

# Spatial data fusion and analysis for soil characterization: a case study in a coastal basin of south-western Sicily (southern Italy)

Donato Sollitto,<sup>1</sup> Daniela De Benedetto,<sup>1</sup> Annamaria Castrignanò,<sup>1</sup> Giuseppina Crescimanno,<sup>2</sup> Giuseppe Provenzano,<sup>2</sup> Domenico Ventrella<sup>1</sup>

<sup>1</sup>Consiglio per la Ricerca e la sperimentazione in Agricoltura - Unità di ricerca per i Sistemi Colturali degli Ambienti caldo-aridi (CRA-SCA), Bari; <sup>2</sup>Dipartimento dei Sistemi Agro-Ambientali, Università di Palermo, Italy

## Abstract

Salinization is one of the most serious problems confronting sustainable agriculture in semi-arid and arid regions. Accurate mapping of soil salinization and the associated risk represent a fundamental step in planning agricultural and remediation activities. Geostatistical analysis is very useful for soil quality assessment because it makes it possible to determine the spatial relationships between selected variables and to produce synthetic maps of spatial variation. The main objective of this paper was to map the soil salinization risk in the Delia-Nivolelli alluvial basin (south-western Sicily, southern Italy), using multivariate geostatistical techniques and a set of topographical, physical and soil hydraulic properties. Elevation data were collected from existing topographic maps and analysed preliminarily to improve the estimate precision of sparsely sampled primary variables. For interpolation multi-collocated cokriging was applied to the dataset, including textural and hydraulic properties and electrical conductivity measurements carried out on 128 collected soil samples, using elevation data as auxiliary variable. Spatial dependence among elevation

and physical soil properties was explored with factorial kriging analysis (FKA) that could isolate and display the sources of variation acting at different spatial scales. FKA isolated significant regionalised factors which give a concise description of the complex soil physical variability at the different selected spatial scales. These factors mapped, allowed the delineation of zones at different salinisation risk to be managed separately to control and prevent salinization risk. The proposed methodology could be a valid support for land use and soil remediation planning at regional scale.

## Introduction

High salt content in topsoil represents a serious threat for agricultural productions because it may limit the plant growth and cause a continuous degradation of soil quality and fertility. There are many possible, both natural and anthropogenic causes, of soil salinization. This is generally related to arid and semiarid climatic conditions and actually is one of the main causes of desertification in many places of the world (Tóth *et al.*, 2008). In Italy desertification risk is more widely recorded in the southern regions, such as Apulia, Calabria, Basilicata and Sicily, where leaching of salt is reduced by low rainfall (Ministry for the Environment, Land and Sea, 2007).

In these regions accurate assessment of soil salinization and its variability represents an important issue which should be taken into account in planning agricultural activities to increase crop yield. Several variables were used by different authors to predict soil salinity, such as electrical conductivity of the saturated soil paste extract EC<sub>e</sub>, exchangeable sodium percentage, pH, and total clay+fine silt content (Rhoades *et al.* 1999; Pozdnyakova and Zhang 1999; Corwin and Lesch 2003; Douaik *et al.*, 2004). This prediction is quite important because excessive amounts of salts may adversely influence biological, physical and chemical properties of soil (Lauchli and Epstein, 1990).

Advanced geostatistical analyses are often used in this research field to characterize soil salinization and determine its causes adequately. Geostatistics offers a collection of statistical and probabilistic tools aimed at understanding and modeling spatial variability and allows to utilize secondary information, that is often available at finer spatial resolution than the sampled values of a primary target variable to increase spatial resolution and improve estimation precision. The estimation of soil salinization risk actually requires to model all the environmental components controlling the process and to analyze their variability at different spatial scales (Castrignanò *et al.*, 2009). At least two processes can cause salinization: the first, operating at large scale, naturally occurs where high levels of soluble salts are stored in the soil and

Correspondence: Annamaria Castrignanò, Consiglio per la Ricerca e la sperimentazione in Agricoltura - Unità di ricerca per i Sistemi Colturali degli Ambienti caldo-aridi (CRA-SCA), via Celso Ulpiani 5, 70125 Bari, Italy.  
E-mail: annamaria.castrignanò@entecra.it

Key words: salinization risk, soil retention curve, geostatistics, factor Kriging, intrinsic random functions.

Acknowledgements: this research was funded by CLIMESCO *Evolution of cropping systems as affected by climate change* project, contract n. 285, 20/02/2006 (Ministry for Education, University and Research).

Received for publication: 1 April 2011.

Accepted for publication: 6 December 2011.

©Copyright D. Sollitto *et al.*, 2012

Licensee PAGEPress srl, Italy

Italian Journal of Agronomy 2012; 7:e4

doi:10.4081/ija.2012.e4

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (by-nc 3.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

ground water. In such cases, the accumulations of salts have originated from landscape processes or sea water intrusion into fractured rock of the aquifer. The second process causing salinization can result from human activities, at short range, such as the excessive use of brackish water in agriculture (Castrignanò *et al.*, 2008). Therefore, since some factors affecting soil salinization are likely to have a short range of influence, whereas others act over longer ranges, soil variables are expected to be correlated in a scale-dependent way. The estimation of their scale-dependent correlation structure then becomes crucial in salinization risk investigations (Sollitto *et al.*, 2010).

The objective of this paper was to delineate areas with different degree of salinization risk in the Delia-Nivolelli sedimentary basin (south-western Sicily, southern Italy) through a multivariate geostatistical approach.

## Material and methods

### The study area

The Delia-Nivolelli sedimentary basin covers an area of about 6000 ha in Trapani province, between Mazara del Vallo and Castelvetrano (south-eastern Sicily, Italy). The basin is characterized by an alluvial plain bordered by hilly landscapes shaped on clayey, sandy and evaporitic rocks (Figure 1). The most widely outcropping lithology is represented by Holocene alluvial deposits of Delia river. In the northern and south-eastern boundaries of the alluvial plain, marly and sandy clay outcrops on a hilly belt, referred to *Cozzo Terravecchia* formation (Miocene). Outside the basin this continental sedimentary formation is covered by a considerable thickness of evaporitic rocks of the *Chalky Solphurous sequence* formation. Pliocene terraced marine sandstones extend in the western side of the Delia floodplain and on the surrounding flat hills. The main morphological elements of the area are the Delia river and the other tributary streams coming from the hilly belts. According to the USDA soil classification (USDA, 1999), only four subgroups of soil can be found in the Delia-Nivolelli catchment: Lithic Xerorthens, Typic Chromoxerert, Vertic Xerichrept and Vertic Xerofluvent.

