




Article

Assessment and Mapping of Soil Salinization Risk in an Egyptian Field Using a Probabilistic Approach

Sameh M. Shaddad ¹, Gabriele Buttafuoco ^{2,*} and Annamaria Castrignano ²

¹ Soil Science Department, Zagazig University, Zagazig 44511, Egypt; smshaddad@agri.zu.edu.eg

² Institute for Agricultural and Forest Systems in the Mediterranean (ISAFOM), Rende (CS), National Research Council of Italy, 00185 Roma, Italy; annamaria.castrignano@ba.irsra.cnr.it

* Correspondence: gabriele.buttafuoco@cnr.it; Tel.: +39-0984-841466

Received: 20 November 2019; Accepted: 5 January 2020; Published: 7 January 2020



Abstract: The assessment of soil salinization risk at the field scale requires modeling of the spatial variability of soil salinity. This paper presents a probabilistic approach to estimate and map a risk index using all available auxiliary information. A probabilistic methodology is proposed to estimate the conditional probability of exceeding the assigned threshold value of a generic indicator of soil salinity. A geostatistical non-parametric technique, probability kriging, was used to assess the risk of soil salinization and delineate different hazard zones within a field. The technique relies on indicator coding of information. The approach was applied to soil electrical conductivity measurements collected in an experimental field located in the Nile Delta region in Egypt, and submitted over time to trials with different fertilization treatments. The application of the method allowed delineation of a north-eastern zone in the field with a high risk of soil salinization due to its lack of cultivation for a long time and nearness to buildings that prevent water infiltration. The method proved to be quite promising from the perspective of precision agriculture and it is easily extendable to any sort of remote and proximal sensing auxiliary information, including information on the deepest layers of soil.

Keywords: GIS; indicator kriging; uniform transformation; multivariate geostatistics; probability kriging; precision agriculture

1. Introduction

Soil salinity is a hazard for agriculture productivity because of its direct and indirect effects on plant growth, which can cause a decline in crop yield and even total crop failure in severe cases [1–3]. The direct effects limit plant water uptake by reducing the osmotic potential and thus the total soil water potential, as well as through the toxicity of specific ions such as boron, chloride, and sodium [2]. The indirect effects are mainly due to an increase in the concentration of sodium ions in the exchange complex of soil, which affects some soil attributes (e.g., porosity, water retention and permeability) and processes (e.g., swelling, compaction, surface sealing) [2].

The classification of soil salinity is based on the total salt concentration of an equilibrated solution, which is commonly measured by the electrical conductivity of the saturated paste extract (EC_e), as well as on the concentration of sodium ions relative to calcium and magnesium ions [4].

Nile Delta areas are very vulnerable to soil salinization due to geomorphic, climatic and human factors [5,6]; recent studies showed that 12% of these areas are at a high hazard level (especially in the eastern Nile Delta) [6]. Intense soil salinization processes have often been started or intensified by unsustainable agricultural practices, mostly related to cereal cropping (wheat, rice, maize) [7]. Farmers in Egypt over-apply agrochemicals (fertilizers, pesticides, etc.) in order to maximize yield

productivity without paying attention to the risks that may result from such practices. The process of soil salinization is more frequent in heavy soils with high clay content due to poor drainage.

Soil salinity is then a very crucial problem for the Egyptian farmers of the Nile Delta. Also several other factors tend to increase the salinity of the soils including intense evapotranspiration, high clay content of the soils that inhibits the internal drainage of water, and soil sealing resulting from the high sodium content on the surface. Therefore, soil amelioration is often required in order to avoid significant economic losses in rural areas.

Assessment and mapping of soil salinization risk are useful tools to support sustainable land planning, and thereby, to mitigate soil degradation. This can be done by individuating specific indicators of soil salinity, and defining a methodology for appropriately weighting and combining them with a set of auxiliary information that is assumed to influence the processes of interest [8–11].

Risk assessment is generally based on the estimation of variables, which are subject to extreme uncertainty [12,13]. This uncertainty depends on both the intrinsic nature of the variables and the cost of obtaining information about them, which limits the number of samples. Risk assessment consists of observing the frequency (probability) with which specified criteria are exceeded or fail to be met [14]. A probabilistic approach is preferred to a deterministic one because mechanistic models are generally unable to describe the complex spatial and temporal variability observed in the attributes of saline soil. More specifically, the numerous factors involved in the process of soil salinization and the lack of exact knowledge of their number and interactions lead us to prefer a probabilistic approach. Such an approach is based on the assumption that our knowledge of natural phenomena is imperfect; therefore, it is necessary to determine the uncertainty of our inferences or predictions.

A statistical and probabilistic approach relies on the available information (observations) and consists of estimating the uncertainty of any particular unknown value of the variable of interest by modeling the data distribution conditional on available information. In Bayesian formalism, any information that corresponds to knowledge existing prior to the observations of the variable of interest, is called a prior distribution. Any observation of the variable of interest will update such a distribution since it contains additional information about its actual value. Such an updated distribution is called a posterior distribution. The proposed approach of risk assessment is based on this process of updating the prior distribution to create a posterior distribution. This will be discussed in detail in the methodological section.

Geostatistics is a set of probabilistic and statistical tools that are suitable for estimating attributes distributed in space [15] and can then be used for probabilistic interpolation (prediction), where an error variance is attached to each estimate using, for example, ordinary kriging [16,17]. Unfortunately, unless a particular parametric (for example, Gaussian) distribution of spatial errors is assumed, error variance cannot provide a reliable error probability distribution, which is required for risk assessment. Gaussian error distribution models are symmetric and are fully characterized by two parameters, error mean and error variance. While such models are suitable for the distribution of measurement errors in the highly controlled environment of a laboratory, they are of questionable use for spatial interpolation errors [18]. The spatial distribution of chemical concentrations in soil are generally highly skewed and thus, they are associated with strongly asymmetric probability density function models that are used for uncertainty estimation [18]. Moreover, as mentioned above, nature is rarely Gaussian and the spatial distribution of environmental/earth attributes that do not fall under human control, is almost always non-Gaussian, unlike the distribution of repetitive measurements in a “highly controlled, carefully planned, laboratory experiment” [18]. Unfortunately, this is often ignored in a lot of scientific literature.

