollution come from Factor 1, As pollution comes from Factor 2, and the contribution rate of Factor 1 and Factor 2 to Pb and Ni is roughly equal.

4. Conclusions

According to the evaluation results of the grey relational model, among the nine soil samples along the Shuimo River, six belonged to Grade 3 with slight pollution, and three belonged to Grade 2 without pollution. The soil evaluation grades of the upstream and middle reaches along the river are similar to slight pollution, and the soil environment of the downstream area is good and unpolluted. According to the results of the factor analysis model, there are two factors that cause slight pollution in the upstream and middle reaches of the river. Factor 1 pollution is mainly Zn, Cu, Cr, Pb, and Ni, and Factor 2 pollution is mainly As, Pb, and Ni. The upstream and midstream are the main industrial areas, analyzed by the elements in Factor 1. Factor 1 is industrial pollution. Factor 2 is mainly As pollution, and As mainly comes from various insecticides, rodenticides, etc., so it can be inferred that Factor 2 is mainly a source of human household pollution. Farmland and vegetable fields are distributed along both sides of the river in the downstream area, so the soil quality is relatively good. Using the combination of the grey relational analysis and factor analysis models to evaluate the degree of soil heavy metal pollution and the pollution sources, we can reduce the influence of the subjective factors, and the calculation is simple. The evaluation results obtained are basically consistent with the actual heavy metal pollution situation and pollution sources, which can more objectively, conveniently, and truly reflect the environmental quality of the soil along the river.

Author Contributions: Y.Z.: data curation, investigation, funding acquisition, supervision, writing—review & editing, resources, investigation. Y.W.: conceptualization, formal analysis, writing—original draft, software, methodology, writing review & editing, methodology. H.Z.: conceptualization, visualization, investigation, data curation. J.Y.: data curation, project administration, validation, resources. H.M.: writing—review & editing, resources, supervision. All authors have read and agreed to the published version of the manuscript.

Funding: Xinjiang University: 620312339.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the data involve sensitive information such as the environmental status of Urumqi.

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

References

1. Zhang, W.I.; Kolbe, H.; Zhang, S.L.; Zhang, D.X.; Bai, Z.G.; Zhang, J.; Shi, H.D. Overview of Soil Survey Works in Main Countries of World. *Sci. Agric. Sin.* **2022**, *55*, 3565–3583. [CrossRef]
2. Yang, Z.; Liao, C.-L.; Zou, Y.; Xie, W.-J.; Chen, X.-W.; Zhang, Y.-F. Distribution Characteristics, Source Analysis, and Potential Ecological Risk Assessment of Soil Heavy Metals in Liuyang River Source Area. *Environ. Sci.* **2022**, *47*, 1–15. [CrossRef]
3. Hou, F.; Lv, G.; Ma, H.; Chen, Y. Contamination characterization and source apportionment of trace metals in riparian soils of Shuimo River in Urumqi. *Environ. Chem.* **2017**, *36*, 2451–2461. [CrossRef]
4. Ma, H.; Yang, X.; Gong, X.; Chen, Y.; Lv, G. Status quo and evaluation of heavy metal pollution in the riparian soils of the Shuimo River in Urumqi. *J. Soil Water Conserv.* **2016**, *30*, 300–307. [CrossRef]
5. Liu, Y.; Niu, H. Grey relational analysis of heavy metal pollution in soil of yindongliang mine area. *J. Liaoning Tech. Univ.* **2013**, *32*, 1626–1629. [CrossRef]
6. Xiao, H.; Huang, B.; Sun, K.; Xu, M.; Wu, J. Grey relational analysis for evaluating the effects of lime combined with organic fertilizer on a cadmium-contaminated soil-plant system. *J. Agro-Environ. Sci.* **2022**, *41*, 1966–1974. [CrossRef]
7. Wei, J.; Zheng, X.; Liu, J. Modeling Analysis of Heavy Metal Evaluation in Complex Geological Soil Based on Nemerow Index Method. *Metals* **2023**, *13*, 439. [CrossRef]

