We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Statistical Methods for the Analysis of Soil Spatial and Temporal Variability

Ahmed Douaik[1], Marc van Meirvenne[2] and Tibor Tóth[3,4]
*[1]Research Unit on Environment and Conservation of Natural Resources,
National Institute of Agricultural Research (INRA), Rabat,
[2]Department of Soil Management, Ghent University, Ghent,
[3]Department of Soil Science, Research Institute for Soil Science and Agricultural
Chemistry, Budapest,
[4]Rural, Water and Ecosystem Resources Unit, Institute for Environment and
Sustainability, European Commission – Joint Research Center, Ispra
[1]Morocco
[2]Belgium
[3]Hungary
[4]Italy*

1. Introduction

Soils vary in space and time due to the combined effects of physical, chemical, and biological processes that operate at different scales and with different intensities. This variability is caused by soil forming factors, i.e., climate, parent material, time, topography, and vegetation (Jenny, 1980), farming practices (fertilization, irrigation, tillage, etc.), erosion, etc. Each cause may operate independently or in combination with other factors and over a wide range of spatial and/or temporal scales. Spatial scales can manifest at microscale to watershed and more, while temporal scale can span from seconds to decades and longer. These processes and causes create a pattern of nested variability or soil heterogeneity. This scale-dependency means that soil properties may display spatial and/or temporal patterns only over certain distances and not over others.

Characterizing spatial variability of soil properties and crop parameters by inventorying them is needed to evaluate the effectiveness of their management. Also, characterizing their temporal variability by monitoring them is important in order to compare different management systems for sustainability and environmental quality. In addition, an understanding of the spatial and/or temporal variability can provide a framework for developing effective sampling schemes for future site management and efficient experimental designs for research approaches.

Due to the spatial and/or temporal variability of soil properties, numerous samples need to be taken and the measurements need to be repeated as conditions change or to determine if they are changing. Soil properties are traditionally determined in laboratory from samples taken from field. However, this approach, although more precise, is more time consuming and expensive. Alternatively, soil properties can be easily, quickly, cheaply, and more

intensively measured, even indirectly, using terrain attributes derived from a DEM and using emerging soil-landscape mapping technologies like GPS, geophysical techniques (EMI, GPR, TDR, etc.), soil spectroscopy and remote sensors, etc. Even with these auxiliary or proxy data sources, a full inventory of soil resources and properties is non-realistic and impossible. Therefore, soil can be measured at only limited space locations and time instants. Soil properties at non-sampled locations and/or times should be predicted. These predictions and related uncertainties can be obtained by using statistical models for soil spatial and/or temporal variability. In this way, soil variability can be considered as the combination of a systematic (deterministic) and a random (stochastic) components. The former represents the gradual or distinct changes (trends) while the latter involves differences not related to a known cause and those not discerned by the nature or the scale of investigation. The choice of the adequate model depends on the behaviour of the phenomenon of interest (a priori knowledge), the objective of the analysis, the nature of the available data, and, from a practical point of view, the existence of good statistical methods and adequate software packages.

Excellent review papers on modelling spatial and temporal variability were published a decade ago (Heuvelink & Webster, 2001; Kyriakidis & Journel, 1999). The latter review focused only on geostatistical methods for modelling particularly the joint space-time variability while the former discussed methods for modelling purely spatial, purely temporal, and joint space-time variability.

Our chapter is basically a review of statistical methods that have been used in order to describe and model the spatial and temporal variability of soil resources and properties. Once this variability is characterized and modelled, it can be advantageously used for different practical applications like optimal sampling of soil resources, efficient designing of experiments, interpolation and mapping of soil properties, etc. Therefore, in this chapter, we will present the main statistical methods with a short description of each of them, discuss eventually the possible links between some of them, identify their assumptions, merits and limitations, stress some of their practical applications, and list some relevant bibliographic references. The chapter fills the gap from the Heuvelink & Webster (2001) paper by considering more statistical methods like temporal stability (Douaik et al, 2006; Vachaud et al, 1985) and dynamic spatial variation (Douaik et al, 2007; Lesch et al, 1998), updates the presentation of the methods from the Kyriakidis & Journel (1999) paper by taking into account the research work done, and especially the last developments in pedometrics, during the past decade and adds some examples of practical applications like mapping using the Bayesian maximum entropy (Christakos, 2000; Douaik et al, 2005; 2009).

After presenting some classical statistical methods (Section 2), we provide an overview of the geostatistical methods including structural analysis (Section 3) and prediction (Section 4). Then, we present the Bayesian maximum entropy, another method of prediction (Section 5). Finally, we discuss the different methods and conclude with some suggestions for choosing among them (Section 6).

2. Classical statistical methods

Classical statistical methods are basically aspatial and atemporal methods where the spatial and temporal coordinates are ignored. The different approaches are: coefficient of variation (2.1), variance component analysis (2.2), temporal stability (2.3), and dynamic spatial variation (2.4).

2.1 Coefficient of variation

It is assumed that the observed value for a soil attribute (z) at any location (s) at a given time instant (t) is:

$$E[z(s,t)] = \mu + \varepsilon(s,t) \tag{1}$$

with μ the population mean and $\varepsilon(s,t)$ random, spatially and temporally uncorrelated error assumed normally distributed with zero mean and variance σ^2.

The population mean is estimated from a sample:

$$\hat{\mu} = \overline{z} = \frac{1}{n}\sum_{i=1}^{n} z_i \tag{2}$$

with n: the sample size and the variance, σ^2, is estimated by:

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(z_i - \overline{z}\right)^2 \tag{3}$$

Its root square, $\hat{\sigma}$, is the standard deviation. The coefficient of variation is defined as:

$$CV = 100 * \frac{\hat{\sigma}}{\overline{z}} \tag{4}$$

The coefficient of variation can be calculated for the temporal probability distribution function (PDF) at each space location as well as for the spatial PDF at each time instant. As an example, Van Wesenbeeck & Kachanoski (1988) studied the spatial distribution (from 100 locations) of the surface soil water content throughout the growing season (65 measurement dates). They concluded that there exists a temporal dependence of the spatial variability and a spatial dependence of the temporal variability. They also suggested that random sampling of the locations should be avoided. James et al. (2003) defined the temporal variability of soil moisture content as the coefficient of variation at each site across the whole year (52 time instants) and used an analysis of variance (ANOVA) to test for differences between habitats (three sites) in the temporal variability of soil moisture content. They found that soil moisture varied significantly with time but not between habitats while the interaction between time and habitat was significant which reflects differences between habitats in temporal pattern of soil moisture. The coefficients of variation were significantly different between habitats. Other examples of the use of the coefficient of variation for the assessment of the spatial and temporal variabilities can be found in Ehrenfeld et al. (1997) and Guo et al. (2002).

The standard deviation and, hence, the coefficient of variation does not incorporate spatial and/or temporal information and therefore doesn't provide a representation of the nature of the spatial and/or temporal behaviour of a given soil property. It is considered as an indicator of global variability as opposed to local variability (Carter & Pearen, 1985), the latter is identified using correlogram, covariance function, or variogram which will be presented later in this chapter (sections 3.3 to 3.6).

2.2 Variance component analysis

In the variance component analysis approach, the spatial and temporal contributions are considered as factors with random effects in the ANOVA model. In this way the different

components of the total variance can be determined. Campbell et al. (1989) adopted the following model, for the log concentrations of a given solute (log c) in subsample l (3) from sampling area k (2) at site j (2) and at time i (6 dates over 20 days):

$$\log c_{ijkl} = \mu + T_i + S_j + A_k + (A.U)_{kl} + \varepsilon_{ijkl} \tag{5}$$

with μ: the grand mean; T: the time effect; S: the site effect; A: the sampling area effect; $(A.U)$: the subsampling within sampling area effect, and ε: an error term. The temporal and spatial variances were compared to the sampling area variance while the latter was compared to the subsampling within sampling area variance. They found that the temporal variation was far less significant than the variation between the two sites. Also they found that the variation between areas a few meters apart was greater than that between the two sites even though the texture and the management of the sites were quiet different. Other examples of the evaluation of the temporal and spatial variabilities based on the concept of variance component analysis can be found in Van Es (1993) and Van Es et al. (1999).

The variance component analysis is directly related to the coefficient of variation presented in the precedent section. The variance component estimates for each variability source are computed, and then the square root of these estimates is divided by the mean value. The output is exactly the definition of the coefficient of variation. Thus, the variance component analysis is suffering from the same drawbacks as the coefficient of variation approach.