Non-parametric geostatistics is based on no assumed parametric model of error distribution and makes the modeling of uncertainty a priority [18,19]. This uncertainty is modeled as a probability distribution of the variable of interest rather than an estimation error, as in kriging [19]. More specifically, modeling the probability that the variable at any unsampled location is greater than a given threshold value is related to the proportion of neighboring data values exceeding the same threshold. Indicator kriging [19], a non-parametric approach, is based on assigning a binary indicator of 1 or 0 depending

on whether the observation is greater or less than a threshold value. However, the binary indicator has no uncertainty because observation is assumed to be certain. The indicator approach has become quite popular in natural resources studies and many authors have used it [20–24]. However, as Journel [18] points out, observations are certainly affected by measurement errors and indicator kriging does not take uncertainty related to measurement errors into account. Moreover, indicator kriging is often faced with problems that arise when the order relationships are not respected [19]. In order to improve indicator kriging estimation, another non-parametric approach based on the cokriging estimator has also been proposed [18]. This method uses the information from all available attributes and considers the order relations of the observed values [25]. The uniform values, also called standardized ranks, of all available observed values are considered as auxiliary variables in order to improve the estimation of the probability of the attribute being greater/lower than the given threshold.

The probability kriging has been shown to perform better than indicator kriging for delineating contaminated soils [19,21] or groundwater [26].

The objective of this paper is to describe a risk assessment method based on multivariate non-parametric geostatistics and GIS, in which nine variables affecting soil salinization were selected and their spatial variability was modeled by using a probability kriging approach.

2. Materials and Methods

2.1. Study Area

An experimental field (3.1 ha) planted with different irrigated vegetable crops within the Faculty of Agriculture of Zagazig University (Sharkia Governorate, Egypt) ($31^{\circ}34'16.256''$ E, $30^{\circ}34'6.734''$ N, centroid coordinates) [27] was the study area (Figure 1). Soil in the study area is classified as Nile alluvium soil, which represents 2.58% of Egypt, and the area is old, flat agricultural land according to the land use/land cover (LULC) classification with a slope gradient of 0.0%–0.2% [28]. The soil texture is classified as heavy clay according to USDA classification [29]. Water is supplied through a surface irrigation system using water from the Nile river, and it has electrical conductivity values ranging between 0.31 and 0.46 dS m^{-1} [30].

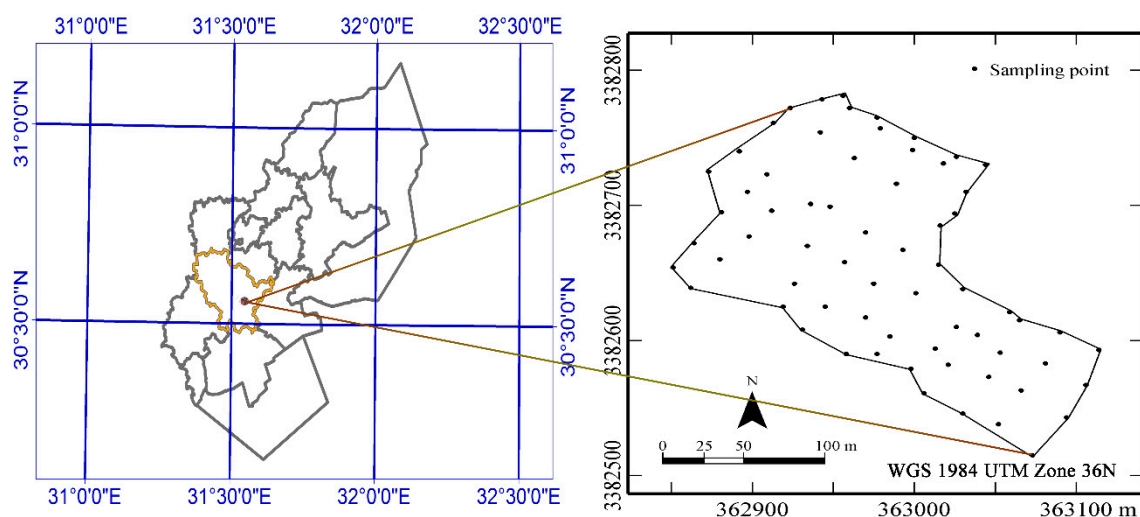


Figure 1. Location of study area and soil sampling scheme.

The climate is mainly characterized by hot dry summers and mild winters with very low annual precipitation of 90–125 mm. Climatic data were measured at the Zagazig meteorological station for the period 1983–2016. The minimum mean air temperature occurs in January (13°C) whereas the maximum mean value occurs in August (29.3°C). In the same months, the relative air humidity is 57% and 43%, respectively [31].

A total of sixty-six georeferenced soil samples from 0.0 m to 0.20 m in depth were collected using a core cylinder with a 10 cm diameter and a height of 20 cm at the nodes of a semi-regular 25 m × 25 m grid (Figure 1). At the time of sampling, the field was planted with wheat and clover.

Twelve soil chemical properties, deemed relevant for characterizing soil salinity and fertility, were selected: pH, electrical conductivity (ECe, dS m⁻¹), organic matter (OM, %), available nitrogen (N, mg kg⁻¹), available phosphorus (P, mg kg⁻¹), available potassium (K, mg kg⁻¹), soluble ions (Ca, Mg, Na, K, mEq L⁻¹), sodium adsorption ratio (SAR) and calcium carbonate content (CaCO₃, g kg⁻¹). In particular, ECe was used as the only indicator variable and its critical threshold was set to 6 dS m⁻¹, since higher values may negatively affect the potential yield of wheat crops [32]. Wheat was chosen as the reference crop because it is the main winter crop in the Nile Delta area.

In the laboratory, the soil samples were dried, crushed and sieved through a 2 mm sieve and analyzed for the selected soil properties: (a) electrical conductivity (EC, dS m⁻¹) was determined in 1:5 soil water extract (ECe) [33] and the results were converted to soil paste extract using the multiplicative factor of 7 for clayey soils [34]; (b) soluble ions (Na, K, Ca, Mg, CO₃, HCO₃, Cl, SO₄) [33]; (c) soil organic carbon (SOC) as a percent [35], which was then converted into organic matter (SOM); (d) available nitrogen NH₄-N and NO₃-N extracted by KCl 2 N solution and the extracted nitrogen was determined using the steam distillation procedure using MgO and Devarda alloy (consisting of aluminum (44%–46%), copper (49%–51%) and zinc (4%–6%) [35]; (e) the available phosphorus content (P) (mg kg⁻¹) was extracted by the Olsen method [35] with the extracted phosphorus measured calorimetrically using the ascorbic acid method with a UV-Vis-NIR spectrophotometer; and (f) available potassium (K) (mg kg⁻¹), which was extracted using 1.0 N ammonium acetate at pH 7.0 and determined using the flame photometer method [35].