8. Yan, Q.; Li, C.; Qin, Q.; Deng, X.; Song, X.; Tang, J. Evaluation on Soil Fertility of *Pinus massoniana* Resin Stands Based on Grey Relation Analysis. *J. Fujian For. Sci. Technol.* **2021**, *48*, 62–66. [CrossRef]
9. Xing, H.; He, Q.; Pu, Z.; Wang, G.; Jin, C. Analysis of precipitation characteristics during the vegetation growth season in the Urumqi River Basin in arid region of northwest China. *J. Arid. Meteorol.* **2023**, *41*, 34–42. [CrossRef]
10. Tian, T. Research on the Evaluation of Water Environmental Quality and the Control Measures of Pollution for Shuimo River in Urumqi. Ph.D. Thesis, Southwest University of Science and Technology, Mianyang, China, 2018.
11. Han, Q.; Song, F.-Y.; Chen, J.-f.; Qi, Q.-q.; Wang, T.; Li, F. Analysis on the Trend of Sewage Interception and Pollution Control and Water Quality Status of Shuimo River (Gumudi River and Laolong River) in Urumqi from 2018 to 2021. *Environ. Sci. Surv.* **2022**, *41*, 57–64. [CrossRef]
12. Chen, Y.-J.; Qu, X.-C.; Zhang, B.; Li, X.-T.; Zhu, P.-G.; Zha, T.-G. Ecological and Health Risk Assessment of Heavy Metal Pollution in Farmland Soil of Xianghe County. *Environ. Sci.* **2022**, *43*, 5728–5741. [CrossRef]
13. Tian, T. Analysis of Water Quality Trend of Shuimo River Based on Pearman Rank Coefficient Method. *Energy Conserv. Environ. Prot.* **2019**, *9*, 52–53.
14. Yang, Z.; Liu, X. Research on ecosystem service value and spatial evolution in Urumqi City. *Environ. Sci. Technol.* **2021**, *44*, 226–236. [CrossRef]
15. Liu, F.; Liu, X. *Soil and Solid Waste Monitoring and Analysis Technology*; Chemical Industry Press: Beijing, China, 2007.
16. Shen, L.; Hu, W.; Zhang, L.; Feng, J.; Sun, B.; Mou, Y. Characteristics and Evaluation of Heavy Metal Contamination in Soil of Typical Waste Dumping Area. *Nonferrous Met. Eng.* **2022**, *12*, 187–197. [CrossRef]
17. Liu, S.; Wang, W.; Han, Y.; Xue, H. Assessment of Heavy Metal Pollution Situation of Farmland Soil Located in Zhangmai Village, Nyingchi City. *J. Plateau Agric.* **2022**, *6*, 349–355. [CrossRef]
18. Hu, Y.-C. Nonadditive tourism forecast combination using grey relational analysis. *Grey Syst. Theory Appl.* **2023**, *13*, 277–296. [CrossRef]
19. Kinuthia, G.K.; Ngure, V.; Beti, D.; Lugalia, R.; Wangila, A.; Kamau, L. Levels of heavy metals in wastewater and soil samples from open drainage channels in Nairobi, Kenya: Community health implication. *Sci. Rep.* **2020**, *10*, 8434. [CrossRef] [PubMed]
20. Ren, W.; Hu, D.; Cui, Y.; Feng, J.; Sun, B.; Mou, Y. Ecological Risk Evaluation on Soil Heavy Metal Around Typical Coal Mines in Northwest Guizhou. *Gui Zhou Agric. Sci.* **2022**, *50*, 103–111. [CrossRef]
21. Kanianska, R.; Varga, J.; Benková, N.; Kizekova, M.; Jancova, L. Floodplain soils contamination assessment using the sequential extraction method of heavy metals from past mining activities. *Sci. Rep.* **2022**, *12*, 2927. [CrossRef]
22. Li, Y.; Yang, S.; Hu, R.; Xie, Y.; Li, Y.; Huang, X. Distribution characteristics and risk assessment of soil heavy metals in different geomorphological units of Yinchuan Plain. *Environ. Chem.* **2022**, *41*, 4130–4144. [CrossRef]