2.3 Temporal stability

The concept of temporal stability was first introduced by Vachaud et al. (1985). It is defined as the time invariant association between spatial location and statistical parameters of soil properties. They distinguished mainly two approaches: relative differences and Spearman rank correlation.

2.3.1 Relative differences

Let z_{ij} be the observed value of a soil property at location i ($i=1, …, n$) and time j ($j=1, …, m$). The relative differences δ_{ij} are defined as:

$$\delta_{ij} = \frac{z_{ij} - \overline{z}_j}{\overline{z}_j} \tag{6}$$

with $\overline{z}_j$: the spatial average for the time instant j. The temporal average of δ_{ij} is:

$$\overline{\delta}_i = \frac{1}{m} \sum_{j=1}^{m} \delta_{ij} \tag{7}$$

and its corresponding temporal standard deviation is defined as:

$$\sigma\left(\overline{\delta}_i\right) = \sqrt{\frac{1}{m-1} \sum_{j=1}^{m} \left(\delta_{ij} - \overline{\delta}_i\right)^2} \tag{8}$$

A zero value for $\overline{\delta}_i$ indicates that the temporal average $\overline{z}_j$ represents the average value over the whole study area at any time. The field average value is overestimated if $\overline{\delta}_i > 0$ while it is underestimated when $\overline{\delta}_i < 0$. A more time stable location will be indicated by a

small value of $\sigma\left(\bar{\delta}_i\right)$ whereas a high value of the latter is an indication of a less time stable location. The $\bar{\delta}_i$ values can be plotted graphically against their rank with the corresponding temporal standard deviations.

Based on the relative differences, Kachanoski & De Jong (1988) showed that a useful test for temporal stability is the Pearson correlation coefficient between soil property values measured at two consecutive time instants. They also showed that temporal stability exists if the relative differences remain constant between two time instants: $\delta_{i2} = \delta_{i1}$. This equation implies that:

$$z_{i2} = \frac{\bar{z_2}}{\bar{z_1}} z_{i1} \tag{9}$$

$\bar{z_1}$ and $\bar{z_2}$ being spatial averages for the time instants j = 1 and 2, respectively. This equation establishes a linear relationship between the soil property at two different times with an intercept equal to zero and a slope equal to the ratio of the mean value observed at the second time frame to the mean value observed at the first time frame. Consequently, a good test for the temporal stability is the correlation between soil property values measured at consecutive time frames (Kachanoski & De Jong, 1988).

Furthermore, the former equation implies that the regression between relative differences from two consecutive time frames should have a zero intercept and a unity slope. This is another way to check the existence of temporal stability of the spatial pattern of a variable. However, the regression between z_{i2} and z_{i1} is, in general, of the following form:

$$z_{i2} = a * z_{i1} + b \tag{10}$$

Using equations (9) and (10) we expect that, if temporal stability exists, the following conditions are verified:

$$a = \frac{\bar{z_2}}{\bar{z_1}} \quad \text{and} \quad b = 0 \tag{11}$$

So, another way to check the temporal stability is to test for a zero intercept and a unity slope of the regression between two consecutive time instants.

If there is a constant increase or decrease in the soil property, w, in all the locations between two time frames, we expect from Eq. (10) that (Kachanoski & De Jong, 1988):

$$a = 1 \text{ and } b = w \tag{12}$$

Equation (12) indicates that if the slope of the regression between values at two consecutive time instants is not significantly different from one, then the regression intercept represents the constant change that has occurred between the two time instants.

Based on whether the intercept (b) is significantly different from 0, the slope (a) is significantly different from 1, or both, four possible scenarios can be distinguished (Douaik et al, 2007; Grant et al., 2004):

- *Scenario* 1 ($a = 1$ *and* $b = 0$): there is no net change between the two sampling dates
- *Scenario* 2 ($a \neq 1$ *and* $b = 0$): the mean soil property level has changed and there is a static (uniform) change in the soil property spatial pattern
- *Scenario* 3 ($a = 1$ *and* $b \neq 0$): the mean soil property level has not changed and the change in the soil property spatial pattern is dynamic (non-uniform) but, when averaged, this change is not significant

-   *Scenario* 4 (*a* ≠ 1 *and b* ≠ 0): the mean soil property level has changed and the change in the soil property spatial pattern is dynamic (non-uniform).

These refinements of the concept of temporal stability were used by, among others, Da Silva et al. (2001), Douaik (2005), Douaik et al (2006; 2007), Ehrenfeldt et al. (1997), Goovaerts & Chiang (1993), Petrone et al. (2004), Van Wesenbeeck et al. (1988).

The concept of relative differences is based on the same statistical parameters (mean and standard deviation) as the two precedent approaches. However, it uses the standardized data (relative differences) for each time instant instead of the original data and the mean and standard deviation are computed for each space location. This concept is based, like the two precedent approaches, on the Gaussian distribution of data and the absence of autocorrelation. Both assumptions must be checked before the use of relative differences in the study of temporal stability. The normality can be checked using the standard statistical tests and the independence can be verified using either a test for autocorrelation like the Moran test or by computing the variogram and determining its range. Normality can be gained via data transformation and independence by the selection of sample locations separated by distances larger than the autocorrelation range. For non normally distributed data, but still non autocorrelated, the Spearman rank correlation can be used. The latter is presented in the next section.

2.3.2 Spearman rank correlation

It refers to the tendency of a soil property, measured at different locations in space, to maintain their relative ranking over time. It is defined as:

$$r_s = 1 - \frac{6\sum_{i=1}^{n}\left(R_{ij} - R_{ik}\right)^2}{n\left(n^2 - 1\right)} \tag{13}$$

with R_{ij} and R_{ik} the ranks of z_{ij} observed at location *i* on time instants *j* and *k*, respectively. A value of this coefficient equal to one indicates a perfect temporal stability between time instants *j* and *k*, and thus identity of ranks for any location whereas a lack of temporal stability implies that $r_s = 0$. Many researchers used this concept to evaluate the temporal stability of different soil properties, among them, Campbell et al. (2001), Douaik (2005), Farley & Fitter (1999), Guo et al. (2002), Heathman et al. (2009), Hu et al (2011), Reichardt et al. (1993), Si (2002), and Wendroth et al. (1999).

The Spearman rank correlation coefficient is a non parametric statistics therefore it can be used even for non normally distributed data. However, in the case of autocorrelation, its use is not recommended and it will be cautious to use statistical methods which consider the spatial and temporal nature of the given soil property. In addition, even for independent data, the Spearman rank correlation inform us only about the relative ranking of space locations between two time instants and do not provide a quantification of the temporal change. This quantification can be obtained using either the paired-*t* test for comparing mean values or the dynamic spatial variation approach. The former is presented in what follows while the latter is discussed in section 2.5.

2.4 The Paired-*t* test

Equation (9) just indicates if there is a constant change or not between two time instants, but we still need to check if the mean values of a soil property between these two times are

significantly different or not. To answer this question, the paired-t test (McClave & Sincich, 2006) is used. The test statistic t is computed as:

$$t = \frac{\bar{d}}{s_d / \sqrt{n}} \tag{14}$$

with $\bar{d}$ the mean of the differences, d_i, between the observations for the second and the first time frames:

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i = \frac{1}{n}\sum_{i=1}^{n}(z_{i2} - z_{i1}) \tag{15}$$

s_d is the standard deviation of the differences:

$$s_d = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2} \tag{16}$$

and n the number of the differences, i.e. the number of locations.

The test statistic t (equation 14) has (n-1) degrees of freedom. The calculated statistic is compared to a tabulated t value with the same degrees of freedom and a probability of error of type I (taken generally equal to 5%). Equivalently, the corresponding probability to the test statistic t is compared to 5% to check if the difference of mean values is significant or not. As an example of application of the paired-t test, Kenny et al. (2002) checked the temporal trend of the mean thickness of the *Ap* horizon.

2.5 Dynamic Spatial variation

The concept of temporal stability handles each soil property separately and, as described before (section 2.3), it can quantify the temporal change even without testing if this change is significant or not. Lesch et al. (1998) discussed another approach that can test for change in the mean value as well as for a dynamic spatial variation between two time instants. Dynamic spatial variation refers to the spatially variable change of a soil property between two time instants, i.e., the change is not the same between two time instants for each location.