2.2. Geostatistical Techniques

2.2.1. Indicator Kriging

The proposed geostatistical approach is based on a simple binary transformation, whereby each datum of the soil attribute (ECe), assumed as criteria to evaluate salinization risk, is transformed into an indicator, i.e., a binary value according to the critical threshold. By convention, data are coded as 0 or 1, if they lie below or above the given critical threshold value z_0 , in conformity with the assumed criteria for the definition of soil salinization risk:

$$i(\mathbf{x}_\alpha; z_0) = \begin{cases} 1 & \text{if } z(\mathbf{x}_\alpha) \geq z_0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The characterization of the spatial distribution requires an indicator semivariogram to be calculated using the indicator data $i(\mathbf{x}_\alpha; z_0)$:

$$\gamma_I(\mathbf{h}; z_0) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [i(\mathbf{x}_\alpha; z_0) - i(\mathbf{x}_\alpha + \mathbf{h}; z_0)]^2 \quad (2)$$

The indicator semivariogram measures how often two values $i(\mathbf{x}_\alpha; z_0)$ and $i(\mathbf{x}_\alpha + \mathbf{h}; z_0)$, separated by a vector \mathbf{h} , are on opposite sides of the threshold value z_0 [19]. An estimation of the unknown indicator value $i^*(\mathbf{x}_0; z_0)$ in an unsampled point \mathbf{x}_0 is given by:

$$i^*(\mathbf{x}_0; z_0) = \sum_{\alpha=1}^n \lambda_\alpha(\mathbf{x}_0; z_0) i(\mathbf{x}_\alpha; z_0) \quad (3)$$

where λ_α is the weight assigned to the indicator datum $i(\mathbf{x}_\alpha; z_0)$. Once a semivariogram model of the indicator variable $i(\mathbf{x}_\alpha; z_0)$ is fitted, the weights $\lambda_\alpha(\mathbf{x}_0; z_0)$ can be obtained by applying ordinary kriging to a set of n indicator data $i(\mathbf{x}_\alpha; z_0)$ at the neighboring locations \mathbf{x}_α :

$$\begin{cases} \sum_{\beta=1}^n \lambda_\beta^{OK}(\mathbf{x}_0; z_0) \gamma_I(\mathbf{x}_\alpha - \mathbf{x}_\beta; z_0) + \mu_{OK}(\mathbf{x}_0; z_0) = \gamma_I(\mathbf{x}_\alpha - \mathbf{x}_0; z_0) \\ \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_\beta^{OK}(\mathbf{x}_0; z_0) = 1 \end{cases} \quad (4)$$

where $\gamma_I(\mathbf{x}_\alpha - \mathbf{x}_\beta; z_0)$ is the semivariogram between a couple of the indicator data $i(\mathbf{x}_\alpha; z_0)$ and $i(\mathbf{x}_\beta; z_0)$, and $\gamma_I(\mathbf{x}_\alpha - \mathbf{x}_0; z_0)$ is the semivariogram between each indicator datum at the neighboring location \mathbf{x}_α and the indicator datum being estimated at \mathbf{x}_0 .

Therefore, the unknown value is estimated in Equation (3) as a weighted average of the indicator data of the same soil attribute (ECe) and the geostatistical technique is called indicator kriging (IK). It is worth noting that the previous expression is not used as an estimate of the local indicator value $i(\mathbf{x}_\alpha; z_0)$, although it is built as such, but rather as a probability estimation that the variable of interest (ECe) is greater than the critical value (6 dS m^{-1}) conditional on the available indicator information [18].

2.2.2. Probability Kriging

IK only uses the indicator part of the variable assumed to measure soil salinity rather than all the information gathered for soil characterization that relates to the process of salinization. Probability Kriging (PK) [18] is an enhancement of indicator kriging, because it uses both secondary information (the laboratory measurements of the other soil attributes) and the primary variable (indicator information of ECe) at the sampled points \mathbf{x}_α in a multivariate approach, called coindicator kriging. In our case, the target variable is a binary variable with values of either 0 or 1, whereas the correlated variables (soil attributes) are continuous numerical variables. Large differences may occur between the units of measurement of secondary information and the indicator transforms, which may cause instability problems when solving the cokriging system.

Since indicator data $i(\mathbf{x}_\alpha; z_0)$ are valued either 0 or 1, it is necessary to first transform the auxiliary variables $z(\mathbf{x}_\alpha)$ to the range (0 to 1) through rank order transform [18]:

$$k(\mathbf{x}_\alpha) = \frac{r(\mathbf{x}_\alpha)}{N} \quad (5)$$

where $r(\mathbf{x}_\alpha) \in (1, N)$ is the rank of the datum $z(\mathbf{x}_\alpha)$ when data are ranked in increasing order, and N the total number of sample locations.

Therefore, the PK approach requires calculation of the linear model of regionalization (LMC) of the whole set of experimental semivariograms of both direct indicator and rank-order transformed variables, and the cross-semivariograms between each pair of variables, regardless of its type (indicator or rank) [19,36].

PK is then the ordinary cokriging of the indicator transform $i^*(\mathbf{x}_\alpha; z_0)$ and the rank-order transform $k(\mathbf{x}_\alpha)$ of each secondary variable; the probability kriging estimate $i^*(\mathbf{x}_0; z_0)$ is given by:

$$i^*(\mathbf{x}_0, z_0) = \sum_{\alpha=1}^{n_1} \lambda_\alpha(\mathbf{x}_0, z_0) i(\mathbf{x}_\alpha, z_0) + \sum_{j=1}^{n_b} \sum_{\alpha=1}^{n_j} \nu_{\alpha,j}(\mathbf{x}_0, z_0) k_j(\mathbf{x}_\alpha) \quad (6)$$

where n_1 and n_j are the number of observations of the indicator and rank-order transformed variable j within the neighborhood of \mathbf{x}_0 , respectively; and n_b is the number of secondary variables to take into account. The weights $\lambda_\alpha(\mathbf{x}_0; z_0)$ and $\nu_{\alpha,j}(\mathbf{x}_0; z_0)$ are obtained by solving the ordinary cokriging system [19].

In our study n_1 and n_j are the same (isotopic case), however, this is not the case with heterotopic cokriging or when auxiliary information is exhaustively recorded on a grid (proximal and remote sensing) [37].

As for IK, it is of paramount importance to realize that Equation (3) does not give the expected (mean) value of the indicator, but rather, the estimated conditional probability that the EC_e exceeds the critical value of 6 dS m^{-1} by using all available information, both binary and numerical. Equation (3) can be extended to any type of auxiliary information including subsoil and nominal variables after transforming the latter to indicator/dummy variables.

2.2.3. Comparison of the Procedures

IK and PK were compared by applying cross-validation [38], which consists of removing one observation at a time from the data set and then estimating it from the remaining data and the corresponding error. The error distributions were compared by calculating different statistics: minimum, mean, maximum, values, standard deviation (STD) and root mean squared error (RMSSE), which is the mean squared difference between the observed value and the estimated value standardized by kriging variance. If the model is accurate, RMSSE should be close to unity [39]. Moreover, the correlation coefficient between observation and estimation (r), and between standardized error and estimation (ρ) and the number of outliers (n) were calculated.