23. Yan, K.; Wang, H.; Lan, Z.; Gong, J.; Gao, J.; Tang, S.; Ma, S.; Duan, Z. Heavy metal pollution in the soil of contaminated sites in China: Research status and pollution assessment over the past two decades. *J. Clean. Prod.* **2022**, *373*, 135340. [CrossRef]
24. Zhang, W.; Liu, M.; Li, C. Soil heavy metal contamination assessment in the Hun-Taizi River watershed, China. *J. Sci. Rep.* **2020**, *10*, 8730. [CrossRef] [PubMed]
25. Shi, J.; Zhao, D.; Ren, F.; Huang, L. Spatiotemporal variation of soil heavy metals in China: The pollution status and risk assessment. *Sci. Total Environ.* **2023**, *871*, 161768. [CrossRef] [PubMed]
26. Dong, L.; Hu, W.; Huang, B. Source appointment of heavy metals in suburban farmland soils based on positive matrix factorization. *China Environ. Sci.* **2015**, *35*, 2103–2111.
27. Mazidah, Z.; Amalyn, A.; Shuhada, N.T.; Fuad, M.M.; Fikriah, F.; Izzuddin, M.A. Assessment of heavy metals and nutrients availability in oil palm plantation effected by bauxite mining using geostatistical and multivariate analyses. *IOP Conf. Ser. Earth Environ. Sci.* **2022**, *1064*, 012002. [CrossRef]
28. Chen, Z.; Hua, Y.; Xu, W.; Pei, J. Analysis of heavy metal pollution sources in suburthan famland based on positive definite matrix factor model. *Acta Sci. Circumst.* **2020**, *40*, 276–283. [CrossRef]
29. Gao, Y.; Liang, Z.; Qiao, P. Research on the Evaluation of Business Environment in Guangxi's Based on Factor Analysis-Entropy Method. *China J. Commer.* **2023**, *876*, 17–20. [CrossRef]
30. Lan, Z.Y.; Li, Y.; Huang, Y.T.; Shi, W.F.; She, D.Y.; Jiang, Z.; Liu, L. Construction of a risk assessment indicator system for re-establishment of imported malaria. *Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi—Chin. J. Schistosomiasis Control* **2022**, *34*, 163–171. [CrossRef]
31. Zhang, X.-H.; Li, C.-Y. Research on the Influencing Factors of Problem-Driven Children's Deep Learning. *Front. Psychol.* **2022**, *13*, 764121. [CrossRef]
32. Tian, L.; Wang, X. A Dynamic Prediction Neural Network Model of Cross-Border e-Commerce Sales for Virtual Community Knowledge Sharing. *Comput. Intell. Neurosci.* **2022**, *2022*, 2529372. [CrossRef]
33. Xie, J.; Ding, L.; Wang, M. Spectral clustering based unsupervised feature selection algorithms. *Ruan Jian Xue Bao/J. Softw.* **2020**, *31*, 1009–1024. [CrossRef]
34. Iordache, A.M.; Nechita, C.; Zgavarogea, R.; Voica, C.; Varlam, M.; Ionete, R.E. Accumulation and ecotoxicological risk assessment of heavy metals in surface sediments of the Olt River, Romania. *Sci. Rep.* **2022**, *12*, 880. [CrossRef] [PubMed]
35. Liu, Y.; Gu, W.; Wu, Y. Pollution Characteristics, Health Risks and Source Analysis of Surface Soil Heavy Metals in Urban Parks of Datong City. *Chin. J. Soil Sci.* **2023**, *54*, 180–191. [CrossRef]

36. Luo, T.; Zhao, F.; Liu, X.; Zhu, H.-H.; Zhu, Q.-H.; Xu, C.; Zhang, Q.; Huang, D.-Y. Sources identification of heavy metals contamination in farmland soils of a typical small watershed around industrial area. *Res. Agric. Mod.* **2022**, *43*, 541–550. [CrossRef]
37. Ahirvar, B.P.; Das, P.; Srivastava, V.; Kumar, M. Perspectives of heavy metal pollution indices for soil, sediment, and water pollution evaluation: An insight. *Total Environ. Res. Themes* **2023**, *6*, 100039. [CrossRef]