To test for the existence of dynamic spatial variation, Lesch et al. (1998) identified different steps. First of all, the pairs of soil data obtained at a first time instant (y and z), from n space locations, are used to establish a calibration equation based on the regression model:

$$z = b_1 * y + b_0 \tag{17}$$

The parameters of this regression equation (b_0 and b_1) are estimated using the ordinary least squares method. In a second step, the residuals of the model are tested for spatial independence, normality, homoscedasticity and the linearity of the relationship. The three last assumptions can be verified using classical regression tools. The assumption of spatial independence can be checked using the Moran test (Lesch et al., 1995). In the next step, a mixed ANOVA model is set:

$$z_{ij} = \mu + t_i + b_j + \varepsilon_{ij} \tag{18}$$

i=1, 2 are the two time instants and j=1, … , n are the locations. This model is equivalent to:

$$z_{2j} - z_{1j} = d_j = d_0 + \eta_j + \left(\varepsilon_{2j} - \varepsilon_{1j}\right)$$ (19)

with $d_0 = t_2 - t_1$ is the difference between mean values for the second and first time instants and η_j independently and identically normally distributed with zero mean and variance equal to $\sigma_\theta^2 : \eta_j \approx N\left(0, \theta^2\right)$. If z_{1j} represents the observed values of the first soil property at the first time instant and at locations j (j=1, …, n) and z_{2k} represents the same property observed at the second time instant at the locations k (k=1, …, m), the difference between the two values is given by Eq. (19) with d_0 representing the shift in the average value between the two time instants and η representing the dynamic spatial variation. These two quantities can be statistically tested: the first if it is different from zero and the variance of the second if it is higher than zero.

Based on whether the first parameter (the mean difference, d_0) is significantly different from zero, the second parameter (spatial shift, θ^2) is significantly higher than zero, or both, four possible scenarios can be distinguished (Douaik et al., 2007):

- *Scenario* 1 ($d_0 = 0$ *and* $\theta^2 = 0$): there is no net change with time in the soil property
- *Scenario* 2 ($d_0 \neq 0$ *and* $\theta^2 = 0$): the mean soil property level has changed and there is a static (uniform) change in the soil property spatial pattern
- *Scenario* 3 ($d_0 = 0$ *and* $\theta^2 > 0$): the mean soil property level has not changed and the change in the soil property spatial pattern is dynamic (non-uniform); again, when averaged, this change is not significant
- *Scenario* 4 ($d_0 \neq 0$ *and* $\theta^2 > 0$): the mean soil property level has changed and the change in the spatial pattern of the soil salinity is dynamic (non-uniform).

An application of the approach of Lesch et al (1998) can be found in Douaik (2005), Douaik et al (2007) and Tóth et al (2002).

The approach of Lesch et al (1998) is based on the determination of a linear regression between two soil properties; hence it assumes that the residuals from the regression model are normally distributed with equality of variances and spatially not autocorrelated. All these assumptions must be checked. If spatial autocorrelation is present in the residuals, the estimates of the regression parameters and their variances are biased and the statistical tests are no longer valid. In this special case, spatial regression and/or geostatistical approaches (sections 3 and 4) are recommended.

3. Geostatistics: Structural analysis

Structural analysis is concerned with the description and the modelling of the spatial and/or temporal variability using geostatistical tools like the covariance function, the correlogram, or the variogram. The analysis of space-time data requires the definition of a space-time continuum with a coordinate system and a measure of space-time distance and models and techniques which make possible the link between spatio-temporally distributed data (Christakos et al., 2002). The spatial coordinates are defined, in general, in two dimensions s = (s_1, s_2) and the temporal coordinate t along the temporal axis such that the space-time coordinates are p = (s, t).

3.1 Space-Time Random Field model

Christakos et al. (2002) defined space-time random field (STRF) model as 'a mathematical construction that rigorously represents the distribution of natural phenomena across space and time. The STRF model provides scientifically meaningful representation, explanation, and prediction of these phenomena in uncertain environments. These uncertainties may be due to measurement errors, heterogeneous databases, erratic fluctuations in the space-time variation of the underlying process, and insufficient knowledge'. An STRF, $X(p)$, is a collection of realizations χ of the space-time distribution of a natural variable. It can be viewed as a collection of correlated random variables $x=[x_1, ..., x_n]'$ at the space-time points $p=[p_1, ..., p_n]'$ and a realization of the STRF at these points is $\chi=[\chi_1, ..., \chi_n]'$. It assigns a probability that a realization χ in n dimensions will occur following the multivariate probability distribution function (PDF) of $X(p)$ or, equivalently, its cumulative distribution function (CDF). The PDF and CDF represent a complete characterization of the STRF. A partial, but sufficient, characterization of the STRF is provided by its space-time statistical moments. The first order moment, which is the mean function, expresses trends or systematic structures in space-time. The moment of second order, the covariance function, expresses correlations and dependencies between two different points p and p'. In addition to these two moments, an STRF can be also characterized by its variogram.

3.2 Conceptual approaches

Christakos and Raghu (1996) reported that some forms of averaging over time or space are done before the analysis of the spatial or temporal random fields. In doing so, the joint space-time variability is not accounted for and there is some loss of information. Initially to consider both space and time variability, some authors resorted to some simplifications. For example, Bilonick (1985) and Egbert & Lettenmaeir (1986) decomposed their STRF in a purely spatial and a purely temporal components whereas Rodriguéz-Iturbe & Eagelson (1987) considered separable models where spatial variations were assumed to be independent of temporal variations.

Three main approaches can be distinguished for the analysis of the geostatistical space-time data (Kyriakidis & Journel, 1999; Stein & Sterk, 1999; Stein et al., 1998):

- Methods based on vectors of spatial random fields (SRF) for the case of many space locations and few time instants (Bogaert & Christakos, 1997; Papritz & Fluhler, 1994). It does not include the temporal dependence existing between observations and can predict only at the observed time instants. The spatial variability is modelled either through a separable spatial variogram for each time instant, or by a single spatial variogram considering time instants as replicates;
- Methods based on vectors of time series (temporal random fields, TRF) for the case of long time series with few space locations (Rouhani & Wackernagel, 1990; Solow & Gorelick, 1986). This approach doesn't take into account the spatial dependence and it predicts only at the observed locations. In a similar way as above, independent TRF or TRF as replicates in space can be considered.
- Methods using an STRF, thus considering the joint space-time variability;

For the first approach, the focus of the analysis is on smooth interpolated soil attribute maps over specific time instants. The intention is to capture single instantaneous snapshots (a static picture) of the natural process. The objective can be also the comparison of the various

maps or the detection of the temporal persistence or change in the spatial patterns (Goovaerts & Chiang, 1993; Van Meirvenne et al., 1996). The intention of the second approach is to capture a sequence of successive snapshots at single space locations, which give temporal profiles. Only the third approach includes both the spatial and temporal dependencies so the interpolation is more precise and can be done for unsampled time instants at unsampled space locations. The focus is here on video sequence of successive spatial pictures like a movie.

As a fully satisfactory stochastic model should involve explicitly both spatial and temporal aspects, we present in the following sections the STRF concept and, subsequently, some special cases. The spatial and/or temporal variability can be described and modelled using either the covariance function or the variogram; although the latter is far more used in the geostatistical literature than the former and, therefore, only the variogram forms will be presented here. Also, in the case of two different soil properties or for the same soil property but observed at two different spatial locations or time instants, the joint variability can be characterized using the cross covariance function or the cross variogram. We present in section 3.3 the joint STRF while the spatial random field (SRF) is presented in section 3.4 and the temporal random field (TRF) in section 3.5. Finally, the space-time, the spatial, and the temporal cross variograms are discussed in section 3.6.

3.3 Joint Space-Time Random Field model

The space time variogram is defined as:

$$\gamma_X(h,\tau) = \frac{1}{2}\mathrm{var}\left[X(s+h,t+\tau) - X(s,t)\right] = \frac{1}{2}E\left\{\left[X(s+h,t+\tau) - X(s,t)\right]^2\right\} \tag{20}$$

It is estimated, from the observed data, by (Stein et al., 1998):

$$\hat{\gamma}_X(h,\tau) = \frac{1}{2N(h,\tau)} \sum_{i,j=1}^{N(h,\tau)} \left[x(s_i,t_j) - x(s_i+h,t_j+\tau)\right]^2 \tag{21}$$

with $x(s_i,t_j)$ and $x(s_i+h,t_j+\tau)$ are pairs of observations with a spatial distance equal h at the time instant t_j, the total number of such pairs is $N_j(h)$, and a temporal distance equal τ, the number of such pairs is $N(\tau)$. Bilonick (1985) and Stein et al. (1998) are examples of application of the space-time variogram while examples of application of the space-time covariance function include Bennett (1975) for the analysis of population diffusion, Bell (1987) for the study of rainfall, Douaik (2005) and Douaik et al. (2004a) for the analysis of soil salinity.