Agricultural land use is widespread in the study area and grapevine is the main crop covering about 83% of the whole agricultural area, which represents 93% of the whole study area.

At present, one of the most important problems in the study area is the considerable increase of soil salinization due to high salinity of irrigation water. Both irrigation water from groundwater and surface water coming from the dam are saline water. The irrigation in farm is generally practised with sprinkler or drip systems using the water stored in the dam built upstream of the Delia river, in which the salts coming from the weathering of evaporitic rocks are dissolved. For this reason one of the most important problems in the area is the considerable increase in soil salinization caused by high salinity of the irrigation water.

### Soil sampling

In order to assess soil salinization, a soil survey was carried out within the frame of a European project aimed at preventing salinization and desertification (Crescimanno, 1998). Topsoil samples were randomly collected at 128 locations and soil moisture ( $\text{m}^3 \text{m}^{-3}$ ), electrical conductivity of saturated extract ( $\text{dS m}^{-1}$ ) and grain size distribution ( $\text{kg kg}^{-1}$ ) were measured. The laboratory measurements were carried out according to the methods of soil analysis of American Societies of Agronomy (ASA) and Soil Science (SSSA) (ASA-SSSA, 1986).

To determine the soil retention curves, hanging water column apparatus (Burke *et al.*, 1996) was used to evaluate soil water content,  $\theta$ , corresponding to soil water potential ( $h$ ) values ranging from -0.05 to -

1.0 m on undisturbed soil samples of 0.08-m diameter and 0.05-m height; pressure plate apparatus (Dane and Hopmans, 2002) was used to measure soil water content corresponding to the  $h$  values of -3.37 m, -10.2 m, -30.6 m, and -153.0 m, on sieved soil samples of 0.05-m diameter and 0.01-m height.

Saturated hydraulic conductivity,  $K_s$ , was determined by the constant head permeameter method (Klute, 1986) on the same undisturbed soil samples used to determine the retention curves.

A digital model of the soil at high resolution was produced by the digitization of existing points of known elevation from regional technical maps. Firstly, 2364 elevation were drawn from a technical map on a 1:5000 scale; secondly, to obtain a more even distribution of points, further 1227 points were chosen on a map on a 1:10000 scale. The exhaustive elevation dataset included 3591 points and was split into a calculation set and a validation set: the first one (2635 points) was used for interpolation, whereas the second (668 points) for validation (Figure 2).

### Multivariate geostatistical analysis

#### Intrinsic random functions of order k (IRF-k)

The elevation data showed a spatial trend (unbounded variogram), so that they were interpolated by using a non stationary linear geostatistical technique, called *intrinsic random functions of order k* (IRF-k) theory formulated by Matheron (1973). It is based on the decomposition of

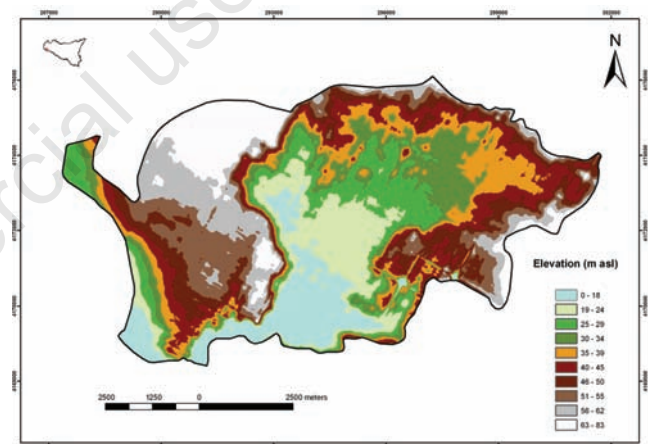


Figure 1. Digital elevation model of the Delia-Nivolelli sedimentary basin (or study area).

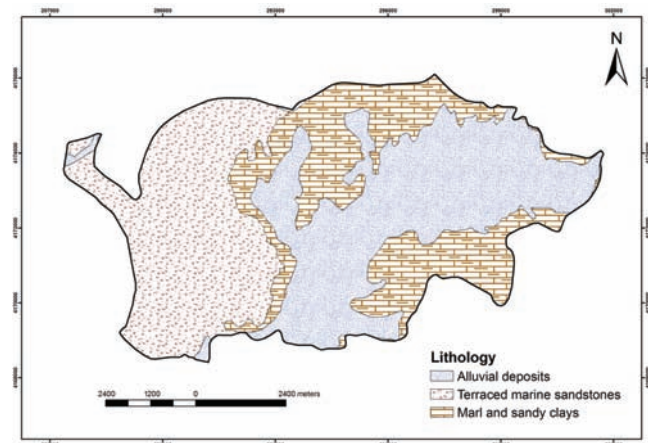


Figure 2. Geolithological map of the study area.

the total spatial variation into a drift component and a stochastic component through the calculation of increments of a sufficiently high order ( $k$ ), so that the drift can be filtered out and stationarity attained (Chilès and Delfiner, 1999; Buttafuoco and Castrignanò, 2005).

The value of some non-stationarity soil attribute  $Z$  at any location  $x$ , expressed as a vector of the coordinates in two dimensional space, can be written as the sum of two components:

$$Z(x) = m(x) + Y(x) \quad (1)$$

where  $m(x)$  is a slowly varying deterministic function, known as drift, and  $Y(x)$  is a rapidly fluctuating spatially correlated random component with zero mean. The basic hypothesis of kriging with drift is that the expectation of the variable, known only at a small set of points in the study area, can be written as the sum of a basis of  $L$  polynomials of the coordinates ( $f_l(x)$ ) (internal drift):

$$E[Z(x)] = \sum_{l=0}^L a_l f_l(x) \quad (2)$$

where  $a_l$  are unknown coefficients.

The correlation structure associated with the random part ( $Y(x)$ ) is expressed by a generalized covariance (GC) function of the distance between two observations  $x_\alpha$  and  $x_\beta$ ,  $K(x_\alpha - x_\beta)$ , denoted by  $K(h)$ . A convenient model for the generalized covariance is the polynomial GC model, namely a linear combination of a given set of generic basic structures under conditions on the coefficients  $b$ :

$$K(|h|) = C_0 \delta(|h|) - b_0 |h| + b_s |h|^2 \log |h| + b_1 |h|^3 \quad (3)$$

where  $\delta(|h|) = 0$  for  $|h| > 0$  else  $\delta(|h|) = 1$ . The coefficients  $C_0$ ,  $b_0$ ,  $b_s$  and  $b_1$ , in a two dimensional space  $R^2$ , must satisfy a set of inequalities for  $K(h)$  being a valid generalized covariance of an IRF- $k$ :

$$C_0 \geq 0, b_0 \geq 0, b_1 \geq 0, b_s \geq -\frac{3}{2} \sqrt{b_0 b_1} \quad (\text{Chilès and Delfiner, 1999}).$$

In intrinsic random function kriging, the structural analysis is done in two stages: firstly the order  $k$  of the drift is established and secondly the generalized covariance,  $K(h)$ , is estimated by fitting a parametric model. To determine the degree of drift, the least-squares errors are ranked in ascending magnitude for each target point and for each option (order) of the drift. The order that produces the smallest rank, averaged over the different target points, corresponds to the optimal model for the drift.

