Assessment of Soil Pollution Levels in North Nile Delta, by
Integrating Contamination Indices, GIS, and Multivariate Modeling

Mohamed E. Abowaly 1, Abdel-Aziz A. Belal 2, Enas E. Abd Elkhalek 1, Salah Elsayed 3, Rasha M. Abou Samra 4, Abdullah S. Alshammari 5, Farahat S. Moghanm 1, Kamal H. Shaltout 6, Saad A. M. Alamri 7 and Ebrahim M. Eid 7,8,*

Abstract: The proper assessment of trace element concentrations in the north Nile Delta of Egypt is needed in order to reduce the high levels of toxic elements in contaminated soils. The objectives of this study were to assess the risks of contamination for four trace elements (nickel (Ni), cobalt (Co), chromium (Cr), and boron (B)) in three different layers of the soil using the geoaccumulation index (I-geo) and pollution load index (PLI) supported by GIS, as well as to evaluate the performance of partial least-square regression (PLSR) and multiple linear regression (MLR) in estimating the PLI based on data for the four studied trace elements in the three different soil layers. The results show a widespread contamination of I-geo Ni, Co, Cr, and B in the three different layers of the soil. The I-geo values varied from 0 to 4.74 for Ni, 0 to 6.56 for Co, 0 to 4.11 for Cr, and 0 to 4.57 for B. According to I-geo classification, the status of Ni, Cr, and B ranged from uncontaminated/moderately contaminated to strongly/extremely contaminated. Co ranged from uncontaminated/moderately contaminated to extremely contaminated. There were no significant differences in the values of I-geo for Ni, Co, Cr, and B in the three different layers of the soil. According to the PLI classification, the majority of the samples were highly polluted. For example, 4.76% and 95.24% of the samples were unpolluted and very highly polluted, respectively, in the subsurface layer of the soil profiles. Both calibration (Cal.) and validation (Val.) models of the PLSR and MLR models showed the highest performance in predicting the PLI based on data for the four studied trace elements, as an alternative method. The validation (Val.) models performed the best in predicting the PLI, with $R^2 = 0.89–0.93$ in the surface layer, 0.91–0.96 in the subsurface layer, 0.89–0.94 in the lowest layers, and 0.92–0.94 across the three different layers. In conclusion, the integration of the I-geo, PLI, GIS technique, and multivariate models is a valuable and applicable approach for the assessment of the risk of contamination for trace elements, and the PLSR and MLR models could be used through applying chemometric techniques to evaluate the PLI in different layers of the soil.

Keywords: trace elements; GIS; geoaccumulation index; PLI; PLSR; north Nile Delta of Egypt
1. Introduction

Land degradation is one of the most serious environmental problems facing the world today. Due to various degradation processes, about 6 million hectares of agricultural land around the world have become less productive [1]. For example, due to the increasing number of factories and their emissions, urban growth, increasing traffic volumes, and the use of wastewater and waste deposits, the quality of the Nile Delta’s land is worsening [2]. The Nile Delta (an area of approximately 20,000 km$^2$) represents only 2.3% of Egypt’s area but 46% of the total cultivated area (55,040 km$^2$) [3,4]. On the Nile Delta, the proportion of agricultural land is 63%, due to the suitable soil characteristics and the existence of irrigation systems [5]. The Nile Delta is dependent on wastewater for irrigation operations. Mixing agricultural and industrial drainage wastewater from the El Gharbia main drain (Kitchenr), Egypt with Nile water was used for irrigation in area of study. The Nile Delta has been recognized for its agricultural activities for thousands of years, and it is one of the most densely populated areas in the world [6]. The types of land degradation in the Nile Delta are water erosion, water logging, salinity, sodicity, and compaction [7].

Uncontrolled activities are the main contributor to the deposition of many toxic substances into the soil, which ultimately leads to soil degradation and harm to human health. The main human factors resulting in the deposition of soil-bearing solids are agricultural practices, especially land use, inorganic and petrochemical fertilizers for organic matter (bio-solids, animal manure and organic fertilizers), and pesticides [8], and soil quality is greatly affected by human activities [9]. The geochemical properties of agricultural soils, especially the levels of trace elements, influence the quality of the soil [10–13]. Contamination with trace elements may pose risks to humans and ecosystems through contact with contaminated soil or direct consumption, the food chain (plant soil–human), groundwater contamination, impaired food quality (safety and marketability) due to phytotoxicity, and reduced land use for agricultural production, causing food insecurity and land tenure problems [14]. The regular application of large quantities of fertilizers in intensive farming systems to provide enough nitrogen (N), phosphorus (P) and potassium (K) is important for crop growth. The compounds used to supply these elements may contain trace amounts of trace elements, which, during the continuous application of fertilizers, may significantly increase the soil mineral content [15].

The mineral pollution of the environment has a significant impact on human health [16]. Furthermore, certain trace elements such as nickel (Ni), cadmium (Cd), cobalt (Co), lead (Pb), arsenic (As), chromium (Cr), and boron (B) directly cause significant health problems in humans, animals, and plants. The study area is considered highly contaminated with heavy metals, especially the elements Cr, Co, B and Ni. The sources of these elements are the industrial drainage in the El Gharbia main drain. For example, Ni help plants to grow, but too much Ni can be harmful to humans and other animals. Excessive Ni in the soil can increase the risk of developing lung, nose, and skin cancers [17,18]. In Egypt, the measured Cr concentrations in the soil range from 11.6 to 179 ppm, depending on the soil type and land management [19]. Cr$^{3+}$’s toxicity depends on its oxidation status; Cr$^{3+}$ is relatively harmless, while Cr$^{6+}$ is highly toxic. If there is too much Cr in the soil, it can cause cancer and damage to the liver and kidneys [20]. Excessive Co concentrations in the soil, transferred to plants, can slow the growth of roots and shoots, which can disrupt the uptake of water and nutrients by the plants, reducing both the quantity and efficiency of yields [21]. According to an FAO study on micronutrients in soils, B deficiency is the most common issue, affecting at least 8 million hectares globally [22]. B deficiency is most common in humid areas with well-drained soils or sand soils, as recorded in China, Japan, and the United States [23]. Both B deficiency and toxicity are related to plant disorders and crop yield losses.

There are currently a plethora of contamination indices and measurement methods available for assessing soil contamination, such as the geoaccumulation index (I-geo) and pollution load index (PLI). The majority of investigators have given their attention to the transport of the trace elements and their distribution in the soil [24,25]. I-geo is useful for
assessing and analyzing trace element contents in soil because it intuitively reflects the effects of human activities on trace elements as well as the impact of trace elements on the environment [26]. The PLI represents the number of times the trace element concentrations in the soil surpass the background concentrations, and it provides a summative measure of the total level of trace element toxicity in a given sample. Both parameters were used to evaluate the contamination levels in the soil and bottom sediment of Upper Egypt [27].