3.4 Separate Spatial Random Field model

To each of the n_t time instants corresponds a spatial random field (SRF) $X(s, t_j)$, j=1, … , n_t. The spatial variogram, corresponding to each of the above SRF, is defined as:

$$\gamma_j(h) = \frac{1}{2}\mathrm{var}\left[X(s,t_j) - X(s+h,t_j)\right] = \frac{1}{2}E\left\{\left[X(s,t_j) - X(s+h,t_j)\right]^2\right\} \tag{22}$$

It is estimated by:

$$\hat{\gamma}_j(h) = \frac{1}{2N_j(h)} \sum_{i=1}^{N_j(h)} \left[x(s_i, t_j) - x(s_i + h, t_j) \right]^2 \tag{23}$$

with $N_j(h)$ is the number of pairs $[x(s_i, t_j), x(s_i + h, t_j)]$ of data separated by a spatial distance h, at the time instant t_j.

The spatial variogram (equation 22) is a special case of the space-time variogram (equation 20), i.e. when the temporal lag is zero meaning the same time instant:

$$\gamma_j(\mathbf{h}) = \gamma_X(\mathbf{h}, 0) \tag{24}$$

The spatial variograms include the spatial dependence but the dependence in time is not accounted for. It is useful in the case of rich data in space and scarce data in time. The spatial variograms can be one, two or three dimensional.

The above procedure was used, for example, by Ettema et al. (1998) and Sterk & Stein (1997), among others. Many applications of separate spatial variograms, one for each time instant, can be found in the soil science literature. Among others, soil carbon (Chevallier et al., 2000; Van Meirvenne et al., 1996), soil salinity and sodicity (Cetin & Kirda, 2003; Mostafa & Yomota, 1998), soil water (Schume et al., 2003; Wendroth et al., 1999), nitrates (Bruckler et al, 1997), soil biological activity (Gorres, 1997), soil strength (Castrignanò et al., 2002), and soil fertility (Delcourt et al, 1996; Shi et al., 2002) were analyzed for their spatial and temporal change.

3.5 Separate Temporal Random Field model

In a similar way to SRF, for a fixed space location s_i, i=1, ..., n_s, corresponds the independent temporal random fields $X(s_i, t)$.

The temporal variogram at the location s_i, i=1, ..., n_s, is:

$$\gamma_i(\tau) = \frac{1}{2}\text{var}\left[X(s_i, t) - X(s_i, t + \tau)\right] = \frac{1}{2}E\left\{\left[X(s_i, t) - X(s_i, t + \tau)\right]^2\right\} \tag{25}$$

It is estimated as:

$$\hat{\gamma}_i(\tau) = \frac{1}{2N_i(\tau)} \sum_{j=1}^{N_i(\tau)} \left[x(s_i, t_j) - x(s_i, t_j + \tau) \right]^2 \tag{26}$$

with $N_i(\tau)$ is the number of pairs $[x(s_i, t_j), x(s_i, t_j + \tau)]$ of data separated by a temporal distance (time period) τ, at the space location s_i.

Similarly to the spatial variogram, the temporal variogram (equation 25) is a special case of the space-time variogram (equation 20) when the spatial lag is zero meaning the same spatial location:

$$\gamma_i(\tau) = \gamma_X(\mathbf{0}, \tau) \tag{27}$$

The temporal dependence is included whereas the spatial dependence is not considered. Thus, it is useful if many data are collected in time and few in space. The temporal

variograms are one dimensional. Application of temporal variograms in the soil literature is scarce and Cichota et al (2006) and Franzluebbers et al (2002) are examples of such application.

3.6 Space-time, spatial, and temporal cross variograms

The space-time cross variogram is defined, for two soil properties X and Y, as:

$$\gamma_{XY}(h,\tau) = \frac{1}{2}E\{[Y(s+h,t+\tau) - Y(s,\tau)][X(s+h,t+\tau) - X(s,\tau)]\} \tag{28}$$

It is estimated by:

$$\hat{\gamma}_{XY}(h,\tau) = \frac{1}{2N(h,\tau)}\sum_{i,j=1}^{N(h,\tau)}\left[x(s_i,t_j) - y(s_i+h,t_j+\tau)\right]^2 \tag{29}$$

In the case of two SRF observed at two time instants t_j and $t_{j'}$, the spatial cross variogram is defined by:

$$\gamma_{jj'}(h) = \frac{1}{2}E\left\{\left[X(s,t_j) - X(s+h,t_j)\right]\left[X(s,t_{j'}) - X(s+h,t_{j'})\right]\right\} \tag{30}$$

It is estimated by:

$$\hat{\gamma}_{jj'}(h) = \frac{1}{2N_\tau(h)}\sum_{i=1}^{N_\tau(h)}\left[x(s_i,t_j) - x(s_i+h,t_j)\right]\left[x(s_i,t_{j'}) - x(s_i+h,t_{j'})\right] \tag{31}$$

with $N_\tau(h)$ is defined as $N(h,\tau)$ for a particular fixed temporal lag τ.

By analogy with the spatial cross variogram, the temporal cross variogram for two TRF observed at two spatial locations s_i and $s_{i'}$ is defined as:

$$\gamma_{ii'}(\tau) = \frac{1}{2}E\{[X(s_i,t) - X(s_i,t+\tau)][X(s_{i'},t) - X(s_{i'},t+\tau)]\} \tag{32}$$

It is estimated by:

$$\hat{\gamma}_{ii'}(\tau) = \frac{1}{2N_h(\tau)}\sum_{j=1}^{N_h(\tau)}\left[x(s_i,t_j) - x(s_i,t_j+\tau)\right]\left[x(s_{i'},t_j) - x(s_{i'},t_j+\tau)\right] \tag{33}$$

with $N_h(\tau)$ is defined as $N(h,\tau)$ for a particular fixed spatial lag h.

Bishop & Lark (2006), Maas et al. (2010), and Motaghian & Mohammadi (2011) are examples of application of cross-variogram in the soil science. The definition of the space-time, the spatial, and the temporal cross-variograms requires that data are observed at the same space locations and time instants, at the same space locations for two different time instants or at the same time instants for two different space locations, respectively. If it is not the case, the space-time, the spatial or the temporal pseudo cross-variogram (Papritz et al., 1993) needs to be used. Vanderlinden et al. (2006) and Zhang et al. (1999) are examples of application of pseudo cross-variogram and cokriging in soil science literature.

4. Geostatistics: Prediction

When a theoretical model is fitted to the experimental (observed) space-time covariance function or variogram, it becomes possible to tackle the problem of predicting attributes at unsampled space locations and/or time instants. Consider the problem of predicting the value of a continuous attribute x at any unsampled space location s_0 and time instant t_0, $x(s_0, t_0)$, using only the x-data available over the space region S and time domain T, say, n data $\{x(s_{ij}, t_j), j = 1, ..., n_t; i = 1, ..., n_j\}$.

Kriging is a family of generalized least squares regression algorithms that allows to predict $x(s_0, t_0)$. The latter is considered, in geostatistics, as a realization of the STRF $X(s_0, t_0)$. The problem of predicting $x(s_0, t_0)$ is defined as (Rouhani & Myers, 1990):

$$\hat{X}(s_0, t_0) = \sum_j \sum_i \lambda_{ij} X(s_{ij}, t_j) \tag{34}$$

with λ_{ij}, the kriging weights for the time instant t_j and the space location s_{ij}.

This is the general form and there are some simplifications which are related to the kind of STRF considered. We present first the simple cases where we consider only SRF or TRF, and then we present the more general case for the single STRF.