To ensure mean and variance of kriging error exist, the following condition must be satisfied:

$$\sum_{\beta=1}^N \lambda_\beta f_l(x_\beta) = f_l(x_0) \quad l = 0, \dots, L \quad (4)$$

where  $x_0$  is the estimated point and  $\lambda_\beta$  are the kriging weights.

The constraint in estimate optimality leads to the traditional kriging equations with the difference of using a generalised covariance function instead of a traditional covariance or variogram function.

#### The factorial CoKriging analysis

The soil dataset was submitted to the multivariate geostatistical technique called *factorial CoKriging analysis* (FCKA) and developed by Matheron (1982) and exhaustively described in several papers (Castrignanò *et al.*, 2000; Wackernagel, 2003).

FCKA decomposes the set of original second-order random variables

$\{Z_i(x); i = 1, \dots, n\}$  into a set of reciprocally orthogonal factors  $\{Y_v^u(x); v = 1, \dots, n; u = 1, \dots, N_s\}$ , defined for each of  $N_s$  spatial scales

(u) through the transformation coefficients  $a_{iv}^u$ :

$$Z_i(x) = \sum_{u=1}^{N_s} \sum_{v=1}^n a_{iv}^u Y_v^u(x)$$

The regionalized factors  $Y_v^u$  are estimated through a modified cokriging system (Wackernagel, 2003) and then mapped, providing a display of the behaviour and relationships among the variables at the different spatial scales.

The applied procedure consists in three steps: i) modelling the co-regionalization of a set of  $n$  variables, using the *Linear Model of Co-regionalization* (LMC) and cokriging the individual variables; ii) analysing the multivariate correlation structure by applying *Principal Component Analysis* (PCA); iii) co-kriging and mapping specific factors at each characteristic spatial scales.

*Linear Model of Co-regionalization*, developed by Journel and Huijbregts (1978), considers all the studied variables as the result of the same independent physical processes, acting at different spatial scales  $u$ . The  $n(n+1)/2$  simple and cross semivariograms of the  $n$  variables are modelled by a linear combination of  $N_s$  semivariograms standardized to unit sill,  $g^u(h)$ , each one specific of the spatial scale  $u$ . Using the matrix notation, LMC can be written as:

$$\Gamma(h) = \sum_{u=1}^{N_s} B^u g^u(h)$$

where  $\Gamma(h) = [\gamma_{ij}(h)]$  is a symmetric matrix of order  $n \times n$ , whose diagonal and out-of-diagonal elements represent simple and cross semivariograms, respectively, for the lag  $h$ ;  $B^u = [b_{ij}^u]$  (co-regionalization matrix) is a symmetric positive semi-definite matrix of order  $n \times n$ , with real elements  $b_{ij}^u$  at the specific spatial scale  $u$ . The functions  $g^u(h)$  are authorized semivariograms models and fitting LMC is performed by weighted least-squares approximation under the constraint of positive semi-definiteness of the  $B^u$ , using the iterative procedures developed by Lajaunie and Béhatéguy (1989).

To perform multivariate analysis on soil chemical properties and elevation, since the soil sampling locations did not match the elevation locations, the elevation was estimated at soil samples locations using IRF- $k$  kriging. A coregionalized dataset was then produced involving the elevation estimates and soil data which were submitted to multivariate analysis. Multi-collocated cokriging is a way of integrating secondary (auxiliary) finer information in primary (target) variable modelling, where the contribution of secondary variable to estimation relies on the cross-correlation with the primary variable.

This approach is quite similar to ordinary cokriging (Wackernagel, 2003) with the only difference in the neighbourhood search. Since using all secondary information contained within the neighbourhood may lead to an intractable solution due to too much information, the secondary variable is used only at the target location and also at all the locations where the primary variable is defined within the neighbourhood. This solution has generally produced reliable and stable results (Rivoirard, 2001; Castrignanò *et al.*, 2009). In contrast to other kriging techniques, such as regression kriging and kriging with external drift, in this approach the influence of secondary variable on primary variable is explicitly taken into account through the estimation of the direct secondary variable variogram and the cross-variogram. Since the co-located secondary datum tends to screen the influence of more distant secondary data, there is actually little loss of information.

In this study case the topographic elevation grid, previously estimated, was used as secondary information for the multi-collocated cokriging of the soil variables (primary variable). *Regionalized Principal*



Component Analysis decomposes each co-regionalization matrix  $B^u$  into eigenvalues and eigenvector matrices (Wackernagel, 2003):

$$B^u = Q^u \Lambda^u Q^{uT} = (Q^u \sqrt{\Lambda^u})(Q^u \sqrt{\Lambda^u})^T = A^u A^{uT}$$

where  $Q^u$  is the matrix of eigenvectors, i.e. the regionalised factors  $Y_v^u(x)$ , and  $\Lambda^u$  is the diagonal matrix of eigenvalues for each spatial scale  $u$ ;  $A^u = Q^u \sqrt{\Lambda^u}$  is the matrix of order  $n \times n$  of the transformation coefficients  $a_{iv}^u$ . The transformation  $a_{iv}^u$  coefficients in the matrix  $A^u$  correspond to the covariances between the original variables  $Z_i(x)$  and the regionalized factors  $Y_v^u(x)$ .

#### The Gaussian anamorphosis

Variogram modelling is sensitive to the presence of outliers when data distribution is highly skewed. In such a case, it is better to normalize and standardise data to mean 0 and variance 1 through *Gaussian Anamorphosis Modelling*.