3. Results and Discussion

Table 1 shows the basic statistics for the twelve study variables. The table shows that there is large variability in soil salinity over the study area with EC_e values varying from 1.86 to 54.78 dS m^{-1} . Moreover, the data distributions of all variables except pH are strongly asymmetrical with positive skewness.

Table 1. Basic statistics of soil properties.

Variable	Min	Max	Mean	Median	St. Dev.	Skewness	Kurtosis
EC_e (dS m^{-1})	1.86	54.78	7.04	3.96	8.58	3.69	18.76
Ca (mEq L^{-1})	3.50	87.50	12.15	8.40	13.80	4.15	20.86
CaCO_3 (g kg^{-1})	29.00	57.00	40.15	39.50	5.93	0.50	2.69
K (mEq L^{-1})	0.41	18.80	3.51	2.18	3.67	2.30	9.16
Mg (mEq L^{-1})	0.00	99.40	12.26	7.00	16.26	3.53	17.11
Na (mEq L^{-1})	25.27	381.90	57.66	44.77	47.25	5.06	34.06
OM (%)	3.22	9.41	4.71	4.57	0.83	2.71	16.54
SAR (-)	10.07	39.51	16.98	16.35	4.03	2.63	15.51
pH (-)	7.64	8.16	7.99	8.00	0.10	-0.98	3.86
Avail. K (mg kg^{-1})	240.84	813.87	420.98	380.66	134.67	1.35	4.02
Avail. N (mg kg^{-1})	23.66	95.55	46.78	46.41	12.82	0.85	4.74
Avail. P (mg kg^{-1})	11.19	28.23	18.70	17.56	3.68	0.39	2.41

In addition, the high values of kurtosis, which are much greater than 3 (characteristic of normal distribution), indicate that most values of each variable are quite similar to the mean but with a few large values that differ significantly from the mean. The only exception is pH, which has an average value very close to the median value. These characteristics of the data distributions lead to the assumption that the soil is quite uniform but with some spots that are characterized by very high, measured attribute values.

Table 2 shows the linear correlation coefficients for the different pairs of variables. Most variables are significantly correlated with the exception of OM and available N.

Table 2. Correlation coefficients of the soil properties.

Variable	ECe	Ca	CaCO ₃	K	Mg	Na	OM	SAR	pH	Av. K	Av. N	Av. P
ECe	1											
Ca	0.96	1										
CaCO ₃	0.42	0.27	1									
K	0.92	0.83	0.51	1								
Mg	0.98	0.94	0.38	0.89	1							
Na	0.95	0.89	0.43	0.85	0.92	1						
OM	0.04	0.01	-0.01	0.05	0.06	0.01	1					
SAR	0.71	0.58	0.55	0.70	0.62	0.83	-0.03	1				
pH	-0.43	-0.49	-0.13	-0.39	-0.41	-0.38	0.21	-0.16	1			
Av. K	0.36	0.23	0.52	0.65	0.37	0.28	0.13	0.33	-0.09	1		
Av. N	0.13	0.20	-0.20	0.14	0.13	0.05	0.05	-0.06	0.02	0.12	1	
Av. P	0.40	0.31	0.41	0.48	0.39	0.38	0.03	0.43	0.16	0.47	0.15	1

In particular, ECe is strongly correlated with Ca, K, Mg, Na and SAR and to a lesser extent with CaCO₃, pH, and available K and P, while it is not significantly correlated with OM or available N. Thus, it seems that the organic component of the soil does not significantly affect the process of salinization.

The indicator transformation of ECe (I_EC) shows that 39% of the data has a value above the critical threshold of 6 dS m⁻¹. This suggests that soil salinization might be fairly confined but might also be a warning to take action with some soil remediation work.

However, the above results assume that there is no spatial correlation between observations. Only a geostatistical structural analysis that specifically considers the cross-semivariograms, can identify the variables that actually influence the spatial process of soil salinization.

A preliminary variographic analysis between the ECe indicator (I_EC) and the rank-order transformed variables showed that the cross-semivariograms between I_EC and the transforms of OM, available N and pH did not exhibit sills significantly different from 0. It was then decided to exclude them from the subsequent multivariate analysis.

An isotropic linear model of co-regionalization (LMC) was fitted to the set of direct and cross-semivariograms of I_EC and the eight rank-order transformed auxiliary variables consisting of the following spatial structures: a nugget effect and a cubic model with a range of 203 m (Table 3). The proportion of the random component (error) with respect to the total variance was obtained by adding the nine eigenvalues of the co-regionalization matrix for the nugget effect (0.132, 0.084, 0.045, 0.042, 0.034, 0.030, 0.016, 0.005, 0.002) [37] and dividing by the total variance of the variables (indicator + uniform transformed variables). It was approximately equal to 0.38, which is quite high and could be reduced by intensifying the point sampling or/and by using a more densely sampled auxiliary variable with proximal sensing [40].

The goodness of fit was evaluated through a cross-validation test, and the two statistics of mean and standard deviation of the error standardized by cokriging standard deviation [39] were equal to 0.024 and 1.024, respectively, therefore, they quite close to the optimal values of 0 and 1.

Figure 2 shows the fitted LMC, which shows a sufficiently good fit although the nugget effect is mostly too high, as previously noted.

Table 3. Fitted sills of the linear model of co-regionalization relative to the different spatial structures. The symbol I stands for indicator whereas u stands for uniform transformed.

(1) Nugget Effect									
Variable	I ECe	u_Ca	u_CaCO ₃	u_K	u_Mg	u_Na	u_SAR	u_av_K	u_av_P
I ECe	0.0864								
u_Ca	0.0272	0.0485							
u_CaCO ₃	0.0056	0.0061	0.0485						
u_K	0.0174	0.0108	0.0141	0.0372					
u_Mg	0.0150	0.0065	0.0094	0.0126	0.0199				
u_Na	0.0131	0.0105	0.0085	0.0144	0.0078	0.0122			
u_SAR	0.0023	−0.0041	0.0061	0.0102	−0.0093	0.0084	0.0349		
u_av_K	0.0095	0.0033	0.0160	0.0316	0.0062	0.0096	0.0098	0.0494	
u_av_P	−0.0154	−0.0152	0.0041	−0.0004	−0.0038	−0.0053	0.0025	0.0050	0.0523
(2) Cubic Model (Range 203 m)									
Variable	I ECe	u_Ca	u_CaCO ₃	u_K	u_Mg	u_Na	u_SAR	u_av_K	u_av_P
I ECe	0.1693								
u_Ca	0.0697	0.0395							
u_CaCO ₃	0.0689	0.0274	0.0421						
u_K	0.0961	0.0459	0.0320	0.0619					
u_Mg	0.0822	0.0371	0.0350	0.0499	0.0523				
u_Na	0.1005	0.0495	0.0474	0.0592	0.0567	0.0727			
u_SAR	0.0931	0.0416	0.0476	0.0509	0.0506	0.0641	0.0605		
u_av_K	0.0764	0.0332	0.0187	0.0510	0.0407	0.0414	0.0350	0.0483	
u_av_P	0.0475	0.0225	0.0153	0.0328	0.0354	0.0306	0.0272	0.0326	0.0354