4.1 Spatial or temporal kriging

In case of spatial or temporal random field (sections 3.4 and 3.5) we saw that the dependence in data is modelled via a spatial or temporal covariance function or variogram. For example, in the case of an SRF, we are dealing in the space domain and the different kriging algorithms defined for SRF can be applied. Let $X(s, t_0)$ denotes the SRF at the time instant t_0. The predictor $X(s_0, t_0)$, at the space location s_0, is a linear combination of the n_s observations $x_1, ..., x_{ns}$ available at the time instant t_0:

$$\hat{X}(s_0, t_0) = \sum_{i=1}^{n_s} \lambda_i X(s_i, t_0) \tag{35}$$

The n_s weights are calculated such that $X(s_0, t_0)$ is unbiased and that the variance of the prediction error is minimal. A detailed description of the computation procedure can be found in geostatistical handbooks (Cressie, 1993; Goovaerts, 1997).

Three kriging algorithms can be distinguished (Goovaerts, 1997):

- Simple kriging: the mean is known and constant thorough the space region;
- Ordinary kriging: the mean is unknown but constant in the sub-domains of the space region;
- Kriging with trend model or universal kriging: the mean is unknown, is local and varies smoothly in the sub-domains of the space region.

In a similar way, the predictor $X(s_0, t_0)$, at the time instant t_0, is a linear combination of the n_t observations $x_1, ..., x_{nt}$ available at the space location s_0.

The prediction is done for each time instant (space location) separately and independently from the other time instants (space locations). Many applications can be found in the soil science literature: ordinary kriging using variograms converted from a basic pooled spatial variogram on the sampling time instants (Sterk & Stein, 1997; Stein & Sterk, 1999); ordinary

kriging and universal kriging (Ettema et al., 1998; Figueira et al., 2001; Liu et al, 2009); and ordinary log-normal kriging (Ettema et al., 2000).

4.2 Spatial or temporal cokriging

For the same cases as before (spatial or temporal random fields) more accurate estimates, based on cokriging, can be obtained by using the information available at a precedent time instant as well as the contemporaneous data for spatial prediction (D'Agostino et al., 1998) or information from two different TRFs observed at two different space locations. Cokriging can also be used to estimate the temporal change for a spatially correlated soil property between two time instants using the pseudo cross-variogram (Lark, 2002; Papritz & Fluhler, 1994).

Bogaert (1996) showed that the simple cokriging system between n_t coregionalized space variables (one for each time instant) is equivalent to the simple space-time kriging. This equivalence was shown for a more general form of the covariance functions (Kyriakidis & Journel, 1999). Bogaert (1996) demonstrated also that space ordinary cokriging system with one unbiasedness condition is equivalent to the space-time ordinary kriging system when it is expressed in terms of covariance functions. Pseudo cross-variograms should replace the cross-variograms if the cokriging system has to be written in terms of variograms. However, the use of the unique unbiasedness constraint is restricted to the second order stationary situations. The author noted also that space-time ordinary kriging is preferable to space ordinary cokriging whenever it is possible because, although the differences in the prediction variances are negligible when data are abundant, it is not the case when some time instants involve limited data.

We recall that for SRF, the prediction can be done at unsampled locations only for one or more of the observed time instants. As examples of application, we cite Bogaert & Christakos (1997) and D'Agostino et al. (1998). In the case of TRF, cokriging allows forecasting and hindcasting, the latter is useful in the estimation of missing values. Examples of application can be found in Rouhani et al. (1992a, b). Cokriging was used also to estimate a missing observation at a given time instant by considering the observations at this time instant as the variable of interest and the observations at other time instants as covariables (Stein, 1998).

Let a given soil property be sampled at n_1 space locations during the first time instant, $X_1(s_i)$, $i = 1, ..., n_1$, and at n_2 (less than n_1) space locations later during the second time instant, $X_2(s_j)$, $j = 1, ..., n_2$. The kriging estimator of X_2 at an unsampled location s_0 is (Goovaerts, 1997):

$$\hat{X}_2(s_0) = \sum_{i=1}^{n_1} \lambda 1_i X_1(s_i) + \sum_{j=1}^{n_2} \lambda 2_j X_2(s_j) \tag{36}$$

with $\lambda 1_i$ and $\lambda 2_j$ are the kriging weights relative to the variables X_1 and X_2, respectively. The weights must satisfy the following conditions:

$$\sum_{i=1}^{n_1} \lambda 1_i = 0 \quad \text{and} \quad \sum_{j=1}^{n_2} \lambda 2_j = 1 \tag{37}$$

4.3 Space-time kriging
For the more general single STRF, two kriging algorithms can be distinguished, i.e. two step space-time kriging and anisotropic space-time kriging. They are presented in the next sections followed by other forms of space-time kriging.

4.3.1 Two-step space-time kriging
The prediction of the value of $X_0 = X(s_0, t_0)$ is a linear combination of n_t predictors in space. In the first step, at each of the n_t time instants t_j, $j = 1, …, n_t$, the predictor of the value of $X(s_0, t_j)$ is determined as (Stein et al., 1994):

$$\hat{X}(s_0, t_j) = \sum_{i=1}^{n_j} \lambda_{ij} X(s_{ij}, t_j) \tag{38}$$

The kriging weights are determined in the standard manner for kriging in the space domain (Goovaerts, 1997). In the second step $X(s_0, t_0)$ is predicted, using a linear combination of the predictors determined at the first step. This predictor is (Stein et al., 1994):

$$\hat{X}(s_0, t_0) = \sum_{j=1}^{n_t} \theta_j \hat{X}(s_0, t_j) = \sum_{j=1}^{n_t} \sum_{i=1}^{n_j} \theta_j \lambda_{ij} X(s_{ij}, t_j) \tag{39}$$

with θ_j are the kriging weights from the second step. The predictor $\hat{X}(s_0, t_0)$ explicitly uses spatial variograms that change over time. This algorithm was applied in different situations (Douaik, 2005; Stein et al., 1994, 1998).

4.3.2 Anisotropic space-time kriging
For this algorithm, the prediction is done in one and unique step capitalizing on a space-time variogram model which incorporates the spatial and temporal dependencies with a space-time anisotropy ratio (Bechini et al., 2000; Rouhani & Myers, 1990; Stein et al., 1994, 1998). The advantage of the anisotropic over the two step space-time kriging is that the former allows predictions to be made at every point in space and time. Also it avoids the uncertainties related to the predictors determined at the first step for the latter.

4.4 Space-time cokriging
De Iaco et al. (2004) extended the Linear Model of Coregionalization (LMC) to the space-time domain using generalized product sum models for the variograms. In addition, they proposed an extension of ordinary cokriging to the space-time domain. Although De Iaco et al. (2004) discussed the more general case of several secondary variables; we present here only the case of one variable of interest and one covariable. A linear space-time predictor of $X(s, t)$ at an unsampled point (s_0, t_0) in the space-time domain is:

$$\hat{X}(s_0, t_0) = \sum_{i=1}^{n} \Lambda_i(s_0, t_0) X(s_i, t_i) \tag{40}$$

where $\Lambda_i(s_0, t_0)$ are the (2x2) matrices of kriging weights whose elements $\lambda_i^{\alpha\beta}(s, t)$ are the weights assigned to the value of the βth variable, $\beta = 1, 2$, at the ith sampled point to predict

the αth variable, $\alpha = 1, 2$, at the unsampled point (s_0, t_0). The predicted value includes two components:

$$\hat{X}(s_0,t_0) = \left[\hat{X}_1(s_0,t_0), \hat{X}_2(s_0,t_0) \right] \tag{41}$$

such that

$$\hat{X}_\alpha(s_0,t_0) = \sum_{i=1}^{n}\sum_{\beta=1}^{2} \lambda_i^{\alpha\beta}(s_0,t_0)X_\beta(s_i,t_i) \ , \ \alpha = 1, 2 \tag{42}$$

This corresponds to an extension of ordinary cokriging to the space-time domain. More detail on how to determine the kriging weights $\lambda_i^{\alpha\beta}(s,t)$ can be found in De Iaco et al. (2004). In their work, they compared space-time ordinary kriging to space-time ordinary cokriging. They found that the correlation coefficients between observed and predicted values were higher when the latter algorithm was used. Also, the cokriging error variances were lower than those from kriging.

4.5 Other forms of space-time kriging

Additional forms of space-time kriging exist in the literature. Among these variants, indicator kriging (Bilonick, 1988), kriging with external drift (Snepvangers et al., 2003), and factorial kriging (De Iaco et al., 2003; Rouhani et al., 1992a, b; Van Meirvenne & Goovaerts, 2002) were used.
Now after that the classical geostatistical techniques of space-time interpolation have been laid out, we present a versatile and more general method of interpolation, i.e., the Bayesian Maximum Entropy (BME).