A procedure known as Gaussian anamorphosis (Chilès and Delfiner, 1999) allows transforming each variable  $Z$  into a Gaussian-shaped variable  $Y$  with zero mean and unit variance, through a mathematical function  $\Phi : Z = \Phi(Y)$ . To transform the raw variable with any distribution into a Gaussian one, it is necessary to invert the previous function as

$$Y = \Phi^{-1}(Z)$$

Such function is estimated by fitting a polynomial expansion (Chilès and Delfiner, 1999):

$$\Phi(Y) = \sum \Psi_i H_i(Y)$$

where  $H_i(Y)$  are the Hermite Polynomials and  $\Psi_i$  are coefficients. In practice, the polynomial expansion is restricted to a generally high order (30-100) and is monotonically increasing within the interval defined by the minimum and the maximum of the sample values (Wackernagel, 2003). The elevation data were interpolated with IRF-k kriging over a  $5 \times 5$  m mesh grid by using the generalized covariance. The estimated kriging values were back transformed to the raw values to produce the map of elevation. The soil data were then interpolated at the nodes of the same grid of DEM, by applying multi-collocated gaussian co-kriging method, and the estimates were then back-transformed to the raw values through the anamorphosis functions previously calculated.

## Results and discussion

### Digital elevation model

The elevation data range between the sea level (0 m) and 160 m asl, with a mean value of 49.08 m. They are quite asymmetrically distributed toward the higher elevations (Figures 3 and 4), as also revealed by the positive skeweness coefficient; therefore, they were transformed through Gaussian anamorphosis procedure.

The variogram map (*not shown*) revealed higher continuity along the N-40° direction, which corresponds to the main development direction of the Delia valley. After different trials, the trend component of variation for the topographic elevation data was modelled by a linear polynomial function of the geographical coordinates  $x$  and  $y$  and an intercept, whereas the stochastic component was estimated by a first order generalized covariance with a 1695.52 m range. The error statistics in the validation test report a quite low mean value of -0.12 m but the errors were widely varying between -47.61 m and 44.86 m, due to the uneven topography of the study area.

The digital elevation model of the Delia-Nivoletti basin (Figure 1) clearly shows the presence of a central sector characterized by a flat morphology with elevation values ranging from about 10 m to 30 m asl, decreasing from NE to SW, with an average topographic gradient of about 0.26%. The lowest elevation values are in the southern area, where they are very close to the sea level. Topographic elevation rapidly increases on the surrounding hilly boundaries up to 70-80 m asl, with quite steep slopes mainly in the northern and south-western sectors of the basin.

### Mapping soil properties

Basic statistics of the soil properties show highly skewed distributions for most of them, with both negative and positive skeweness coefficients (Table 1). Negative coefficients are for all the soil retention curve parameters, whereas the soil physical attributes are characterized by positive skeweness coefficients, with outliers at the higher values mainly for electrical conductivity, saturated hydraulic conductivity and gravel.

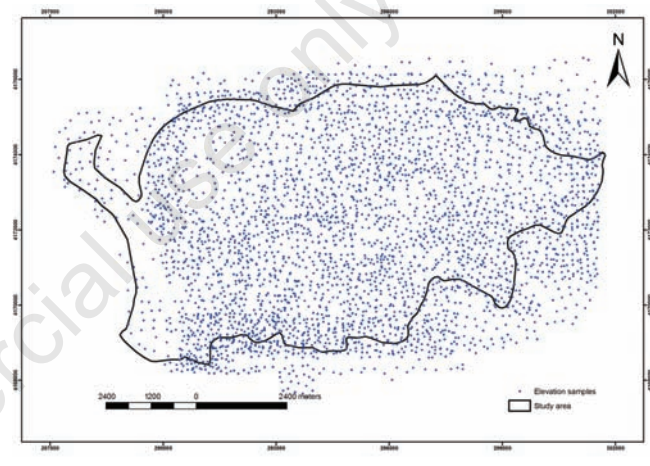


Figure 3. Topographic elevation samples locations.

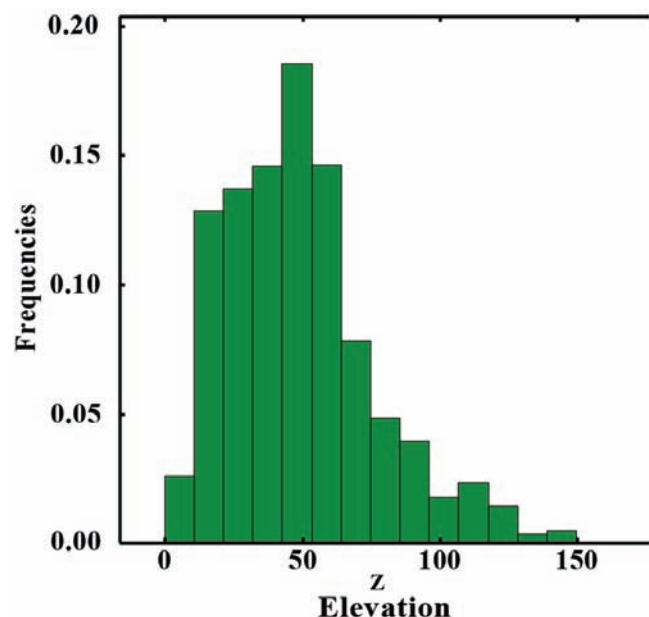


Figure 4. Frequency (-) distribution of elevation data.

The correlation matrix (Table 2) shows that the strongest correlations occur between the adjacent soil water values of the retention curve ( $\theta(h)$ ), whereas they tend to decrease as the distance increases. Clay and sand have positive and negative correlations, respectively, with the water values corresponding to different water potential. The skeleton is negatively correlated with  $\theta$  at the water potentials less than -3.37 m and with the clay content, but positively with the sand.

Elevation is not significantly correlated with the soil properties, whereas it is negatively correlated with the  $\theta$  at the potentials less than -0.10 m. The electrical conductivity is correlated positively with  $\theta$  at the lowest potential and negatively with sand and saturated hydraulic conductivity  $K_s$ . No relevant anisotropy was observed in the variogram

maps of the soil variables and all the variograms looked to be upper bounded (*not shown*), so that an intrinsic stationarity was assumed for all the soil variables. An isotropic linear model of coregionalization was then fitted to model the experimental variograms, including three basic structures, revealed by the visual inspection of the experimental variogram matrix: a nugget effect, a spherical model with a 1000-m range and a spherical model with a 4000-m range. The co-kriging maps of the variables show a clear dependence on the physical characters of the basin (Figure 5 a-r).