By focusing in particular on the cross-semivariograms between I_EC and the uniform transformed variables, and assuming the distance of the cross-semivariogram model from the dashed line (intrinsic correlation) as a measure of the strength of the spatial correlation between the two variables [37] (Figure 2), the following observations can be highlighted. The variables that are most strongly spatially correlated to the indicator variable are the uniform transformed variables of Na, Mg and K, which represent the cations adsorbed into the solid matrix of soil. Submitting the soil to fertilization trials, without monitoring soil salinity and eventually adding leaching requirements might raise the salinity levels in the soil, especially if the soil is left as bare soil for long time, which is actually the case in the north-east area of the field. The source of salinity is the consecutive application of fertilizers, which are dissolved by the irrigation water and then penetrate along the profile. Although soil salinization is not generally expected with irrigated crops, in this case, the high electrical conductivity values might be due to a lack of compensation between upward capillary movement, due to high evaporation, and downward movement of water and solutes, due to the poor soil infiltration.

Differently, the uniform transformed P appears to be the least correlated variable to the indicator variable probably because of its lower mobility and precipitation as di- and/or tri-calcium phosphate and thus it does not affect soil salinity. Clearly, mapping the study variables can help considerably in locating areas at risk and interpreting the probable causes of soil salinization. This is why we have moved from a traditional statistical analysis to a geostatistical analysis. In order to emphasize the differences between the various structures of spatial dependence within a thematic map, we opted for a representation in isoquantiles rather than in a linear scale of colors, so that all color classes, having an equal number of pixels, were equally represented. Figure 3 shows the cokriged map of the probability of EC exceeding the critical threshold.

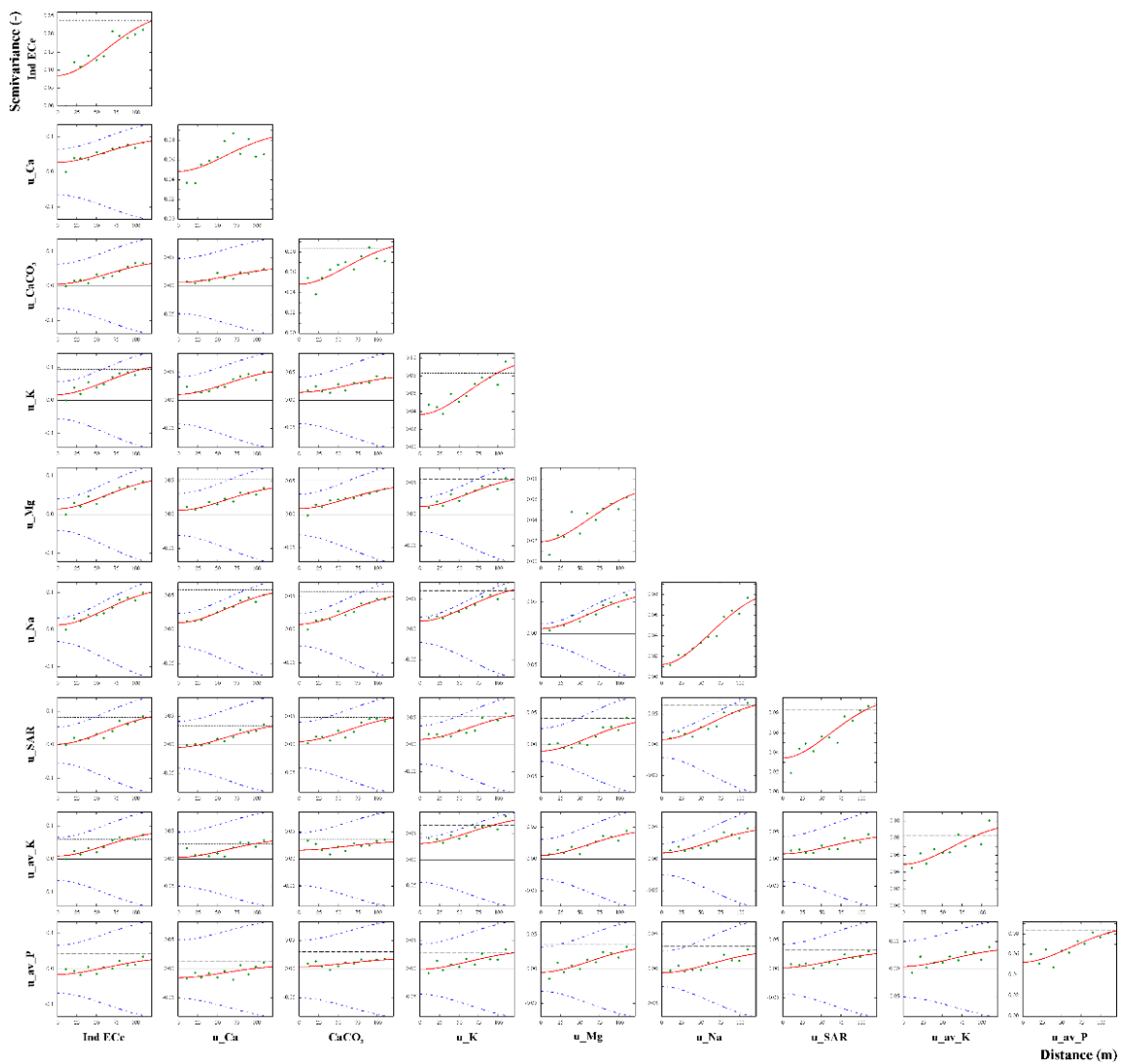


Figure 2. Direct and cross-semivariograms between the indicator transformation of ECe (I_ECe) and the uniform transformed variables. The experimental values are the plotted green points and the red solid lines represent the model of co-regionalization. The blue dash-dotted lines are the hull of perfect correlation [37] and the black dashed lines are the experimental variances.