Natural landscapes and ecosystems have been impacted by global development and unplanned agricultural practices [28]. Therefore, the evaluation of soil pollution needs to consider in depth knowledge of the spatial distribution of contaminants [29]. The use of a GIS database can provide detailed information for low-cost soil surveyance. GIS databases can also help in deriving digital elevation models (DEMs), which can assist in the development of landscape features used to characterize landform [30]. It is important to analyze the distribution and concentration of trace elements. This will enable pollution levels to be determined and the associated impacts on both the environment and human health to be assessed. The evaluation and mapping of soil toxic elements can help in developing strategies to promote the sustainable use of soil resources, reduce soil degradation, and expand crop production. Geophysical completion is used to survey and interpret the spatial distributions of pollutants in soil [31]. The function of the weighted distance (IDW) is useful for comprehensively evaluating pollution patterns [32]. Proper assessments of the toxic element concentrations in soils, supported by GIS databases, are needed to reduce the high levels of toxic elements in contaminated soils [33].

Calculating the PLI requires a series of calculation steps that require significant time and effort to convert several numbers from the trace element data for the soil into a single value describing the soil contamination level [27]. Partial least-square regression (PLSR) and multiple linear regression (MLR) could be used to solve this problem since they are standard methods for specifying a linear relationship between independent variables and dependent variables [34,35]. New modeling frameworks and multivariate regression models such as PLSR and MLR should facilitate a substantial increase in the efficiency of predicting dependent variables based on several independent variables. PLSR is commonly used to create predictive models of the hyperspectral responses of in situ canopy samples [9,36]. PLSR was recently shown to perform well in assessing water quality indices [37]. PLSR and MLR have been proposed for resolving strongly multicollinear and noisy variables and efficiently assessing measured parameters [38]. Both methods can combine data for a large number of trace elements into a single index to enhance the prediction of a measured variable. Therefore, using these methods, the PLI or other pollution indices can be simultaneously estimated from data for several trace elements. PLSR can reduce many collinear factors to a few, non-correlated latent factors, preventing the overfitting or underfitting of the data, and avoiding redundant data [37,39]. There is little information available with which to evaluate the PLSR and MLR methods based on trace elements for assessing the PLI of soil in different layers.

Therefore, the objectives of this study were to (i) map soil pollution according to four trace elements (Ni, Co, Cr, and B) based on I-geo using the GIS technique for soil profiles in three soil layers; (ii) assess the risk of contamination for four trace elements using I-geo and PLI in three different soil layers; and (iii) evaluate the performance of PLSR and MLR models based on four trace elements (Ni, Co, Cr, and B) for predicting PLI, as a new method, in three different soil layers.

2. Materials and Methods

2.1. The Study Area

The research area is situated north of the Nile Delta (Figure 1) in Kafr El-Sheikh Governorate, Egypt. The geographic coordinates are in UTM zone 36 (latitudes 31°0′0″–31°40′0″ N and longitudes 30°30′0″–31°10′0″ E; Figure 1). The northern part of the Nile Delta is in the arid region, while the southern part is in the hyper-arid region, according to a map of the global distribution of arid regions. The study area is encircled by two branches of the
Nile River: the Rosetta in the west and the Damietta to the east. A network of 40,000 km of canals diverts and supplies water from the Nile River to nearly 2 million farmers for cropland irrigation, with a similar network of drainage canals also incorporated in the region [40,41]. These drainage canals cover approximately 18,000 km, leading to a total length of nearly 58,000 km when combined with irrigation canals [40]. In the Nile Delta, the population’s normal growth rate is 21,600/year. An area of 825,100 acres is preserved for cultivation and is famous for agricultural rice, beet, cotton, and wheat. In the north of the Nile Delta, the Lake Burullus flood plain, coastal plain, urban and industrial commercial centers, and some sand dunes in the coastal sections are all important features and major landforms. The alluvial plain, lacustrine plain, and marine plain (71.08%, 19.34%, and 9.57% of the total territory, respectively) are the three major physiographic units.

Figure 1. Study area location in Landsat-8 Image and the sample collection sites.

2.2. Soil Analysis

Soil samples were collected from 21 representative profiles in Kafr El-Sheikh Governorate during September 2018. Twenty one soil profiles were selected according to geomorphologic units in the study area to represent the different agricultural practices and soil samples were collected from different layers according to morphological variations. As each physiographic unit covers more than one profile, and therefore the degree of certainty in the distribution is great because it covers the physiographic units to a large extent. The soil profiles were described according to virtual characterization, and the Mansell book was used to determine the color grade. The studied profiles and soil sample locations were determined using a Global Positioning System (GPS) unit (German model) as shown in Figure 1. Throughout the sampling, a GPS was used to pinpoint the precise locations of the sampling sites. Soil samples were taken from the selected top layers. The study area was classified in to order Entisols, there are no morphological differences in the soil profiles, so soil samples were taken at the three levels. The type of soil is clay. Three soil samples were taken from each profile at depths of 0–30, 30–60, and 60–100 cm. All the samples were placed in sealed polyethylene bags and returned to the laboratory. The samples were composited, homogenized, air-dried at 25 to 35 °C, crushed, and sieved to 2 mm. Soil properties were determined using the prepared samples (pH, EC, Ca$^{2+}$, Mg$^{2+}$, K$^+$, Na$^+$, CO$_3^{2-}$, HCO$_3^-$, Cl$^-$, SO$_4^{2-}$, OM %, CaCO$_3$ %, P% and N%) in Table S1.

For the determination of the total element concentration, exactly 1 g of powdered soil sample was digested with aqua regia (HNO$_3$:HCl, 1:3). The concentrations of Cd, Co, Ni, and B were determined in the Laboratory of Geochemistry and Toxic Elements, Faculty of Agriculture, Kafrelsheikh University (accredited according to ISO/IEC 17025-2017 by...
ICP-OES). The threshold trace element concentrations in the soil, in mg kg\(^{-1}\) dry soil, were determined according to Kabata-Pendias [42] as shown in Table 1.

Table 1. The threshold trace element concentrations (mg kg\(^{-1}\)) in soil as dry soil.

| Trace Elements | Threshold Concentration (mg kg\(^{-1}\)) |
|---------------|------------------------------------------|
| Cr            | 50–200                                   |
| Co            | 20–50                                    |
| B             | 42                                        |
| Ni            | 20–60                                     |

2.3. Assessment of Contamination

2.3.1. Geoaccumulation Index (I-geo)

The I-geo expresses pollution by comparing the measured levels of trace elements with the background levels originally used for evaluating bottom sediments [43,44]. Trace element contamination was assessed using the geographic accumulation index (I-geo) using the following equation:

\[
I_{\text{geo}} = \log_2 \frac{C_n}{1.5B_n}
\]

where \(C_n\) is the concentration of the trace elements measured in the soil, \(B_n\) is the geochemical background concentration of the trace element (medium crust) [45]. The constant 1.5 in Equation (1) was introduced to minimize the effect of potential differences in background values that could be attributed to rocky differences in sediments. The concentration here is between the concentration obtained and that of the elements in the Earth’s crust, because soil is part of the Earth’s crust, and its chemical composition is related to that of the crust [46]. The I-geo classification is shown in Table 2.