As regards the water values of the retention curve, it can be seen that the maps at the highest values of the matric potential look quite variable, without any clear structure (Figure 5 a-e). It can be observed the

**Table 1. Basic statistics of variables.**

Variable	Points	Min	Max	Mean	St. dev.	Variance	Var. coeff.	Skewness
$\theta$ h=0 m*	128	0.23	0.61	0.42	0.05	0.0034	0.14	-0.12
$\theta$ h=-0.05 m*	128	0.22	0.59	0.41	0.05	0.0033	0.14	-0.18
$\theta$ h=-0.10 m*	128	0.22	0.57	0.4	0.05	0.0029	0.13	-0.21
$\theta$ h=-0.40 m*	128	0.16	0.45	0.35	0.04	0.0023	0.13	-0.91
$\theta$ h=-0.70 m*	128	0.12	0.44	0.33	0.04	0.0024	0.15	-1.09
$\theta$ h=-1.20 m*	128	0.11	0.42	0.31	0.04	0.0024	0.16	-0.99
$\theta$ h=-3.37 m*	128	0.1	0.41	0.28	0.05	0.0031	0.19	-0.57
$\theta$ h=-10.2 m*	128	0.08	0.34	0.23	0.05	0.0027	0.22	-0.36
$\theta$ h=-30.6 m*	128	0.06	0.29	0.19	0.04	0.0021	0.23	-0.28
$\theta$ h=-153 m*	128	0.05	0.28	0.17	0.04	0.0019	0.25	-0.07
Clay (%)°	128	12	65	38.31	11.2	125.4	0.29	-0.14
Sand (%)#	128	0	76	36.11	14.4	207.3	0.4	0.41
Skeleton (%)§	128	0.07	37.9	8.15	10.09	101.8	1.24	1.41
$K_s$ (cm/h)^	128	0.28	117.8	23.47	22.76	518.2	0.97	1.32
EC1:5 (mS/cm)§	128	0.08	1.7	0.32	0.16	0.0286	0.53	4.39
pH	128	7.2	8.5	7.93	0.17	0.0307	0.02	-0.11
Elevation	3591	0	160	49.08	26.91	724.4	0.54	0.89

\*Soil water content,  $\theta$ , corresponding to h values ranging from 0 to -153 m; °clay content; #sand content; §skeleton content; ^saturated hydraulic conductivity  $K_s$ ; §electrical conductivity of saturated extract.

**Table 2. Correlation matrix of the studied variables. All correlation coefficients >0.23 are significant at P<0.01.**

Variable	$\theta$ h=0 m	$\theta$ h=-0.05 m	$\theta$ h=-0.10 m	$\theta$ h=-0.40 m	$\theta$ h=-0.70 m	$\theta$ h=-1.20 m	$\theta$ h=-3.37 m	$\theta$ h=-10.2 m	$\theta$ h=-30.6 m	$\theta$ h=-153 m	Clay	Sand	Skeleton	$K_{sat}$	EC <sub>1:5</sub>	pH	Elevation
$\theta$ h=0 m*	1																
$\theta$ h=-0.05 m*	0.98	1															
$\theta$ h=-0.10 m*	0.93	0.98	1														
$\theta$ h=-0.40 m*	0.72	0.77	0.83	1													
$\theta$ h=-0.70 m*	0.61	0.66	0.71	0.97	1												
$\theta$ h=-1.20 m*	0.52	0.55	0.6	0.91	0.97	1											
$\theta$ h=-3.37 m*	0.37	0.4	0.42	0.74	0.84	0.89	1										
$\theta$ h=-10.2 m*	0.41	0.44	0.46	0.75	0.82	0.87	0.97	1									
$\theta$ h=-30.6 m*	0.4	0.43	0.46	0.74	0.82	0.87	0.97	0.99	1								
$\theta$ h=-153 m*	0.38	0.39	0.41	0.66	0.73	0.78	0.91	0.93	0.93	1							
Clay°	0.15	0.15	0.17	0.52	0.65	0.73	0.85	0.86	0.85	0.81	1						
Sand#	-0.26	-0.28	-0.3	-0.64	-0.74	-0.8	-0.9	-0.91	-0.91	-0.87	-0.92	1					
Skeleton§	0.18	0.2	0.21	-0.06	-0.16	-0.21	-0.38	-0.36	-0.36	-0.37	-0.46	0.38	1				
$K_s$ ^	0.27	0.18	0.11	0.06	0.09	0.09	0.11	0.1	0.08	0.09	0.12	-0.08	-0.05	1			
EC1:5§	0.04	0.07	0.08	0.14	0.16	0.18	0.19	0.2	0.22	0.28	0.15	-0.26	-0.05	-0.33	1		
pH	0.01	-0.06	-0.1	0.03	0.1	0.14	0.18	0.19	0.19	0.22	0.34	-0.26	-0.21	0.2	0.04	1	
Elevation data	-0.07	-0.15	-0.22	-0.3	-0.29	-0.27	-0.23	-0.24	-0.23	-0.12	-0.12	0.26	-0.22	0.15	0.05	0.35	1

\*Soil water content,  $\theta$ , corresponding to h values ranging from 0 to -153 m; °clay content; #sand content; §skeleton content; ^saturated hydraulic conductivity  $K_s$ ; §electrical conductivity of saturated extract.

highest soil moisture values (greater than 0.5%) are tendentially distributed in the north-western portion of the basin, whereas the lowest values (0.30÷0.25%) are found in a large area located on the southern and south-western sectors.

As the matric potential decreases, the spatial structures of  $\theta$  look quite similar to the morphological and lithological settings of the basin. The  $\theta$  maps at  $h=-30.6$  m and  $h=-153$  m (Figure 5 i-l) show that the lowest values occur in the western and south-western sectors of the

area, roughly corresponding to the outcropping area of the marine terraced deposits composed by sands and silty-sands. On the contrary, the highest values are located in the eastern sector of the basin and seem to follow the main surface drainage patterns in the flattest portions of the area. The dependence of  $\theta$  at the lowest potentials on the soil texture and lithological composition of parent material is confirmed by the similarity of their maps with the ones of sand, clay and skeleton (Figure 5 o,p,q).