5. Bayesian maximum entropy

For most of the classical geostatistical methods of interpolation, i.e. kriging, the prediction is based solely on the hard data, accurate observed or measured data. However, other data sources can be useful, mostly, for space locations and/or time instants where hard data are missing and they can improve the accuracy of the predictions. Such data can be different types of soft data (interval, probabilistic, etc.), physical laws, moments of higher order, etc.
Bayesian Maximum Entropy (BME) (Christakos, 2000; Christakos & Li, 1998) is a recent approach developed for the spatio-temporal mapping of natural processes using uncertain information. It offers the flexibility to incorporate various sources of physical knowledge. This incorporation of physical knowledge bases enables global prediction features and the adoption of probability distributions without a need for any assumptions like for example to be Gaussian.

5.1 BME knowledge bases

Christakos (2000) defined a knowledge base as 'a collection of knowledge sources relevant to the problem at hand to be involved by a reasoning process aiming at the solution of the problem'. In the BME framework, the total knowledge (K) available regarding a natural process is considered to be formed from two main bases: the general knowledge (G) and the specificatory knowledge (S). The general knowledge represents the knowledge that one has

about the distribution of the natural variable to be mapped before any specific data. It encompasses physical laws, statistical moments of any order (including the mean and variogram or covariance function), multipoint statistics, etc. It is said to be general because it does not depend on the specific random field realization at hand. The specificatory knowledge includes the data specific to a given experiment or situation. It refers to a particular occurrence of the natural variable at a particular space location and a particular time instant. It is divided into two main categories depending on the accuracy of the data. The first category is called hard data which are exact measurements of the natural process and considered to be error-free. It encompasses accurate measurements obtained from real-time observation devices, computational algorithms, simulation processes, etc. The second category is called soft data. It involves uncertain observations, empirical charts, assessments by experts, etc. They are incomplete or qualitative data linked to opinions, intuition, etc. Different kinds of soft data are available, often in the form of interval domain data or of a probabilistic nature.

To illustrate these notions of hard and soft data, we give here an example. Soil salinity can be assessed by measuring the electrical conductivity of the soil. It is measured either in the laboratory (*ECe*) or in the field (*ECa*). The measurement in the laboratory is done with high precision in standard conditions and is directly related to soil salinity. Thus, it is considered as an accurate measurement of the soil salinity and represents the hard data. The apparent electrical conductivity measured in the field (*ECa*) is only an indirect measurement of soil salinity and still needs to be calibrated with *ECe* values. Thus, based on a calibration model (frequently a regression model), *ECa* values are converted into 'predicted' *ECe* values. As these values are only predictions from a stochastic model, they are entailed with some uncertainty, and are consequently considered as soft data. The predicted *ECe* values and their corresponding standard deviations allow us to define the two kinds of soft data: the confidence intervals are built which give the interval soft data while the two statistic parameters are used to describe Gaussian PDFs which represent our probabilistic soft data.

5.2 The three steps of BME analysis

The BME has a double goal: informativeness (prior information maximization given the general knowledge) and cogency (posterior probability maximization given specificatory knowledge). The BME analysis is done in three main steps of knowledge acquisition, integration and processing: the structural or prior step, the meta-prior step, and the integration or posterior step.

The goal of the prior step is the maximization of the information content considering only the general knowledge before any use of the data, which corresponds to the first goal of the BME. This step assumes an inverse relation between information and probability: the more informative an evaluation of a mapping situation is, the less probable it is to occur. This means that if a theory is general and vague, it includes more alternatives so it is more probable however it is less informative. A random field is completely defined by its multivariate probability distribution function (PDF), which forms the prior PDF. The latter should be derived by means of an estimation process that takes into consideration physical constraints under the form of prior information or knowledge. This information is measured, in the context of BME, using the Shannon's entropy function (Shannon, 1948), thus the E in BME. The expected information is given by the entropy function which needs

to be maximized subject to the physical constraints provided by prior information or knowledge *G*, which justifies the M in the BME acronym:

$$f_G(\boldsymbol{\chi}_{map}) = Z^{-1} \exp\left[\sum_{\alpha=1}^{N_c} \mu_\alpha g_\alpha(\boldsymbol{\chi}_{map})\right] \tag{43}$$

where χ_{map} represents the available data and the one to be predicted, $f_G(\boldsymbol{\chi}_{map})$ is the unknown multivariate PDF associated with the general knowledge *G*, $Z = \exp(-\mu_0)$ is a normalization constant, and μ_α are Lagrange multipliers.

During the meta-prior step, the specificatory knowledge is collected and organized in appropriate quantitative forms that can be easily incorporated in the BME framework. The available data can be divided into two main types: the hard data and the soft data. The hard data are considered exact and accurate measurements of the natural process while the soft data are indirect and inaccurate measures of the variable of interest. The soft data may be in the form of intervals, probability distribution functions, etc.

In the posterior step, the two knowledge bases (*G* and *S*) are integrated. The goal is the maximization of the posterior PDF given the total knowledge *K*, which is the second goal of BME. The prior PDF is updated, by considering the available site-specific knowledge (the data). This updating is performed using a Bayesian conditionalization, thus the B in BME. For interval soft data, the posterior PDF $f_K(\chi_k)$ is:

$$f_K(\chi_k) = (AZ)^{-1} \int_I \exp\left[\sum_{\alpha=1}^{N_c} \mu_\alpha g_\alpha(\boldsymbol{\chi}_{map})\right] d\boldsymbol{\chi}_{soft} \tag{44}$$

and for probabilistic soft data, the posterior PDF $f_K(\chi_k)$ is:

$$f_K(\chi_k) = (AZ)^{-1} \int_I \exp\left[\sum_{\alpha=1}^{N_c} \mu_\alpha g_\alpha(\boldsymbol{\chi}_{map})\right] f_S(\boldsymbol{\chi}_{soft}) d\boldsymbol{\chi}_{soft} \tag{45}$$

with χ_{soft} the soft data and *A* is a normalization constant.

5.3 Some BME estimates

The posterior PDF (equations 44 and 45), which is not limited to the Gaussian type, describes fully the random field at the estimation point. It provides a complete picture of the mapping situation as well as different estimators and their associated estimation uncertainty. The prediction points lie, in general, on a regular grid and the predictions are used to create space-time maps. The different estimates presented in the following sections are specific to a general knowledge involving the first two statistical moments and a specificatory knowledge encompassing hard data as well as interval and probabilistic soft data. From the posterior PDF, different statistical parameters can be derived like the mode, the conditional mean, and the variance of the prediction error. The conditional mean is:

$$\hat{\chi}_{k,mean} = \int f_K(\chi_k) \chi_k d\chi_k \tag{46}$$

while the variance of the prediction error is:

$$\hat{\sigma}_k^{\;2} = \int \left(\chi_k - \hat{\chi}_{k,mean} \right)^2 f_K(\chi_k) d\chi_k \qquad (47)$$

5.4 Kriging as a special case of BME
When the general knowledge is limited to the mean and covariance/variogram functions and the specificatory knowledge is restricted to the only hard data, the BME posterior PDF is Gaussian (the mean and the mode are equal) and the BME mean estimate is equivalent to the kriging (best minimum mean squared error, MMSE) estimate, which is the conditional mean. When the space-time random field is Gaussian, the kriging estimate becomes linear and optimal among all MMSE estimators. Consequently, BME is a more general interpolation approach and kriging is a special case in limiting situations.

5.5 Examples of BME application
The BME approach has been applied in different research fields. Examples of such applications can be found in agricultural sciences (Bogaert & D'Or, 2002; D'Or & Bogaert, 2003; D'Or et al., 2001; Douaik, 2005; Douaik et al., 2004b, 2005, 2009), environmental sciences (Christakos & Serre, 2000; Serre et al., 2001), etc.

5.6 Other space-time interpolation methods
The BME is one of the space-time interpolation methods that allow the incorporation of data with different information support and of different degrees of certainty and the knowledge in the form of physical laws. The data assimilation approach (Swinbank et al., 2003) is another method for such situations. It has two main variants: the Kalman filtering (Kalman, 1960) and the Bayesian data fusion (Fasbender et al., 2009). However, their application in the soil science literature is limited (Bierkens, 2001; Heuvelink et al., 2006), mostly for the latter since it was very recently developed.