Table 2. Class, value, and contamination level according to geoaccumulation index (I-geo) in soil [46].

| I-geo Class | I-geo Value | Contamination Level                 |
|------------|-------------|-------------------------------------|
| 0          | I-geo \(\leq\) 0 | Uncontaminated                      |
| 1          | 0 < I-geo < 1 | Uncontaminated/moderately contaminated |
| 2          | 1 < I-geo < 2 | Moderately contaminated              |
| 3          | 2 < I-geo < 3 | Moderately/strongly contaminated     |
| 4          | 3 < I-geo < 4 | Strongly contaminated                |
| 5          | 4 < I-geo < 5 | Strongly/extremely contaminated      |
| 6          | 5 < I-geo     | Extremely contaminated               |

2.3.2. The Pollution Load Index (PLI)

The PLI is a geometric average of impurity coefficients (CF) that defines the contribution of all trace elements in a specific place [47]. Contamination factor (CF) was computed by dividing metal concentration by background value using the following equation [48]:

\[
CF = \frac{C_n}{B_n}
\]

Pollution load index (PLI) is used to estimate elements concentrations in soils relative to the reference concentration and was calculated using the following equation [49]:

\[
PLI = (CF_1 \times CF_2 \times \ldots \times CF_n)^{1/n}
\]

where CF is the contamination factor and \(n\) is the number of metals.

This parameter can be used to determine the level of environmental pollution in order to undertake monitoring activities to improve soil quality. The PLI classification is shown in Table 3.
Table 3. Class, value, and pollution level according to pollution load index (PLI) in soil [47].

| PLI Class | PLI Value | Pollution Level |
|-----------|-----------|-----------------|
| 1         | 0 < PLI ≤ 1 | Unpolluted      |
| 2         | 1 < PLI ≤ 2 | Moderately polluted to unpolluted |
| 3         | 2 < PLI ≤ 3 | Moderately polluted |
| 4         | 3 < PLI ≤ 4 | Moderately to highly polluted |
| 5         | 4 < PLI ≤ 5 | Highly polluted  |
| 6         | 5 ≤ PLI     | Very highly polluted |

2.4. Spatial Distributions of Trace Elements

Spatial completion is widely used when data are collected in distinct locations (such as soil profiles) to produce continuous information [50]. The ArcGIS Spatial Analyst 10.2.1 extension offers spatial data analysis tools that use statistical theory and techniques to model spatially referenced data. Data for four trace elements were used to derive the intervening values using ArcGIS Spatial Analysis’ interpolation methods. The weighted inverse distance (IDW) is an interpolation method that uses the values measured surrounding the prediction site. The values measured closest to the prediction site have a greater effect on the expected values of the distant ones; greater weight is given to the points closest to the prediction site, the weight being a function of distance [51]. The advantage of using IDW in mapping of spatial distribution of heavy metals is that it is efficient. This interpolation method works better with equally distributed points [52]. In this study, the concentration of elements is not due to natural sources, but the presence of other sources such as agricultural and industrial drainage, which directly affects the concentration of elements in the soil, which varies based on the distance from the source, and therefore it is better to choose this IDW method. The statistical relations between the known points were determined using the IDW function of ArcGIS and were used to determine the concentrations of trace elements in the study area. The IDW was used with 12 adjacent samples to estimate each grid point. The nearest point was weighed using two strengths.

\[
\hat{z}(x_0) = \frac{\sum_{i=1}^{n} z(x_i) d_{ij}^{-r}}{\sum_{i=1}^{n} d_{ij}^{-r}}
\]  

where \( x_0 \) represents the estimation point and \( x_i \) represents the data points inside a selected neighborhood. The distance between the estimation point and the data points is connected to the weights \( r \) by \( d_{ij} \). The IDW method has the effect of giving data points near the interpolation point relatively substantial weights; whereas data points further away have less impact. The larger the weight, the more close-to-\( x_0 \) influence points are granted.

2.5. Partial Least-Square Regression (PLSR)

PLSR was evaluated in this study as a new method for predicting the PLI. PLSR is a versatile tool that can easily manage data when the number of input variables is much greater than the number of target variables, and the input variables have much collinearity and noise [53]. In this study, to link the input variables (four trace elements) to the output variables (PLI), PLSR was combined with leave-one-out cross-validation (LOOCV). An important step in PLSR analysis is to select the optimal number of latent variables (LVs) in order to represent the calibration data without overfitting or underfitting. The LV parameter was optimized using the LOOCV in terms of the lowest value of the RMSE. Random 10-fold cross-validation was applied on the datasets to increase the robustness of the results as indicated by Unscrambler X Version 10.2 (CAMO Software AS, Oslo, Norway). The performance of PLSR models based on four trace elements for each soil layer was evaluated to predict the PLI. The best model for both calibration (Cal.) and validation (Val.) was chosen according to the lowest value of the root mean square error (RMSE) and mean absolute deviation (MAD) as well as the highest determination coefficient (\( R^2 \)) and accuracy (Acc).
The absolute variance fraction, $R^2$, is computed as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(PLIo_i - LIo_i)^2}{\sum_{i=1}^{n}(PLIo_i)^2}$$  \hspace{1cm} (5)$$

The RMSE shows how well the model (absolutely) fits the data points. The RMSE is a relative measure of fit that determines the best absolute values, with smaller RMSE values suggesting a better fit. The RMSE is calculated with the following equation:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(PLIo_i - LIo_i)^2}{n}}$$  \hspace{1cm} (6)$$

The MAD evaluates the average magnitude of errors through a series of simulations without taking direction into account. It also determines the precision of constant variables, as seen below:

$$MAD = \frac{\sum_{i=1}^{n}|PLIo_i - LIo_i|}{n}$$  \hspace{1cm} (7)$$

$WQI_o$ represents the observed value, and $n$ represents the number of data points. $WQI_p$, conversely, is the predicted value.

$$Acc = 1 - \text{abs} \left( \frac{\text{mean } PLI_p - \text{PLI}_{\text{act}}}{\text{PLI}_{\text{act}}} \right)$$  \hspace{1cm} (8)$$

$Acc$ represents the accuracy of the model, $PLI_p$ is the predicted or simulated value, and $PLI_{\text{ave}}$ is the average value.

2.6. Multiple Linear Regression (MLR)

MLR was evaluated in this study as a new method like PLSR for predicting the PLI. MLR analyzes a dependent parameter (PLI) using two or more independent parameters (four trace elements). MLR attempts to model the linear relationship between the independent and the response (dependent) variable. The best model for both calibration (Cal.) and validation (Val.) was also chosen according to the lowest RMSE and mean absolute deviation (MAD) as well as the highest $R^2$. The least-square approach was used to calculate the parameters, using the regression equation, that minimized the sum of the errors squared.