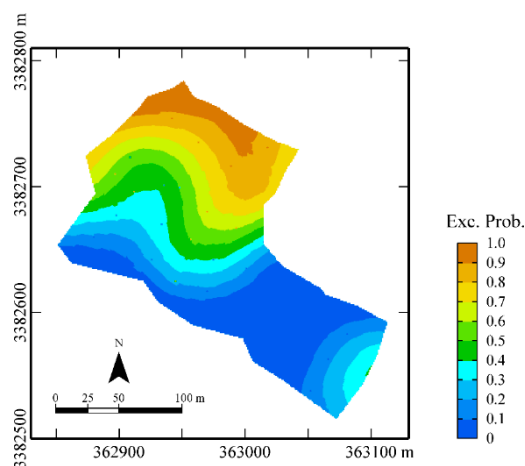


Figure 3. Cokriged map of the probability of ECe exceeding the critical threshold.

It is evident that there is a well-defined area to the north-east, where the probability of salinization is very high, that is, greater than 0.8. Moving from this area to the south and south-east, the probability decreases so we can consider this area is quite safe from salinization. Such a map leads us to search for local causes for the high risk of salinization of the soil in the proximity of the north-east. The territory in the north-east area was previously abandoned for many years without cultivation and subsequent irrigation. The higher salinity levels might be caused by the accumulation of salts on the surface due to intense evaporation, which has not been balanced by the infiltration water due to the low rainfall. Moreover, the absence of plants reduces the removal of cations from the soil, which are absorbed as minor nutrients.

In addition, this area is close to buildings, which hamper the movement of water and cause it to accumulate in the area. Further, the area is at a low level compared to the whole field, which favors water accumulation, and thus the deposit of salts on the surface due to the effect of high evaporation. The salt accumulation is also enhanced by the high clay content in the soil. In order to improve the interpretation of the risk map, the maps for the auxiliary variables (the values varying between 0 and 1) selected in this study, are also displayed (Figure 4).

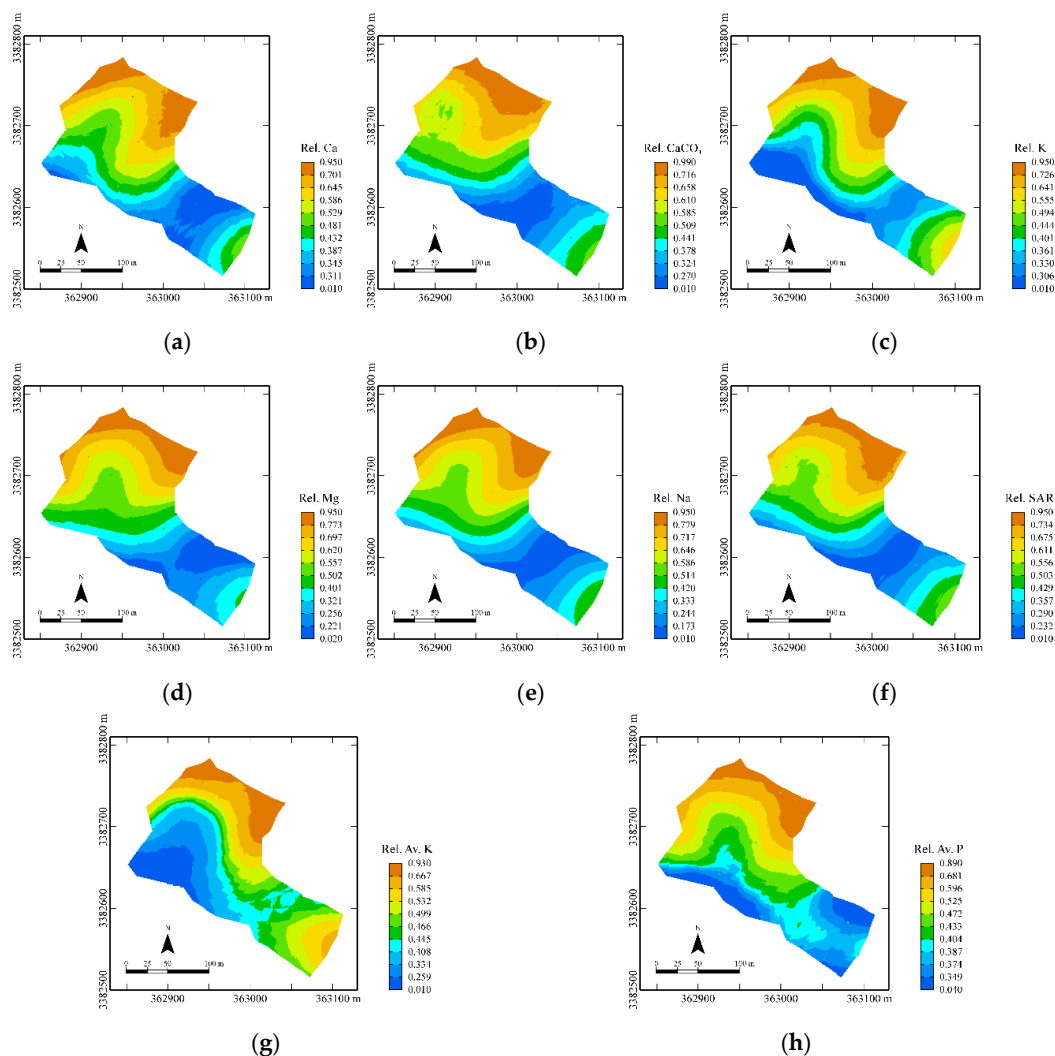


Figure 4. Maps for the auxiliary variables (rank-order transformed). (a) Ca, (b) CaCO₃, (c) K, (d) Mg, (e) Na, (f) SAR, (g) Av. K and (h) Av. P.

An overview of these maps suggests that they are strongly consistent since they all show the highest values in the north-east area. This is an indication that the study variables are different

representations of the same spatial processes that cause soil salinization in this part of the field. Thus, the maps support our hypothesis that the salinization process is essentially attributable to the fact that this area was uncultivated for many years owing to its nearness to buildings, which inhibit the free circulation of water and solutes in the soil/subsoil.

Finally, these maps can be helpful in the practice of precision agriculture, or more generally, in site-specific land management [27]. They allow delineation of the area at highest risk once a probability threshold has been selected, for example 0.8. Different levels of probability can also be set, with different levels of intervention priority assigned to them. In this way, various remediation operations will be performed only on this/these area(s) and not uniformly over the entire field. This will lead to considerable economic savings (less use of chemicals, energy and man/machine working hours), and the environmental impact will also be significantly reduced. Above all, our analysis clearly stresses the need to carry out a local analysis, since the causes of pollution can differ significantly in space and time, even in a small agricultural field.