6. Discussion and conclusions

This chapter has the main objective of reviewing the mostly used statistical methods for appraising the spatial and temporal variability of soil properties. There is a large panoply of techniques and, as stated in the introduction, the choice of the adequate model will depend on the behaviour of the phenomenon of interest (a priori knowledge), the objective of the analysis, and the nature of the available data.
As a first step, the classical statistical method of coefficient of variation, which ignores the spatial and temporal coordinates, can be used as an exploratory data analysis to get an idea about the variability of data either for a given location, a given time instant or both. The coefficient of variation is considered as an indicator of global variability as opposed to local variability; the latter is identified using, for example, the variogram, a geostatistical tool.
Another classical statistical method, the variance component analysis, in which the spatial and temporal contributions are considered as factors with random effects and the joint space-time contribution as interaction in the ANOVA model, can be used to determine their relative contribution to the total variance. Since this approach is directly related to the coefficient of variation, it is suffering from the same drawbacks as the later.

Still using classical statistical methods, there are opportunities to consider the spatial and temporal dimensions of data using, for example, the concept of temporal stability. It is defined as the time invariant association between spatial location and statistical parameters of soil properties and can be assessed using either the relative differences or the Spearman rank correlation. A good test for the temporal stability is the correlation/regression between soil property values measured at consecutive time frames. A way to check the temporal stability of the spatial pattern is to test for a zero intercept and a unity slope of the regression between two consecutive time instants. If the slope of the regression between values at two consecutive time instants is not significantly different from one, then the regression intercept represents the constant change that has occurred between the two time instants. Based on the significance test of these two regression parameters, four possible scenarios can be distinguished: no net change between the two sampling dates; the mean soil property level has changed and there is a static (uniform) change in the soil property spatial pattern; the mean soil property level has not changed and the change in the soil property spatial pattern is dynamic (non-uniform) but, when averaged, this change is not significant; and the mean soil property level has changed and the change in the soil property spatial pattern is dynamic (non-uniform). The concept of relative differences is based on the same statistical parameters (mean and standard deviation) as the two precedent approaches (coefficient of variation and variance component analysis). This concept is based, like the two precedent approaches, on the assumptions of Gaussian distribution of data and the absence of autocorrelation. Both assumptions must be checked before the use of relative differences in the study of temporal stability. The normality can be checked using the standard statistical tests and the independence can be verified using either a test for autocorrelation like the Moran test or by computing the variogram and determining its range. Normality can be gained via data transformation and independence by the selection of sample locations separated by distances larger than the autocorrelation range. For non normally distributed data, but still non autocorrelated, the Spearman rank correlation can be used. It refers to the tendency of a soil property, measured at different locations in space, to maintain their relative ranking over time. In the existence of autocorrelation, its use is not recommended and it will be cautious to use statistical methods which consider the spatial and temporal nature of the given soil property. In addition, even for independent data, the Spearman rank correlation inform us only about the relative ranking of space locations between two time instants and do not provide a quantification of the temporal change. This quantification can be obtained using the paired-*t* test for comparing mean values. Thus, the paired-*t* test, still a classical statistical method, can be used following a temporal stability analysis based on the concept of relative differences.

The concept of temporal stability handles each soil property separately and quantifies the temporal change even without testing if this change is significant or not. Dynamic spatial variation approach can test for change in the mean value (mean shift) as well as for uniform/non-uniform spatial variation between two time instants (spatial shift). Dynamic spatial variation refers to the spatially variable change of a soil property between two time instants, i.e., the change is not the same between two time instants for each location. Again, like with the relative differences concept, based on whether the mean difference is significantly different from zero, the spatial shift parameter is significantly higher than zero, or both, four possible scenarios can be distinguished: no net change with time in the soil property; the mean soil property level has changed and there is a static (uniform) change in

the soil property spatial pattern; the mean soil property level has not changed and the change in the soil property spatial pattern is dynamic (non-uniform), again, when averaged, this change is not significant; and finally, the mean soil property level has changed and the change in the spatial pattern of the soil property is dynamic (non-uniform).

The dynamic spatial variation approach is based on the determination of linear regression between two soil properties; hence, it assumes that the residuals from the regression model are normally distributed with equality of variances and spatially non autocorrelated. All these assumptions must be checked. If spatial autocorrelation is present in the residuals, the estimates of the regression parameters and their variances are biased and the statistical tests are no longer valid. In this special case, spatial regression and/or geostatistical approaches are recommended.

The concept of temporal stability requires a moderate number of measurements in space and time, but once it is checked, the number of future measurements can be reduced to locations with measurements representing important statistics such as the average value. Likewise, based on the dynamic spatial variation test, if it was found that there was a static (uniform) spatial variation between two dates, the sampling effort could be limited to the first soil property survey (an easy to measure and cheap variable) for the first date, while the second soil property (a hard and expensive variable) could be done for two dates but at a reduced number of locations. So the joint use of the concept of temporal stability and the temporal mean shift and spatial shift tests could result in a drastically reduced sampling effort. As illustration, Douaik et al. (2006) found that it was sufficient to measure soil salinity at just two locations and nine dates, from initially 20 locations and 19 time instants, to characterize reliably the average field soil salinity of 25 km² study area. The statistical method to use will depend on the data availability and the aim of the study. If data are available only for one of the two soil properties, there is only one possibility for checking the spatial pattern: the concept of temporal stability. For checking the temporal change in the average soil property, there is a choice between the paired-*t* test and the concept of temporal stability. If, on the other hand, data are available for both soil properties, there are more options: the paired-*t* test and the concept of temporal stability for checking the temporal change in the average level separately for both soil properties or the temporal mean shift test for checking the temporal change in the average level combining both soil properties. Regarding the check of the temporal stability of the spatial patterns, the same choices are available except the paired-*t* test.

For an implicit and formal assessment of the spatial and temporal variability of soil properties, structural analysis is used with the goal of describing and modelling this variability using geostatistical tools like the covariance function, the correlogram, or the variogram. The models of spatial and temporal variability can be used, later, for interpolating between space locations and time instants for mapping soil properties. Basically two interpolating methods can be used for space-time data: the geostatistical method of kriging and the Bayesian maximum entropy.

Three main approaches for the analysis of space-time data can be distinguished:

- Methods based on vectors of spatial random fields (SRF) for the case of many space locations and few time instants. It does not include the temporal dependence existing between observations and can predict only at the observed time instants. The spatial variability is modelled either through a separable spatial variogram for each time instant, or by a single spatial variogram considering time instants as replicates;

- Methods based on vectors of time series (temporal random fields, TRF) for the case of long time series with few space locations. This approach doesn't take into account the spatial dependence and it predicts only at the observed locations. In a similar way as above, independent TRF or TRF as replicates in space can be considered.
- Methods using an STRF, thus considering the joint space-time variability.

For the first approach, the focus of the analysis is on smooth interpolated soil attribute maps over specific time instants. The intention is to capture single instantaneous snapshots (a static picture) of the natural process. The objective can be also the comparison of the various maps or the detection of the temporal persistence or change in the spatial patterns. The intention of the second approach is to capture a sequence of successive snapshots at single space locations, which give temporal profiles. Only the third approach includes both the spatial and temporal dependencies so the interpolation is more precise and can be done for unsampled time instants at unsampled space locations. The focus is here on video sequence of successive spatial pictures like a movie.

When a theoretical model is fitted to the experimental (observed) space-time variogram, it becomes possible to tackle the problem of predicting attributes at unsampled space locations and time instants. Kriging, a family of generalized least squares regression algorithms, is one of the methods allowing these predictions. In case of spatial (temporal) random field, the dependence in data is modelled via a spatial (temporal) variogram. The prediction is done for each time instant (space location) separately and independently from the other time instants (space locations). More accurate estimates, based on cokriging, can be obtained by using the information available at a precedent time instant as well as the contemporaneous data for spatial prediction or information from two different TRFs observed at two different space locations. Cokriging can also be used to estimate the temporal change for a spatially correlated soil property between two time instants using the pseudo cross variogram. Space-time kriging is preferable to spatial cokriging whenever it is possible because, although the differences in the prediction variances are negligible when data are abundant, it is not the case when some time instants involve limited data. For the more general single STRF, two kriging algorithms can be distinguished: two-step space-time kriging and anisotropic space-time kriging. For the latter algorithm, the prediction is done in one and unique step capitalizing on a space-time variogram model which incorporates the spatial and temporal dependencies with a space-time anisotropy ratio. The advantage of the anisotropic over the two-step space-time kriging is that the former allows predictions to be made at every point in space and time. Also it avoids the uncertainties related to the predictors determined at the first step for the latter. In the case of the availability of space-time data for more than one soil property, space-time cokriging can be used for improving the interpolation of the soil property of interest by considering its spatial and temporal dependence with one or more auxiliary soil properties. Additional forms of space-time kriging exist in the literature. Among these variants, indicator kriging, kriging with external drift, and factorial kriging were used in soil science.