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + \epsilon$$  \hspace{1cm} (9)$$

where, for $i = n$ observations, $Y_i$ is the dependent variable, $x_i$ = explanatory variables, $\beta_0$ = the y-intercept (constant term), $\beta_p$ = the slope coefficients for each explanatory variable, and $\epsilon$ = the model’s error term (also known as the residuals).

2.7. Statistical Analysis

This statistical analysis was performed using SPSS (v. 12.0, SPSS Inc., Chicago, IL, USA). The relationship between the observed and predicted value of the PLI derived from PLSR was modeled using a simple linear regression. The significance level of the coefficients of determination ($R^2$) for these relationships was set at 0.05.

3. Results and Discussion

3.1. The Variation of Four Trace Elements in Three Different Layers of Soil

In this study, contamination with Ni, Co, Cr, and B in three different layers of soil was assessed. There were wide variations in the values of the four trace elements in the three different layers. The Ni varied from 0 to 2720 mg kg$^{-1}$, Co varied from 0 to 2694 mg kg$^{-1}$, Cr varied from 0 to 2327 mg kg$^{-1}$, and B varied from 0 to 3551 mg kg$^{-1}$ (Table S2 and Table 4) across the three different layers. The trace element concentrations in the soil therefore ranked, in descending order, as B > Ni > Cr > Co. There were no
significant differences in the values of Ni, Co, Cr, and B between the three layers of the soil, as shown in Table 4.

Table 4. Minimum (min), maximum (max), mean, and standard deviation (SD) of four trace elements. Means in the same row followed by different letters are significantly different at $p < 0.05$.

| Measured Parameters | Surface Soil (0–30 cm) | Subsurface Soil (30–60 cm) | Underground Soil (60–100 cm) |
|---------------------|------------------------|-----------------------------|-------------------------------|
|                     | Min  | Max  | Mean | SD  | Min  | Max  | Mean | SD  | Min  | Max  | Mean | SD  |
| Ni (mg kg$^{-1}$)   | 10   | 2580 | 1737 | a  | 711  | 2702 | 1672 | a  | 671  | 2720 | 1584 | a  | 542 |
| Co (mg kg$^{-1}$)   | 495  | 1905 | 1392 | a  | 579  | 2158 | 1058 | ab | 452  | 2694 | 1295 | b  | 502 |
| Cr (mg kg$^{-1}$)   | 452  | 2327 | 1454 | a  | 373  | 717  | 2217 | a  | 496  | 783  | 2018 | a  | 420 |
| B (mg kg$^{-1}$)    | 479  | 3231 | 2021 | a  | 998  | 40   | 3072 | a  | 797  | 3551 | 1820 | a  | 804 |

3.2. Assessment of Contamination Risk Using Geoaccumulation Index

In this study, the contamination of soils was assessed based on the I-geo. The I-geo values indicated widespread pollution by Ni, Co, Cr, and B in the different layers of soil. I-geo (Ni) varied from 0 to 4.74, I-geo (Co) varied from 0 to 6.56, I-geo (Cr) varied from 0 to 4.11, and I-geo (B) varied from 0 to 4.57 (Table 5). There were no significant differences in the values of I-geo (Ni), I-geo (Co), I-geo (Cr), and I-geo (B) between the three layers of the soil, as shown in Table 5.

Table 5. Minimum (min), maximum (max), mean, and standard deviation (SD) of geoaccumulation index (I-geo) for four trace elements. Means in the same row followed by different letters are significantly different at $p < 0.05$.

| Measured Parameters | Surface Soil (0–30 cm) | Subsurface Soil (30–60 cm) | Underground Soil (60–100 cm) |
|---------------------|------------------------|-----------------------------|-------------------------------|
|                     | Min  | Max  | Mean | SD  | Min  | Max  | Mean | SD  | Min  | Max  | Mean | SD  |
| I-geo (Ni)          | 0.00 | 4.66 | 3.82 | a  | 1.05 | 2.45 | 3.87 | a  | 0.55 | 0.00 | 3.66 | a  | 1.33 |
| I-geo (Co)          | 0.00 | 6.06 | 4.95 | a  | 1.25 | 2.70 | 5.34 | a  | 0.84 | 0.00 | 5.20 | a  | 1.36 |
| I-geo (Cr)          | 1.74 | 4.11 | 3.39 | a  | 0.58 | 2.41 | 3.28 | a  | 0.45 | 0.00 | 3.17 | a  | 0.84 |
| I-geo (B)           | 1.67 | 4.43 | 3.52 | a  | 0.74 | 0.00 | 3.35 | a  | 0.96 | 0.00 | 3.25 | a  | 1.27 |

The spatial distributions of I-geo for heavy metals in the study area are illustrated in Figures 2–5. The beryl green color represents unpolluted soils related to the heavy metals analyzed, while the strongly polluted soils are represented by the red color. The maps of the Co distributions are wholly covered by red color, indicating that the study area was strongly polluted with Co contamination in the different layers of soil. Regarding the B and Cr maps, most of the study area was mostly covered by orange color, indicating that a large part of the research area was moderately to highly polluted with B and Cr. Based on the Ni maps, it was deduced that about half of the research area was highly polluted with this metal. Additionally, the spatial distribution maps of Ni in soil layers were recognized by increasing concentrations toward the eastern parts, which may be due to the proximity to El Gharbia main drain.

The results show that I-geo values of four trace elements indicated non-polluted to highly contaminated conditions. The I-geo values show wide variations in the values Ni, Co, Cr, and B in the different layers of soil (Table S3 & Figures 2–5). The Ni, Cr, and B according to I-geo classification ranged from uncontaminated/moderately contaminated to strongly/extremely contaminated. While, Co according to I-geo classification ranged from uncontaminated/moderately contaminated to extremely contaminated (Table S3 and Table 6). The highest values of I-geo Ni, Co, Cr, B were found at 1Und, 14Und, 17S and 4Und, respectively. The I-geo (Ni) revealed that all the soil samples were contaminated except for three samples (5S, 20Und, and 21Und), which were in the non-polluted category, as shown in Table S3. The I-geo (Co) showed a high accumulation effect in many samples and indicates severe contamination in most samples. There was also a non-polluted in the samples 15S and 21Und as shown in Table S3. I-geo (Cr) and I-geo (B) showed high
contamination in most samples, but they indicate a medium degree of contamination in many samples, and non-polluted was present in the samples of Cr at 21Und and B at 11Sub and 9-21Und.

These results indicate that the three soil layers in Kafr El-Sheikh Governorate contain large quantities of Ni, Co, Cr and B. The main sources of Ni, Co and Cr pollution in soils are the metal plating, fossil fuel combustion, Ni mining, and electroplating industries, anthropogenic sources such as sewage sludge and other wastes used as soil conditioners, agricultural fertilizers especially phosphates, atmospheric deposition and in inorganic fertilizers [54–59]. The main sources of B contamination in soil are the borosilicate mineral tourmaline, underground currents, groundwater, and seawater [60].