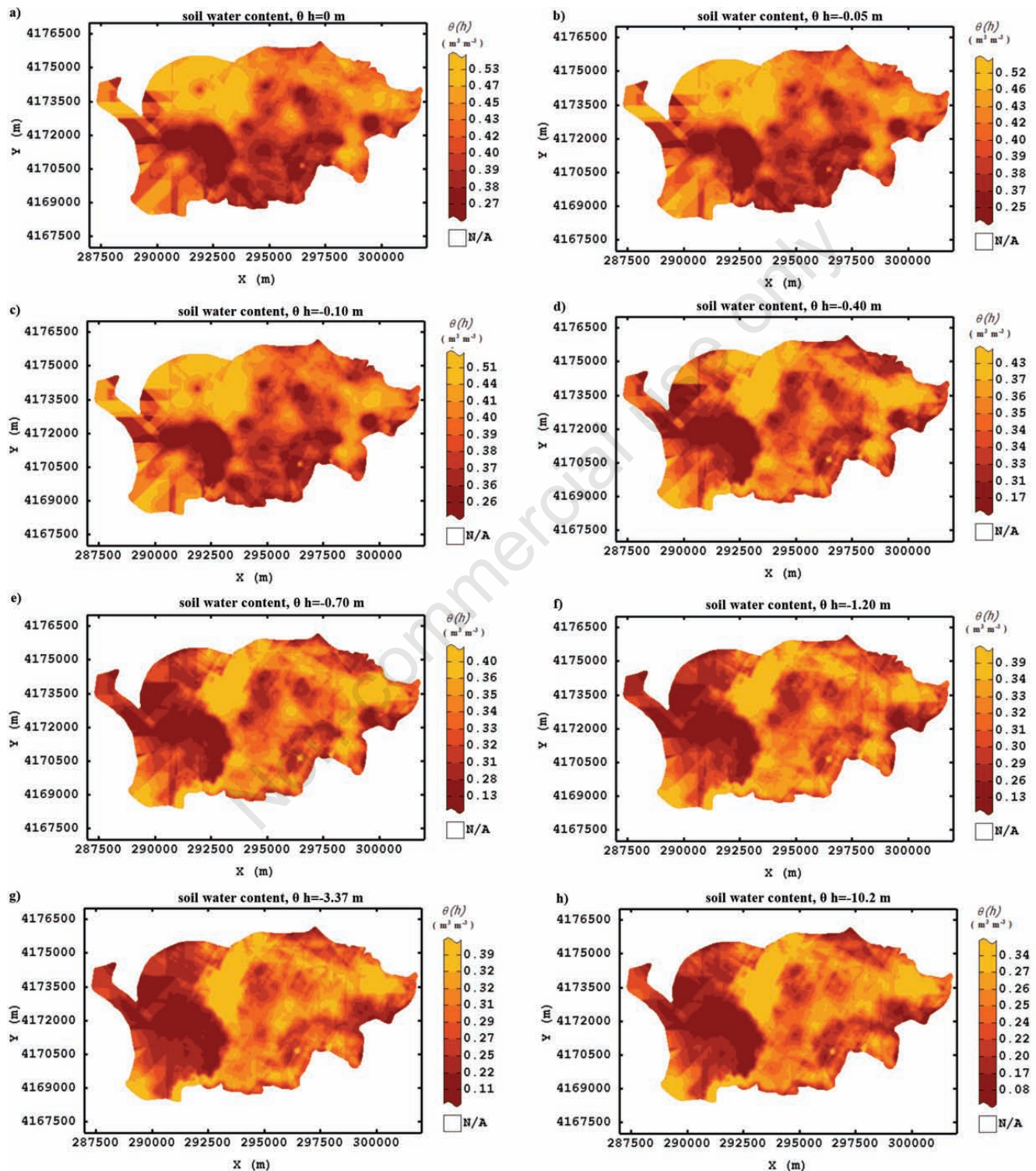


Figure 5. Co-kriging maps of the analysed variables (a-h).



The western sector of the basin, where the terraced deposits outcrop, is characterized by high sand and silty sand content but very low clay contents. In this area the soil is also characterized by abundant skeleton and low  $\theta$  at the lowest matric potential.

On the contrary, the clay content is higher and the sand content very poor in some areas of the floodplain, where  $\theta$  at the lowest potential is tendentially higher. The electrical conductivity map shows a significant

dependence on the topography, since the lowest values of EC are located in areas corresponding to the steeper slopes, whereas the flat area in the floodplain is characterized by a more conductive soil.

Most spatial variation occurs at the shorter scale, *i.e.* within a distance of 1000 m, as results from the sum of the eigenvalues at the different spatial scales (Table 3). Retaining only the eigenvectors with eigenvalues greater than one, the first factor at the shorter range