The above discussion of the results has highlighted the advantages of using a multivariate approach, which assesses and models the spatial relationships between the various soil attributes through LMC in order to provide useful insights on the possible causes of the ongoing salinization process. However, in comparing the two procedures, applying cross-validation made the differences between the univariate (IK) and the multivariate (PK) non-parametric approaches more objective (Table 4).

Table 4. Results of cross-validation test for the procedures indicator kriging (Err IK) and probability kriging (Err PK). RMSSE is the root mean squared standardized error; r is the correlation coefficient between observation and estimation; ρ is the correlation coefficient between standardized error and estimation; n is the number of outliers.

Variable	Min	Max	Mean	St. Dev.	RMSSE	r	ρ	n
Err IK	−1.00	0.87	0.005	0.33	1.01	0.73	0.002	3
Err PK	−0.54	0.68	0.007	0.26	1.02	0.84	0.004	1

The results of cross-validation show that both procedures perform quite well (Table 4), so the probability estimates of exceeding the critical threshold for E_{Ce} can be regarded as unbiased and accurate for both methods. However, compared to IK, PK has less variability of errors, smaller minimum and maximum absolute errors, higher correlation between observations and estimates, and fewer outliers.

4. Conclusions

A probabilistic multivariate geostatistical approach was proposed to assess salinization risk. Maps were produced for both the individual soil indicator and the soil attributes expressed in relative terms. Thus, the procedure allowed for the identification of both the areas of high salinization risk, and the variables that mainly affect the process in those zones.

The method could be extended to any type of soil contaminant and auxiliary information, even with different spatial support after data homogenization, using the geostatistical techniques of data fusion, which explicitly take change of support into account [41]. The approach also proved to be quite promising from the perspective of precision agriculture, because it can effectively delineate the field into zones with different levels of priority and management for soil recovery. Moreover, the profusion of big data from current technology makes the approach easily extendable to any sort of remote and proximal sensing auxiliary information for assessment and modelling of soil contamination in view of precision remediation of potentially contaminated sites. Actually, risk analysis can benefit greatly from the use of remote and proximal sensing because the outcomes of the sensors can be used, according to the proposed approach, in conjunction with sparse ground measurements of the variables more directly related to the risk being assessed. In particular, the use of geophysical sensors such as ground

penetrating radar exploring the deeper depths of subsoil, in conjunction with analytical measurements of shallow soil, would allow for a better interpretation of the possible causes of soil salinization and the separation of natural from anthropogenic sources of variation, which provides the basis for efficient and effective remediation action.

Author Contributions: S.M.S. conceived and designed the experiments, analyzed and interpreted the data, and wrote the paper; A.C. conceived and designed the experiments, analyzed and interpreted the data, and wrote the paper; G.B. conceived and designed the experiments, analyzed and interpreted the data, designed the figures, and wrote the paper. All authors revised and approved the final paper version to be published. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors are very grateful to the National Research Council of Italy (Consiglio Nazionale delle Ricerche, CNR) for funding the period spent by the first author in Italy within the framework of the Short-Term Mobility Program 2016 (STM 2016). The authors thank the Reviewers of this paper for providing constructive comments, which have contributed to the improvement of the published version.

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

References

1. Corwin, D.L.; Rhoades, J.D.; Šimůnek, J. Leaching requirement for soil salinity control: Steady-state versus transient models. *Agric. Water Manag.* **2007**, *90*, 165–180. [CrossRef]
2. Hillel, D. *Environmental Soil Physics*; Academic Press: San Diego, CA, USA, 1998; p. 771.
3. Rhoades, J.D.; Loveday, J. *Salinity in Irrigated Agriculture*; Stewart, B.A., Nielsen, D.R., Eds.; American Society of Agronomists: Madison, WI, USA, 1990; ISBN 0891181024.
4. Richards, L.A. Diagnosis and Improvement of Saline and Alkaline Soils. *Soil Sci. Soc. Am. J.* **1954**, *18*, 348. [CrossRef]
5. Abrol, I.P.; Yadav, J.S.P.; Massoud, F.I. *Salt-Affected Soils and Their Management*; Food and Agriculture Organization of the United Nations: Rome, Italy, 1988; ISBN 9251026866.
6. Mohamed, E.; Belal, A.-A.; Ali, R.R.; Saleh, A.; Hendawy, E.A. Land Degradation. In *The Soils of Egypt*; Springer: Cham, Switzerland, 2019; pp. 159–174. ISBN 978-3-319-95516-2.
7. Ministry of Agriculture and Land Reclamation. Sustainable Agricultural Development Strategy towards 2030. Cairo, Egypt. Available online: <http://far-malr.gov.eg/pdf/en/Full%20SADS2030.pdf> (accessed on 19 November 2019).
8. Corwin, D.L.; Scudiero, E. Field-Scale Apparent Soil Electrical Conductivity. In *Methods of Soil Analysis*; Logsdon, S., Ed.; Soil Science Society of America: Madison, WI, USA, 2016; Volume 1, pp. 1–29.
9. Corwin, D.L. Site-specific management and delineating management zones. In *Precision Agriculture for Sustainability and Environmental Protection*; Oliver, M., Bishop, T., Marchant, B.P., Eds.; Routledge: London, UK, 2013; Volume 9780203128, pp. 135–157. ISBN 9780203128329.
10. Zovko, M.; Romić, D.; Colombo, C.; Di Iorio, E.; Romić, M.; Buttafuoco, G.; Castrignanò, A. A geostatistical Vis-NIR spectroscopy index to assess the incipient soil salinization in the Neretva River valley, Croatia. *Geoderma* **2018**, *332*, 60–72. [CrossRef]
11. Langella, G.; Agrillo, A.; Basile, A.; De Mascellis, R.; Manna, P.; Moretti, P.; Mileti, F.A.; Terribile, F.; Vingiani, S. Geography of soil contamination for characterization and precision remediation of potentially contaminated sites. *Ital. J. Agron.* **2018**, *13*, 6–15.
12. Bourgault, G.; Journel, A.G.; Rhoades, J.D.; Corwin, D.L.; Lesch, S.M. Geostatistical Analysis of a Soil Salinity Data Set. *Adv. Agron.* **1996**, *58*, 241–292. [CrossRef]
13. Castrignanò, A.; Buttafuoco, G.; Canu, A.; Zucca, C.; Madrau, S. Modelling spatial uncertainty of soil erodibility factor using joint stochastic simulation. *L. Degrad. Dev.* **2008**, *19*, 193–213. [CrossRef]
14. Dowd, P.A.; Pardo-Igúzquiza, E. The incorporation of model uncertainty in geostatistical simulation. *Geogr. Environ. Model.* **2002**, *6*, 147–169. [CrossRef]
15. Matheron, G. *The Theory of Regionalized Variables and Its Applications*; Ecole Nationale Supérieure des Mines de Paris: Fontainebleau, France, 1971; Volume 5.
16. Chilès, J.-P.; Delfiner, P. *Geostatistics: Modeling Spatial Uncertainty*, 2nd ed.; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2012; ISBN 9781118136188.