The main sources of B contamination in soil are the borosilicate mineral tourmaline, underground currents, groundwater, and seawater [60].

Figure 2. Spatial distribution of I-geo (Ni) in the three different layers of soils.
Figure 3. Spatial distribution of I-geo (Co) in the three different layers of soils.

Figure 4. Cont.
According to the I-geo classification, 33.4% and 57.1% of the soil samples in the surface layer (S), 52.3% and 38.1% in the subsurface layer (Sub), and 28.5% and 57.1% in the underground layer (Und) were strongly contaminated and strongly/extremely contaminated, respectively, with Ni (Table 6). For Co, about 71.4% and 90.4% were strongly/extremely contaminated in the S and Sub layers, respectively, while 66.6% of the soil samples in the Und layer were extremely contaminated (Table 6). For Cr, about 76.1%, 76.1%, and 71.4% were strongly contaminated in the S, Sub, and Und layers, respectively. About 47.6%
and 33.3% of the soil samples from the S layer, as well as 52.3% and 23.8% from the Sub layer, were strongly contaminated and strongly/extremely contaminated, respectively. Additionally, For B, 14.3% and 76.1% of the soil samples in Und were moderately/strongly contaminated and strongly contaminated. The rest of the soil samples from the three layers were uncontaminated/moderately contaminated to moderately contaminated for all the trace elements.

Table 6. Classification of soil contamination level based on I-geo value for Ni, Co, Cr, and B in three soil layers.

| Soil            | I-geo Value | Number of Samples (Percent) | Ni   | Co   | Cr   | B    |
|-----------------|-------------|-------------------------------|------|------|------|------|
| Surface layer   | I-geo ≤ 0   | - (0%)                        | - (0%)| - (0%)| - (0%)| - (0%)|
|                 | 0 < I-geo < 1| 1 (4.8%)                      | 1 (4.8%)| - (0%)| - (0%)| - (0%)|
|                 | 1 < I-geo < 2| - (0%)                        | - (0%)| 1 (4.8%)| 1 (4.8%)| - (0%)|
|                 | 2 < I-geo < 3| 1 (4.8%)                      | - (0%)| 1 (4.8%)| 1 (4.8%)| - (0%)|
|                 | 3 < I-geo < 4| 7 (33.4%)                     | 5 (23.8%)| 16 (76.1%)| 10 (47.6%)| - (0%)|
|                 | 4 < I-geo < 5| 12 (57.1%)                    | 15 (71.4%)| 1 (4.8%)| 7 (33.3%)| - (0%)|
|                 | 5 < I-geo   | - (0%)                        | - (0%)| - (0%)| - (0%)| - (0%)|
| Subsurface layer| I-geo ≤ 0   | - (0%)                        | - (0%)| - (0%)| - (0%)| - (0%)|
|                 | 0 < I-geo < 1| - (0%)                        | - (0%)| - (0%)| - (0%)| 1 (4.8%)|
|                 | 1 < I-geo < 2| - (0%)                        | - (0%)| - (0%)| - (0%)| - (0%)|
|                 | 2 < I-geo < 3| 2 (9.6%)                      | 1 (4.8%)| 4 (19.1%)| 4 (19.1%)| - (0%)|
|                 | 3 < I-geo < 4| 11 (52.3%)                    | 1 (4.8%)| 16 (76.1%)| 11 (52.3%)| - (0%)|
|                 | 4 < I-geo < 5| 8 (38.1%)                     | 19 (90.4%)| 1 (4.8%)| 5 (23.8%)| - (0%)|
|                 | 5 < I-geo   | - (0%)                        | - (0%)| - (0%)| - (0%)| - (0%)|
| Underground layer| I-geo ≤ 0   | - (0%)                        | - (0%)| - (0%)| - (0%)| - (0%)|
|                 | 0 < I-geo < 1| 2 (9.6%)                      | 1 (4.8%)| 1 (4.8%)| 2 (9.6%)| - (0%)|
|                 | 1 < I-geo < 2| - (0%)                        | - (0%)| - (0%)| - (0%)| - (0%)|
|                 | 2 < I-geo < 3| 1 (4.8%)                      | - (0%)| 5 (23.8%)| 3 (14.3%)| - (0%)|
|                 | 3 < I-geo < 4| 6 (28.5%)                     | - (0%)| 15 (71.4%)| 16 (76.1%)| - (0%)|
|                 | 4 < I-geo < 5| 12 (57.1%)                    | 6 (28.6%)| - (0%)| - (0%)| - (0%)|
|                 | 5 < I-geo   | - (0%)                        | 14 (66.6%)| - (0%)| - (0%)| - (0%)|

3.3. Assessment of Contamination Risk Using Pollution Load Index

In this study, the pollution of soils was assessed on the basis of PLI values. The PLI_S values varied from 0 to 35.50, and the mean value was 22.86. The PLI_Sub values varied from 11 to 34.43, and the mean was 23.87. The PLI_Und values varied from 0 to 32.61, and the mean was 223.20 (Table 7). These results agree with those of Elbehiry et al. [61], who evaluated the risks for four trace elements based on the PLI in soils of the Nile Delta close to the studied area. They found that the PLI ranged from 0.03 to 23.36 across the studied soils.

Table 7. Minimum (Min), maximum (Max), mean, and standard deviation (SD) of pollution load index under different soil layers. PLI_S is pollution load index in surface layer; PLI_Sub is pollution load index in subsurface layer; PLI_Und is pollution load index in underground layer.

|                   | PLI_S   | PLI_Sub  | PLI_Und  |
|-------------------|---------|----------|----------|
| Min               | 0.0     | 11.9     | 0.0      |
| Max               | 35.5    | 34.4     | 32.6     |
| Mean              | 22.9    | 23.9     | 22.2     |
| SD                | 8.2     | 6.2      | 10.0     |

According to the PLI classification, 4.76% and 95.24% of the samples were unpolluted and very highly polluted, respectively, in the S layer of the soil profiles. All the samples were very highly polluted in the Sub layer of the soil profiles. Additionally, 14.29% and
85.71% of the samples were unpolluted and very highly polluted, respectively, in the Und layers of the soil profiles (Table 8). This region is one of Egypt’s most inhabited, fertile, and cultivated, which means it supports a large population. Furthermore, since it is close to the sea and hosts many of Egypt’s industrial and agricultural resources, as well as much domestic drainage in the Nile Delta, this region is subjected to many stresses. As well as the sources of these elements are the industrial drainage in the El Gharbia main drain [61].