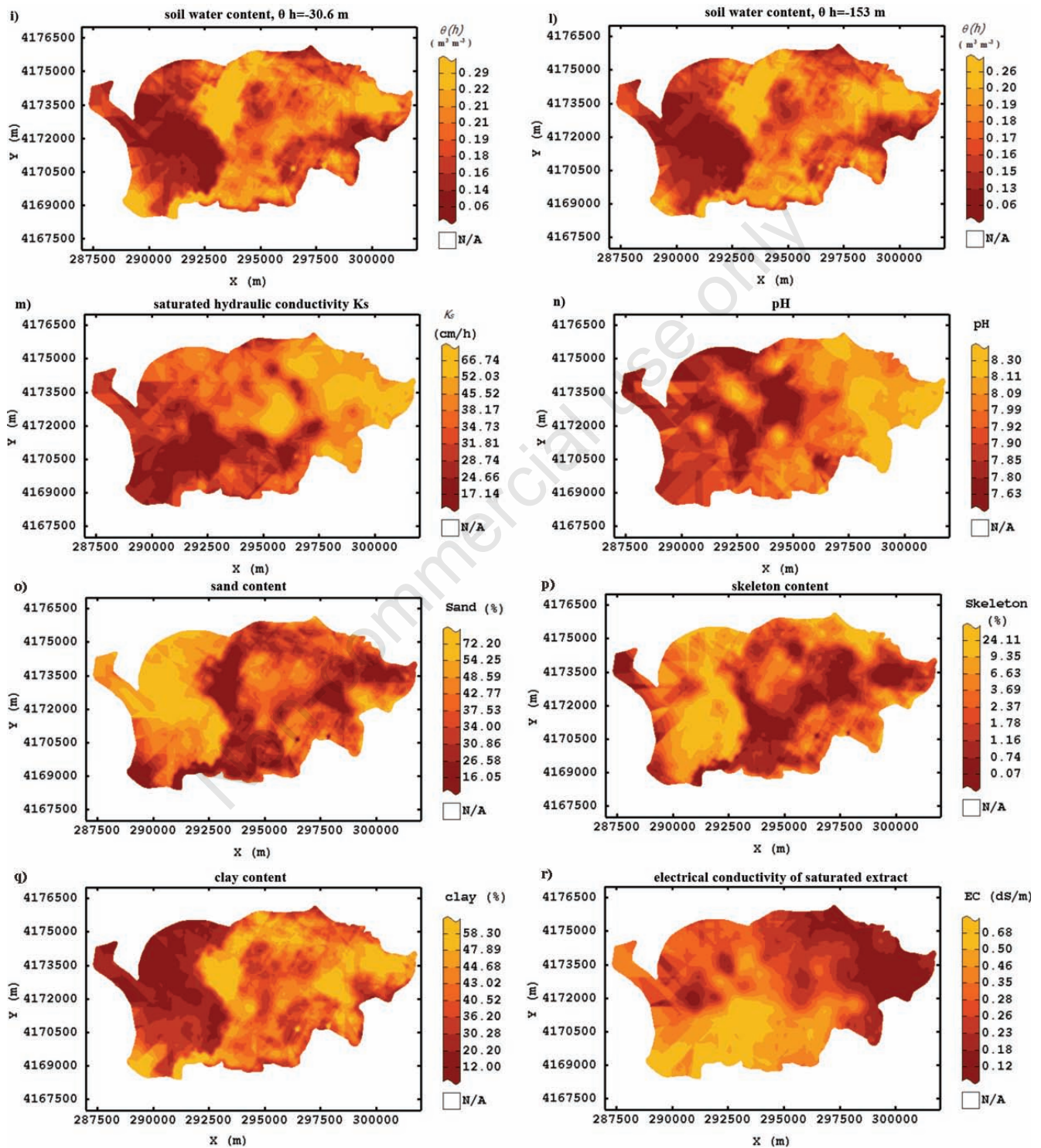


Figure 5. Co-kriging maps of the analysed variables (i-r).

Table 3. Decomposition into regionalised factors of the Gaussian variables. Only Eigenvectors corresponding to Eigenvalues greater than 1 are reported.

S1: Spherical range=1000 m. Sum of the eigen values=6.29.																			
	g $\theta$ h= 0 m*	g $\theta$ h= -0.05 m*	g $\theta$ h= -0.10 m*	g $\theta$ h= -0.40 m*	g $\theta$ h= -0.70 m*	g $\theta$ h= -1.20 m*	g $\theta$ h= -3.37 m*	g $\theta$ h= -10.2 m*	g $\theta$ h= -30.6 m*	g $\theta$ h= -153 m*	gClay°	gSand#	gSkeleton§	gKs^	gEC1:5§	gPH	gElevation	Eigen value	Var. %
Factor 1	0.3199	0.3218	0.3183	0.303	0.2875	0.2623	0.2221	0.2642	0.2389	0.2367	0.2178	-0.256	-0.201	0.143	0.0178	0.17	-0.116	6.2014	74.72
S2: Spherical range=4000 m. Sum of the eigen values=4.93.																			
	g $\theta$ h= 0 m*	g $\theta$ h= -0.05 m*	g $\theta$ h= -0.10 m*	g $\theta$ h= -0.40 m*	g $\theta$ h= -0.70 m*	g $\theta$ h= -1.20 m*	g $\theta$ h= -3.37 m*	g $\theta$ h= -10.2 m*	g $\theta$ h= -30.6 m*	g $\theta$ h= -153 m*	gClay°	gSand#	gSkeleton§	gKs^	gEC1:5§	gPH	gElevation	Eigen value	Var. %
Factor 1	0.0346	0.0467	0.0481	0.1362	0.1471	0.1641	0.2956	0.3089	0.3175	0.3345	0.214	-0.272	-0.214	-0.014	0.2321	-0.18	-0.53	2.8013	56.76
Factor 2	0.3777	0.3327	0.3136	0.1549	0.07	0.0292	-0.1315	0.1905	0.1929	0.1848	-0.016	0.0322	0.3546	0.313	-0.441	-0.13	0.2414	1.3432	27.21

\*Soil water content,  $\theta$ , corresponding to h values ranging from 0 to -153 m; °clay content; §skeleton content; #sand content; ^saturated hydraulic conductivity Ks; §electrical conductivity of saturated extract.

explains most of the spatial variability, more than 74% of the variance at this scale, whereas, at the longer range, the first factor explains the 57% and the second factor the 27% of the variance.

On the first factor at shorter range all  $\theta$ , of the retention curve and clay weigh mainly and positively; on the contrary, on the first factor at longer range the elevation weighs more and negatively, whereas clay, electrical conductivity and  $\theta$ , at the potentials less than -3.37 m, weigh positively (Table 3).

The second factor at longer range, which describes only a small proportion of the variation, is more related to the behaviour of the soil at the saturation and at the higher water potentials (Table 3).

In summary, we can say that the first factor at shorter range provides an overall description of the soil variability occurring within a very short distance (less than 1000 m), whereas the two factors at longer range contribute to describe two complementary conditions: the first factor is more related to water stress conditions, preferentially occurring in the lowest sectors of the area, more likely at soil salinisation hazard; the second factor provides a description of the behaviour of the soil in no water stress conditions. The cokriging maps were displayed only for the first regionalised factors at the two spatial scales, because of the variance percentage explained by them; they can then be used to localize the natural and anthropogenic processes affecting soil quality.

The map of the first factor at shorter range ( $F1_{1000\text{m}}$ ) shows a sensi-

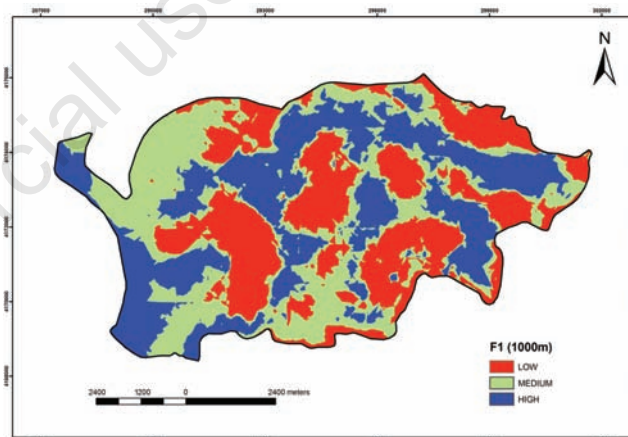


Figure 6. Map of the first regionalized factor at shorter range.