17. Webster, R.; Oliver, M.A. *Geostatistics for Environmental Scientists*; Statistics in Practice; John Wiley & Sons, Ltd: Chichester, UK, 2007; ISBN 9780470517277.
18. Journel, A.G. *Fundamentals of Geostatistics in Five Lessons*; American Geophysical Union: Washington, DC, USA, 1989; ISBN 0-87590-708-3.
19. Goovaerts, P. *Geostatistics for Natural Resources Evaluation*; Oxford University Press: New York, NY, USA, 1997; ISBN 0195115384.
20. Journel, A.G. Nonparametric Geostatistics for Risk and Additional Sampling Assessment. In *Principles of Environmental Sampling*; Keith, L.H., Ed.; American Chemical Society: Washington, DC, USA, 1988; pp. 45–72.
21. Juang, K.W.; Lee, D.Y. Comparison of three nonparametric kriging methods for delineating heavy-metal contaminated soils. In *Journal of Environmental Quality*; American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America: Madison, WI, USA, 2000; Volume 29, pp. 197–205.
22. Goovaerts, P.; Avruskin, G.; Meliker, J.; Slotnick, M.; Jacquez, G.; Nriagu, J. Geostatistical modeling of the spatial variability of arsenic in groundwater of southeast Michigan. *Water Resour. Res.* **2005**, *41*, 1–19. [[CrossRef](#)]
23. Lee, J.-J.; Jang, C.-S.; Wang, S.-W.; Liu, C.-W. Evaluation of potential health risk of arsenic-affected groundwater using indicator kriging and dose response model. *Sci. Total Environ.* **2007**, *384*, 151–162. [[CrossRef](#)] [[PubMed](#)]
24. Castrignanò, A.; Buttafuoco, G.; Puddu, R. Multi-scale assessment of the risk of soil salinization in an area of south-eastern Sardinia (Italy). *Precis. Agric.* **2008**, *9*, 17–31. [[CrossRef](#)]
25. Carr, J.R. Order relation correction experiments for probability kriging. *Math. Geol.* **1994**, *26*, 605–621. [[CrossRef](#)]
26. Adhikary, P.P.; Dash, C.J.; Bej, R.; Chandrasekharan, H. Indicator and probability kriging methods for delineating Cu, Fe, and Mn contamination in groundwater of Najafgarh Block, Delhi, India. *Environ. Monit. Assess.* **2011**, *176*, 663–676. [[CrossRef](#)] [[PubMed](#)]
27. Shaddad, S.M.; Buttafuoco, G.; Elrys, A.; Castrignanò, A. Site-specific management of salt affected soils: A case study from Egypt. *Sci. Total Environ.* **2019**, *688*, 153–161. [[CrossRef](#)] [[PubMed](#)]
28. Bakr, N.; Bahnassy, M.H. Egyptian Natural Resources. In *The Soils of Egypt*; El-Ramady, H., Alshaal, T., Bakr, N., Elbana, T., Mohamed, E., Belal, A.-A., Eds.; Springer: Cham, Switzerland, 2019; pp. 33–49. ISBN 978-3-319-95516-2.
29. Soil Science Division Staff. *Soil Survey Manual. USDA Handbook 18*; Ditzler, C., Scheffe, K., Monger, H.C., Eds.; Government Printing Office: Washington, DC, USA, 2017.
30. El Bouraie, M.M.; Motawea, E.A.; Mohamed, G.G.; Yehia, M.M. Water quality of Rosetta branch in Nile delta, Egypt. *Suo* **2011**, *62*, 31–37.
31. El-Marsafawy, S.; Bakr, N.; Elbana, T.; El-Ramady, H. *Climate*; El-Ramady, H., Alshaal, T., Bakr, N., Elbana, T., Mohamed, E., Belal, A.-A., Eds.; Springer: Cham, Switzerland, 2019; pp. 69–92.
32. Ayers, R.S.; Wescot, D.W. *Water Quality for Agriculture. FAO Irrigation and Drainage Paper 29, Rev. 1, Reprinted 1994*; Ayers, R.S., Westcot, D.W., Eds.; Food and Agriculture Organization of the United Nations: Rome, Italy, 1994; ISBN 9251022631.
33. Miller, J.J.; Curtin, D. Electrical Conductivity and Soluble Ions. In *Soil Sampling and Methods of Analysis*; Carter, M.R., Gregorich, E.G., Eds.; CRC Press: Boca Raton, FL, USA, 2007; p. 1224. ISBN 9780849335860.
34. Slavich, P.G.; Petterson, G.H. Estimating the electrical conductivity of saturated paste extracts from 1:5 soil: Water suspensions and texture. *Aust. J. Soil Res.* **1993**, *31*, 73. [[CrossRef](#)]
35. Sparks, D.L.; Page, A.L.; Helmke, P.A.; Loeppert, R.H. *Methods of Soil Analysis Part 3—Chemical Methods. SSSA Book Ser. 5.3. SSSA*; SSSA Book Series; Soil Science Society of America, American Society of Agronomy: Madison, WI, USA, 1996; ISBN 978-0-89118-866-7.
36. Castrignanò, A.; Giugliarini, L.; Risaliti, R.; Martinelli, N. Study of spatial relationships among some soil physico-chemical properties of a field in central Italy using multivariate geostatistics. *Geoderma* **2000**, *97*, 39–60. [[CrossRef](#)]
37. Wackernagel, H. *Multivariate Geostatistics: An Introduction with Applications*; Springer: Berlin/Heidelberg, Germany, 2003; ISBN 3540441425.
38. Isaaks, E.H.; Srivastava, R.M. *Applied Geostatistics*; Oxford University Press: Oxford, UK, 1989; Volume 21, ISBN 0195050134.
39. Cressie. *Statistics for Spatial Data-Cressie.pdf*; Wiley-Blackwell: Hoboken, NJ, USA, 1993; ISBN 0-471-84336-9.

40. Castrignanò, A.; Wong, M.T.F.; Stelluti, M.; De Benedetto, D.; Sollitto, D. Use of EMI, gamma-ray emission and GPS height as multi-sensor data for soil characterisation. *Geoderma* **2012**, *175–176*, 78–89. [[CrossRef](#)]
41. Castrignanò, A.; Quarto, R.; Venezia, A.; Buttafuoco, G. A comparison between mixed support kriging and block cokriging for modelling and combining spatial data with different support. *Precis. Agric.* **2019**, 1–21. [[CrossRef](#)]



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