Table 8. Pollution load index values and pollution levels in soil.

| Profile No. | PLI_S | Pollution Level | PLI_Sub | Pollution Level | PLI_Und | Pollution Level |
|-------------|-------|----------------|---------|----------------|---------|----------------|
| 1           | 22.31 | V.H.P.         | 26.22   | V.H.P.         | 21.95   | V.H.P.         |
| 2           | 30.18 | V.H.P.         | 19.88   | V.H.P.         | 29.17   | V.H.P.         |
| 3           | 19.81 | V.H.P.         | 29.71   | V.H.P.         | 32.61   | V.H.P.         |
| 4           | 31.55 | V.H.P.         | 32.95   | V.H.P.         | 30.33   | V.H.P.         |
| 5           | 6.42  | V.H.P.         | 22.74   | V.H.P.         | 21.38   | V.H.P.         |
| 6           | 17.58 | V.H.P.         | 30.60   | V.H.P.         | 31.07   | V.H.P.         |
| 7           | 35.50 | V.H.P.         | 34.43   | V.H.P.         | 27.76   | V.H.P.         |
| 8           | 24.72 | V.H.P.         | 23.64   | V.H.P.         | 24.10   | V.H.P.         |
| 9           | 30.83 | V.H.P.         | 26.91   | V.H.P.         | 0.00    | U.P.           |
| 10          | 33.32 | V.H.P.         | 21.37   | V.H.P.         | 24.88   | V.H.P.         |
| 11          | 25.16 | V.H.P.         | 11.90   | V.H.P.         | 26.70   | V.H.P.         |
| 12          | 24.09 | V.H.P.         | 34.33   | V.H.P.         | 25.32   | V.H.P.         |
| 13          | 20.56 | V.H.P.         | 20.77   | V.H.P.         | 22.43   | V.H.P.         |
| 14          | 19.82 | V.H.P.         | 24.61   | V.H.P.         | 25.37   | V.H.P.         |
| 15          | 0.00  | U.P.           | 20.93   | V.H.P.         | 30.81   | V.H.P.         |
| 16          | 24.11 | V.H.P.         | 21.60   | V.H.P.         | 26.87   | V.H.P.         |
| 17          | 21.06 | V.H.P.         | 14.94   | V.H.P.         | 16.29   | V.H.P.         |
| 18          | 24.14 | V.H.P.         | 23.54   | V.H.P.         | 23.14   | V.H.P.         |
| 19          | 21.02 | V.H.P.         | 26.03   | V.H.P.         | 26.06   | V.H.P.         |
| 20          | 23.44 | V.H.P.         | 16.75   | V.H.P.         | 0.00    | U.P.           |
| 21          | 24.56 | V.H.P.         | 17.50   | V.H.P.         | 0.00    | U.P.           |

PLI_S is pollution load index in the surface layer; PLI_Sub is pollution load index in subsurface layer; PLI_Und is pollution load index in underground layer. U.P. is unpolluted; V.H.P. is very highly polluted.

3.4. Performance of PLSR and MLR Models in Predicting the PLI

Mathematical techniques can be used to calculate the PLIs of soil sites with high accuracy [47]. These methods, however, are time consuming, since they require many mathematical equations to convert several numbers of trace element data into a single value that represents the soil pollution levels. By contrast, the PLSR and PLR methods are easy and do not need several steps for calculating the PLI. The multivariate regression models such as PLSR and MLR have recently been used as alternative methods to predict the water quality indices based on data for several trace elements [37,39].

To reduce the strongly collinear independent variables to a small minority of orthogonal factors, multivariate statistical techniques, PLSR models, based on the four trace elements, were tested for predicting the PLSR techniques can be used to identify optimized models that enhance the efficiency when searching for optimized relationships [53,62–65]. The calibration (Cal.) models of PLSR and MLR performed the best in predicting the PLI based on four trace elements, with $R^2 = 0.91–0.95$ in the surface layer, $0.94–0.97$ in the subsurface layer, $0.92–0.99$ in the underground layers, and $0.93–0.97$ across the three different layers (Tables 9 and 10). The validation (Val.) models performed the best in predicting the PLI based on data for four trace elements, with $R^2 = 0.89–0.93$ in the surface layer, $0.91–0.96$ in the subsurface layer, $0.89–0.94$ in the underground layers, and $0.92–0.94$ across the three different layers (Tables 9 and 10). In general, the Cal. and Val. of PLSR models performed better in predicting the PLI for the three different layers than the MLR models (Tables 9 and 10; Figures 6 and 7). The Cal. and Val. of PLSR models showed a higher $R^2$ and lower RMSE and MAD than the MLR models. For example, the RMSE and MAD for the Cal. Models of PLSR were 2.19 and 1.68 in the surface layer, 1.07 and
0.83 in the subsurface layer, 1.71 and 1.24 in the underground layers, and 0.94 and 1.19 across the three different layers. The RMSE and MAD for the Cal. Models of PLSR were 2.61 and 2.21 in the surface layer, 1.53 and 1.27 in the subsurface layer, 3.15 and 2.19 in the underground layers, and 1.44 and 1.16, respectively, across the three different layers (Tables 9 and 10). PLSR analysis selected the optimal number of latent variables (LVs) from 1 to 2 in order to represent the calibration data without overfitting or underfitting. The PLSR and MLR models showed a very small drop in the quality of the performance measures ($R^2$, RMSE, MAD, and Acc) when moving from the calibrating stage to the testing stage (Tables 9 and 10; Figures 6 and 7). A significant positive relationship was also obtained between the measured and predicted of I-geo and PLI in case of calibrating and validating values (Tables 9 and 10; Figures 6 and 7).

Table 9. Results of calibration ($R^2_{\text{cal}}$, RMSE$_C$, MAD$_C$, and Acc$_C$), and ten-fold cross-validation ($R^2_{\text{val}}$, RMSE$_v$, MAD$_v$, and Acc$_v$): partial least-square regression models of the relationships between four trace elements and pollution load index (PLI). ***: $p < 0.001$.

| Pollution Load Index | Layers    | LVs | Calibration | Validation |
|----------------------|-----------|-----|-------------|------------|
|                      |           |     | $R^2_{\text{cal}}$ | RMSE$_C$ | MAD$_C$ | Acc$_C$ | $R^2_{\text{val}}$ | RMSE$_v$ | MAD$_v$ | Acc$_v$ |
| PLI                  | Surface   | 2   | 0.95 *** | 2.19 | 1.68 | 0.94 | 0.93 *** | 2.31 | 1.80 | 0.96 |
|                      | Subsurface| 1   | 0.97 *** | 1.07 | 0.83 | 0.98 | 0.96 *** | 1.23 | 0.97 | 0.98 |
|                      | Underground| 2   | 0.99 *** | 1.71 | 1.24 | 0.99 | 0.94 *** | 2.46 | 1.51 | 0.97 |
|                      | All layers| 2   | 0.97 *** | 1.19 | 0.94 | 0.98 | 0.94 *** | 1.50 | 1.15 | 0.97 |

Table 10. Results of calibration ($R^2_{\text{cal}}$, RMSE$_C$, MAD$_C$, and Acc$_C$), and cross-validation ($R^2_{\text{val}}$, RMSE$_v$, MAD$_v$, and Acc$_v$): multiple linear regression models of the relationships between four trace elements and pollution load index. ***: $p < 0.001$.