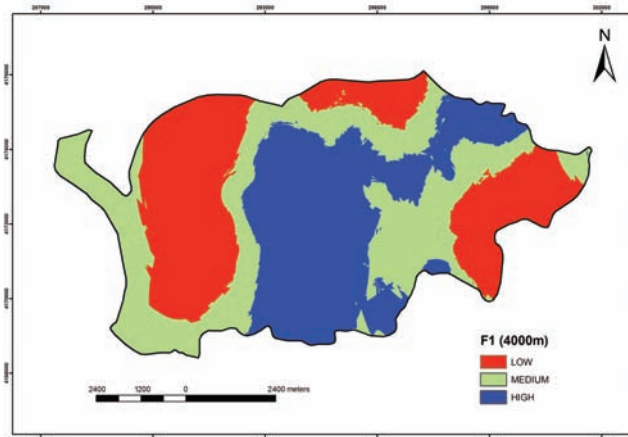


Figure 7. Map of the first regionalized factor at longer range.



ble erratic distribution, characterized by several hot spots due to local variation of the soil hydraulic properties, which might be attributed to differences in the intrinsic properties of the soil and/or different soil management (Figure 6). The map of the first regionalized factor at longer range ( $F_{1,4000\text{ m}}$ ) highlights the presence of some wide areas characterized by different scores of the factor (Figure 7), ascribable to a different degree of soil salinization hazard, owing to the positive relationships between the factor and the electrical conductivity. The higher values of this factor are mainly distributed in the central and the lowest sectors of the floodplain, characterized by high clay contents which favour salt accumulation in the soil.

On the contrary, the lowest factor values are mainly distributed on the western portion and on the northern and south-eastern boundaries of the basin, where the lithological and morphological conditions make easy the soil drainage. In the western sector the presence of terraced sandy deposits determines a quite permeable soil, easily draining water in depth. In the other sectors of the basin, characterized by the lowest factor values, the steep slopes determine high energy of surface water, which causes a washout action that does not allow salt accumulation in the top soil.

## Conclusions

The proposed approach, based on a set of geostatistical techniques, allowed partitioning a basin of south-western Sicily into areas characterized by different degree of soil salinization.

Multivariate geostatistics has allowed to disclose the spatial structures of coregionalization of the variables that are involved in the soil salinization processes and to map the most relevant regionalized factors, so to delineate regions at different salinization risk.

The geostatistical analysis has revealed that the soil salinization risk in the study site might be related to both anthropogenic actions, mainly acting at local scale and due to agricultural activities, and environmental settings of the basin, such as the lithological and textural compositions, the topography and the hydraulic soil properties.

The proposed methodology could be a valid support for land use and soil remediation planning at regional scale. Moreover, the obtained results could also be used to direct the farmers to intensify soil sampling and monitoring in those areas at high risk of salinization.

## References

- ASA-SSSA, 1986. Methods of soil analysis. Part 1 - Physical and mineralogical methods. A. Klute editor, Madison, WI, USA.
- Burke W, Gabriels D, Bouma J, 1986. Soil structure assessment. Balkema Ed., Rotterdam, The Netherlands.
- Buttafuoco G, Castrignanò A, 2005. Study of the spatio-temporal variation of soil moisture under forest using intrinsic random functions of order k. *Geoderma* 128:208-220.
- Castrignanò A, Giugliarini L, Risaliti R, Martinelli N, 2000. Study of spatial relationships among some soil physico-chemical properties of a field in central Italy using multivariate Geostatistics. *Geoderma* 97:39-60.
- Castrignanò A, Buttafuoco G, Puddu R, 2008. Multi-scale assessment of the risk of soil salinization in an area of south-eastern Sardinia (Italy). *Precis. Agri.* 9:17-31.
- Castrignanò A, Costantini AC, Barbetti ER, Sollitto D, 2009. Accounting for extensive topographic and pedologic secondary information to improve soil mapping. *Catena* 77:28-38.
- Chiles JP, Delfiner P, 1999. *Geostatistics: Modelling Spatial Uncertainty*. Wiley Ed., New York, NY, USA.
- Corwin DL, Lesch SM, 2003. Application of soil electrical conductivity to precision agriculture: Theory, principles, and guidelines. *Agr. J.* 95:455-471.
- Crescimanno G, 1988. An integrated approach for sustainable management of irrigated land susceptible to degradation/desertification. *Proc. European Climate Science Conf., Vienna, Austria*, 11-26.
- Dane JH, Hopman JW, 2002. Water retention and storage. In: JH Dane and GC Topp (eds.) *Methods of soil analysis. Part 4 - Physical methods*. SSSA Book Ser. 5. SSSA, Madison, WI, USA.
- Douaik A, Van Meirvenne M, Tóth T, 2004. Spatio-temporal kriging of soil salinity rescaled from bulk soil electrical conductivity. In: X Sanchez-Vila, J Carrera and J Gomez-Hernandez (eds.) *GeoEnv IV: geostatistics for environmental applications*. Kluwer Academic Publ., Dordrecht, The Netherlands.
- Klute A, 1986. *Methods of Soil Analysis. Part 1 -Physical and Mineralogical properties*. 2nd ed., American Society of Agronomy Publ., Madison, WI, USA.
- Lajaunie C, Béhaxétéguy JP, 1989. Elaboration d'un programme d'ajustement semi-automatique d'un modèle de corégionalisation - Theorie. Technical report N21/89/G. ENSMP Publ., Paris, France.
- Lauchli A, Epstein E., 1990. Plant response to salinity and sodic conditions. In: KK Tanji (ed.) *Agricultural salinity assessment and management*. American Society of Civil Engineers Publ., New York, NY, USA.
- Matheron G, 1973. The intrinsic random functions and their applications. *Adv. Appl. Prob.* 5:239-465.
- Ministry for the Environment, Land and Sea, 2007. Fourth National Communication under the UN Framework Convention on Climate Change, Italy. Available from: <http://unfccc.int/resource/docs/natc/itanc4.pdf>
- Pozdnyakova L, Zhang R, 1999. Geostatistical analyses of soil salinity in a large field. *Precis. Agri.* 1:153-165.
- Rhoades JD, Chanduvi F, Lesch SM, 1999. Soil salinity assessment methods and interpretation of electrical conductivity measurements. *Irrigation and drainage. Paper 57*. FAO Publ., Roma, Italy.
- Rivoirard J, 2001. Which models for collocated cokriging? *Math. Geol.* 33:117-131.
- Soil Survey Staff, 1999. *Soil Taxonomy. Agriculture Handbook 436*. USDA-NRCS Publ. Washington, DC, USA.
- Sollitto D, Romic M, Castrignanò A, Romic D, Bakic H, 2010. Assessing heavy metal contamination in soils of the Zagreb region (Northwest Croatia) using multivariate geostatistics. *Catena* 80:182-194.
- Tóth G, Montanarella L, Rusco E, 2008. Updated map of salt affected soils in the European Union. In: G Tóth, L Montanarella and E Rusco (eds.) *Threats to Soil Quality in Europe EUR23438-cientific and Technical Research series*. Official Publications of the European Communities Luxembourg, pp 61-74.
- Wackernagel H, 2003. *Multivariate Geostatistics: an introduction with application*. Springer Ed., New York, USA.