| Pollution Load Index | Layers    | Calibration | Validation |
|----------------------|-----------|-------------|------------|
|                      |           | $R^2_{\text{cal}}$ | RMSE$_C$ | MAD$_C$ | Acc$_C$ | $R^2_{\text{val}}$ | RMSE$_v$ | MAD$_v$ | Acc$_v$ |
| PLI                  | Surface   | 0.91 *** | 2.61 | 2.21 | 0.99 | 0.89 *** | 2.89 | 2.48 | 0.89 |
|                      | Subsurface| 0.94 *** | 1.53 | 1.27 | 0.99 | 0.91 *** | 1.88 | 1.59 | 0.99 |
|                      | Underground| 0.92 *** | 3.15 | 2.19 | 0.95 | 0.89 *** | 3.65 | 2.64 | 0.95 |
|                      | All layers| 0.93 *** | 1.44 | 1.16 | 0.97 | 0.92 *** | 1.69 | 1.37 | 0.89 |

To the best of our knowledge, the issue of predicting PLI using PLSR and MLR models, based on trace element data, has not been addressed to date. The multivariate regression models were recently shown to perform well in predicting the water quality indices [37,39]. For example, Gad et al. [37] found that PLSR based on data for several trace elements accurately estimated four pollution indices for water and the drinking water quality index (DWQI) for both the Cal. and Val. Models; $R^2$ varied from 0.98 to 1.00 for the Cal. and from 0.88 to 0.99 for the Val. Elsayed et al. [39] found that principal component regression (PCR) and support vector machine regression (SVMR) represented robust models for predicting six water quality indices in the Cal. and Val. models; $R^2$ varied from 0.48 to 0.99. Finally, the results obtained from this study prove that both PLSR and MLR have the potential to predict the PLI in and across three different layers.
Figure 6. Comparison between measuring series, calibrating series, and validating series for pollution load index in and across different layers (a) surface soil, (b) subsurface soil, (c) underground soil and (d) all layers using the PLSR models based on data for four trace elements. Statistical analysis was presented in Table 9.
Figure 7. Comparison between measuring series, calibrating series, and validating series for pollution load index (PLI) in and across different layers (a) surface soil, (b) subsurface soil, (c) underground soil and (d) all layers using the MLR models based on data for four trace elements. Statistical analysis was presented in Table 10.
4. Conclusions

Trace metal (Ni, Co, Cr, and B) contamination in the north Nile Delta of Egypt was assessed based on the I-geo and PLI. The distribution patterns of trace metals in the three layers of the soil profiles indicate high pollution. The status regarding Ni, Cr, and B, according to I-geo classification, ranged from uncontaminated/moderately contaminated to strongly/extremely contaminated. Co ranged from uncontaminated/moderately contaminated to extremely contaminated. According to the I-geo classification, 33.4% and 57.1% of the soil samples in the surface layer (S), 52.3% and 38.1% in the subsurface layer (Sub), and 28.5% and 57.1% in the underground layer (Und) were strongly contaminated and strongly/extremely contaminated, respectively, with Ni. The PLI showed that the majority of the samples were very highly polluted. According to the PLI classification, 4.76% and 95.24% of the samples were unpolluted and very highly polluted, respectively, in the S layer of the soil profiles. All the samples were very highly polluted in the Sub layer of the soil profiles. Additionally, 14.29% and 85.71% of the samples were unpolluted and very highly polluted, respectively, in the Und layers of the soil profiles; thus, high concentrations of the four trace elements may present potential health risks for the human populations residing in the surrounding area. The deterioration of soil quality in this area can be attributed to large applications of agricultural resources, industrial activities, and poor drainage and its location close to the sea and Lake Burullus. The PLSR and MLR models showed robust performance in estimating the PLI in and across different soil layers, showing the highest $R^2$ values, lowest RMSE and MAD values, and greatest slope values in calibration and validation datasets. The PLSR and MLR models showed a very small drop in the quality of the performance measures ($R^2$, RMSE, MAD, and Acc) when moving from the calibrating stage to the testing stage. The Cal. and Val. of PLSR models showed a higher $R^2$ and lower RMSE and MAD than the MLR models. For example, the RMSE and MAD for the Cal. Models of PLSR were 2.19 and 1.68 in the surface layer, 1.07 and 0.83 in the subsurface layer, 1.71 and 1.24 in the underground layers, and 0.94 and 1.19 across the three different layers, respectively. Future studies should evaluate both the PLSR and MLR models under different environmental conditions for different soils.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/su13148027/s1, Table S1. Physiochemical parameters in three layers for different soil profiles; Table S2. Trace element concentrations in soil samples; Table S3. Geoaccumulation index values and contamination levels in soil.

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

Funding: The authors extend their appreciation to the Scientific Research Deanship at King Khalid University and the Ministry of Education in Saudi Arabia for funding this research work through the project number IFP-KKU-2020/3.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article and in the Supplementary Materials file.

Conflicts of Interest: The authors declare no conflict of interest.
31. He, J.Y.; Jia, X. Arc GIS geostatistical analyst application in assessment of MTBE contamination. In Proceedings of the ESRI User Conference, Fremont, CA, USA, 9–13 August 2004. [CrossRef]

32. Zheng, C. Using multivariate analyses and GIS to identify pollutants and their spatial patterns in urban soils in Galway, Ireland. Environ. Pollut. 2006, 142, 501–511. [CrossRef]

33. Moor, C.; Lymberopoulou, T.; Dietrich, VJ. Determination of heavy metals in soils, sediments and geological materials by ICP-AES and ICP-MS. Microchim. Acta 2001, 136, 123–128. [CrossRef]

34. Garriga, M.; Romero-Bravo, S.; Estrada, F.; Escobar, A.; Matus, I.A.; del Pozo, A.;Astudillo, C.A.; Lobos, G.A. Assessing wheat traits by spectral reflectance: Do we really need to focus on predicted trait-values or directly identify the elite genotypes group? Front. Plant Sci. 2017, 8, 280. [CrossRef]

35. Cho, H.J.; Lee, H.J. Multiple linear regression models for predicting nonpoint-source pollutant discharge from a highland agricultural region. Water 2018, 10, 1156. [CrossRef]

36. Wei, L.; Huang, C.; Zhong, Y.; Wang, Z.; Hu, X.; Lin, L. Inland waters suspended solids concentration retrieval based on PSO-LSSVM for UAV-borne hyperspectral remote sensing imagery. Remote Sens. 2019, 11, 1455. [CrossRef]

37. Gad, M.; Elsayed, S.; Moghann, FS.; Almarshadi, M.H.; Alshammarri, A.S.; Khedher, K.M.; Eid, E.M.; Hussein, H. Combining water quality indices and multivariate modeling to assess Surface water quality in the Northern Nile Delta, Egypt. Water 2020, 12, 2142. [CrossRef]

38. Feng, M.; Guo, X.; Wang, C.; Yang, W.; Shi, C.; Ding, G.; Zhang, X.; Xiao, L.; Zhang, M.; Song, X. Monitoring and evaluation in freeze stress of winter wheat (Triticum aestivum L.) through canopy hyper-spectrum reflectance and multiple statistical analysis. Ecol. Indic. 2018, 84, 290–297. [CrossRef]

39. Elsayed, S.; Hussein, H.; Moghann, FS.; Khedher, K.M.; Eid, E.M.; Gad, M. Application of irrigation water quality indices and multivariate statistical techniques for surface water quality assessments in the Northern Nile Delta, Egypt. Water 2020, 12, 3300. [CrossRef]

40. Shaltout, K.H.; Galal, T.M.; El-Komi, T.M. Evaluation of the nutrient status of some hydrophytes in the water courses of Nile Delta, Egypt. J. Bot. 2009, 36, 77–87. [CrossRef]

41. Molle, F.; Gaafar, I.; El-Agha, D.E.; Rap, E. The Nile delta’s water and salt balances and implications for management. Agric. Water Manag. 2018, 197, 110–121. [CrossRef]

42. Kabata-Pendias, A. Trace Elements in Soils and Plants, 4th ed.; CRC Press: Boca Raton, FL, USA, 2011.

43. Gowd, S.S.; Reddy, M.R.; Govil, P.K. Assessment of heavy metal contamination in soils at Jajmau (Kanpur) and Unnao industrial areas of the Ganga Plain, Uttar Pradesh, India. J. Hazard. Mater. 2010, 174, 113–121. [CrossRef] [PubMed]

44. Awadh, S.M. Assessment of the potential nickel and lead in the road-side dust in the Karkh City along the highway between Ramadi and Rutba, West of Iraq. Merit Res. J. Environ. Sci. Toxicol. 2003, 1, 126–135.

45. Lu, X.; Wang, L.; Lei, K.; Huang, J.; Zhai, Y. Contamination assessment of copper, lead, zinc, manganese and nickel in street dust of Baoji, NW China. J. Hazard. Mater. 2009, 161, 1058–1062. [CrossRef] [PubMed]

46. Rahman, S.; Khanam, D.; Adyel, T.; Islam, M.S.; Mohammad Ahsan, A.; Akbor, M.A. Assessment of heavy metal contamination of agricultural soil around Dhaka Export Processing Zone (DEPZ), Bangladesh: Implication of seasonal variation and indices. Appl. Sci. 2012, 2, 584–601. [CrossRef]

47. Jorfi, S.; Maleki, R.; Jiafarzadeh, N.; Ahmadi, M. Pollution load index for heavy metals in Mian-Ab plain soil, Khuzestan, Iran. Data Brief 2017, 15, 584–590. [CrossRef]

48. Karimi, A.; Naghizadeh, A.; Biglari, H.; Peirovi, R.; Ghasemi, A.; Zarei, A. Assessment of human health risks and pollution index for heavy metals in farmlands irrigated by effluents of stabilization ponds. Environ. Sci. Pollut. Res. 2020, 27, 10317–10327. [CrossRef]

49. Jafari, A.; Ghaderpoori, M.; Kamarehi, B.; Abdipour, H. Soil pollution evaluation and health risk assessment of heavy metals around Doroud cement factory. Iran. Environ. Earth. Sci. 2019, 78, 250. [CrossRef]

50. Ali, R.R.; Moghanm, F.S. Variation of soil properties over the landforms around Idku Lake, Egypt. Egypt. J. Remote Sens. Space Sci. 2013, 16, 91–101. [CrossRef]

51. Shepard, D. A two-dimensional interpolation function for irregularly-spaced data. In Proceedings of the 1968 23rd ACM National Conference, New York, NY, USA, 27–29 August 1968. [CrossRef]

52. Azpurua, M.A.; Ramos, K.D. A comparison of spatial interpolation methods for estimation of average electromagnetic field magnitude. Prog. Electromagn. Res. 2010, 14, 135–145. [CrossRef]

53. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. Chemom. Intell. Lab. Syst. 2001, 58, 109–130. [CrossRef]

54. Raymond, A.W.; Okeieimen, F.E. Heavy Metals in Contaminated Soils: A Review of Sources, Chemistry, Risks and Best Available Strategies for Remediation. ISRN Ecol. 2011, 2011, 402647.

55. Megrath, S.P. Nickel in Heavy Metals in Soils, 2nd ed.; Alloway, B.J., Ed.; Blackie Academic & Professional: London, UK, 1995; 371p. [CrossRef]

56. Parth, V.; Murthy, N.N.; Saxena, P.R. Assessment of heavy metal contamination in soil around hazardous waste disposal sites in Hyderabad City (India): Natural and anthropogenic implications. J. Environ. Res. Manag. 2011, 2, 27–34.

57. Senesi, G.S.; Baldassarre, G.; Senesi, N.; Radina, B. Trace element input into soils by anthropogenic activities and implications for human health. Chemosphere 1999, 39, 343–377. [CrossRef]
58. Luo, L.; Ma, Y.B.; Zhang, S.Z.; Wei, D.P.; Zhu, Y.G. An inventory of heavy metal inputs to agricultural soils in China. *J. Environ. Manag.* 2009, 90, 2524–2530. [CrossRef] [PubMed]

59. Nicholson, F.A.; Smith, S.R.; Alloway, B.J.; Carlton-Smith, C.; Chambers, B.J. An inventory of heavy metals inputs to agricultural soils in England and Wales. *Sci. Total Environ.* 2003, 311, 205–219. [CrossRef]

60. Rocks, I.T.T. *ITT Corporation ENVI Software 4.7*, White Plains: New York, NY, USA, 2009.

61. Elbehiry, F.; Elbasiouny, H.; El-Ramady, H.; Brevik, E.C. Mobility, distribution, and potential risk assessment of selected trace elements in soils of the Nile Delta, Egypt. *Environ. Monit. Assess.* 2019, 191, 713. [CrossRef]

62. El-Hendawy, S.E.; Al-Suhaibani, N.; Al-Ashkar, I.; Alotaibi, M.; Tahir, M.U.; Solieman, T.; Hassan, W.M. Combining genetic analysis and multivariate modeling to evaluate spectral reflectance indices as indirect selection tools in wheat breeding under water deficit stress conditions. *Remote Sens.* 2020, 12, 1480. [CrossRef]

63. Elsayed, S.; Elhoweity, M.; Ibrahim, H.H.; Dewir, Y.H.; Migdadic, H.M.; Schmidhalter, U. Thermal imaging and passive reflectance sensing to estimate the water status and grain yield of wheat under different irrigation regimes. *Agric. Water Manag.* 2017, 189, 98–110. [CrossRef]

64. Haenlein, M.; Kaplan, A.M. A beginner’s guide to partial least squares analysis. *Underst. Stat.* 2004, 3, 283–297. [CrossRef]

65. Çakir, R. Effect of water stress at different development stages on vegetative and reproductive growth of corn. *Field Crop. Res.* 2004, 89, 1–16. [CrossRef]