The prediction for most of the geostatistical methods of interpolation, i.e. kriging, is based solely on the accurate observed or measured data called hard data. However, other data sources can be useful, mostly, for space locations and/or time instants where hard data are missing and they can improve the accuracy of the predictions. Such data can be different types of uncertain data (soft data), physical laws, statistical moments of higher order, etc. Bayesian Maximum Entropy (BME) is a recent approach developed for the spatio-temporal

mapping of natural processes using uncertain information in addition to accurate measured data. It offers the flexibility to incorporate various sources of physical knowledge enabling global prediction features and the adoption of probability distributions without a need for any assumptions like for example to be Gaussian. The BME has a double goal of informativeness (prior information maximization given the general knowledge) and cogency (posterior probability maximization given specificatory knowledge or the data at hand). The BME probability distribution function describes fully the random field at the prediction point. It provides a complete picture of the mapping situation and different statistical parameters can be derived like the mode, the conditional mean, and the variance of the prediction error. The prediction points lie, in general, on a regular grid and the predictions are used to create space-time maps. In limiting situations (when the general knowledge is limited to the mean and variogram and the specificatory knowledge is restricted to the only hard data), the BME probability distribution function is Gaussian and the BME mean estimate is equivalent to the kriging estimate. Consequently, BME is a versatile and more general method of interpolation and kriging is a special case in limiting situations.

Besides BME, there are other space-time interpolation methods allowing the incorporation of data with different information support and of different degrees of certainty and the knowledge in the form of physical laws. The general data assimilation approach is one of these methods, with its two main variants: the Kalman filtering and the Bayesian data fusion. However, their application in the soil science literature is limited, mostly for the latter since it was very recently developed.

In conclusion, different statistical methods are presently available for dealing with spatial and temporal variability of soil properties and others may be developed in the future. The behaviour of the phenomenon of interest (a priori knowledge), the objective of the research work, and the nature of the available data should help the practitioner in the choice of the adequate method(s).

7. Acknowledgements

8. References

Bechini, L.; Ducco, G.; Donatelli, M. & Stein, A. (2000). Modelling, interpolation, and stochastic simulation in space and time of global solar radiation. *Agriculture, Ecosystems and Environment*, Vol. 81, No. 1, pp. 29–42, ISSN 0167-8809.

Bell, T.L. (1987). A space-time stochastic model of rainfall for satellite remote-sensing studies. *Journal of Geophysical Research*, Vol. 92, No. D8, pp. 9631–9643, ISSN 0148–0227.

Bennett, R.J. (1975). The representation and identification of spatio-temporal systems: an example of population diffusion in North-West England. *Transactions of the Institute of the British Geographers*, Vol. 66, No. 1, pp. 73–94, ISSN 0020-2754.

Bierkens, M.F.P. (2001). Spatio-temporal modelling of the soil water balance using a stochastic model and soil profile descriptions. *Geoderma*, Vol. 103, No. 1, pp. 27–50, ISSN 0016-7061.

Bilonick, R.A. (1985). The space-time distribution of sulphate deposition in the Northern US. *Atmospheric Environment*, Vol. 19, No. 11, pp. 1829–1845, ISSN 0004-6981.

Bilonick, R.A. (1988). Monthly hydrogen ion deposition maps for the North Eastern US from July 1982 to September 1984. *Atmospheric Environment*, Vol. 22, No. 9, pp. 1909–1924, ISSN 0004-6981.

Bishop, T.F.A. & Lark, R.M. (2006). The geostatistical analysis of experiments at the landscape-scale. *Geoderma*, Vol. 133, No. 1-2, pp. 87–106, ISSN 0016-7061.

Bogaert, P. (1996). Comparison of kriging in a space-time context. *Mathematical Geology*, Vol. 28, No. 1, pp. 73–86, ISSN 0882-8121.

Bogaert, P. & Christakos, G. (1997). Stochastic analysis of spatio-temporal solute content measurements using a regression model. *Stochastic Hydrology and Hydraulics*, Vol. 11, No. 4, pp. 267–295, ISSN 0931-1955.

Bogaert, P. & D'Or, D. (2002). Estimating soil properties from thematic soil maps: the Bayesian Maximum Entropy. *Soil Science Society of America Journal*, Vol. 66, No. 5, pp. 1492–1500, ISSN 0361-5995.

Bruckler, L.; De Cockborne, A.M.; Renault, P. & Claudot, P. (1997). Spatial and temporal variability of nitrate in irrigated salad crops. *Irrigation Science*, Vol. 17, No. 2, pp. 53–61, ISSN 0342-7188.

Campbell, C.G.; Ghodrati, M. & Garrido, F. (2001). Temporal consistency of solute transport in a heterogeneous field plot. *Soil Science*, Vol. 166, No. 8, pp. 491–506, ISSN 0038-075X.

Campbell, D.J.; Kinnigurgh, D.G. & Beckett, P.H.T. (1989). The soil solution chemistry of some Oxfordshire soils: temporal and spatial variability. *Journal of Soil Science*, Vol. 40, No. 2, pp. 321–339, ISSN 0022-4588.

Carter, M.R. & Pearen, J.R. (1985). General and spatial variability of Solonetzic soils in north central Alberta. *Canadian Journal of Soil Science*, Vol. 65, No. 1, pp. 157–167, ISSN 0008-4271.

Castrignanò, A.; Maiorana, M.; Fornaro, F. & Lopez, N. (2002). 3D spatial variability of soil strength and its change over time in a durum wheat field in southern Italy. *Soil and Tillage Research*, Vol. 65, No. 1, pp. 95–108, ISSN 0167-1987.

Cetin, M. & Kirda, C. (2003). Spatial and temporal changes of soil salinity in a cotton field irrigated with low-quality water. *Journal of Hydrology*, Vol. 272, No. 1-4, pp. 238–249, ISSN 0022-1694.

Chevallier, T.; Voltz, M.; Blanchart, E.; Chotte, J.L.; Eschenbrenner, V.; Mahieu, M. & Albrecht, A. (2000). Spatial and temporal changes of soil carbon after establishment of a pasture on a long-term cultivated vertisol (Martinique). *Geoderma*, Vol. 94, No. 1, pp. 43–58, ISSN 0016-7061.

Christakos, G. (2000). *Modern spatiotemporal geostatistics*, Oxford University Press, ISBN 978-019-5138-95-5, New York, USA.

Christakos, G.; Bogaert, P. & Serre, M.L. (2002). *Temporal GIS: advanced functions for field-based applications*, Springer-Verlag, ISBN 978-354-0414-76-6, New York, USA.

Christakos, G. & Li, X. (1998). Bayesian maximum entropy analysis and mapping: a farewell to kriging estimators. *Mathematical Geology*, Vol. 30, No. 4, pp. 435–462, ISSN 0882-8121.

Christakos, G. & Raghu, V.R. (1996). Dynamic stochastic estimation of physical variables. *Mathematical Geology*, Vol. 28, No. 3, pp. 341–365, ISSN 0882-8121.

Christakos, G. & Serre, M.L. (2000). BME analysis of spatiotemporal particulate matter distributions in North Carolina. *Atmospheric Environment*, Vol. 34, No. 20, pp. 3393–3406, ISSN 0004-6981.

Cichota, R.; Hortado, A.L.B. & Van Lier, Q.D.J. (2006). Spatio-temporal variability of soil water tension in a tropical soil in Brazil. *Geoderma*, Vol. 133, No. 3-4, pp. 231–243, ISSN 0016-7061

Cressie, N. (1993). *Statistics for spatial data*, Wiley, ISBN 978-047-1002-55-0, New York, USA.

D'Agostino, V.; Greene, E.A.; Passarella, G. & Vurro, M. (1998). Spatial and temporal study of nitrate concentration in groundwater by means of coregionalisation. *Environmental Geology*, Vol. 36, No. 3-4, pp. 285–295, ISSN 0943-0105.

D'Or, D.; Bogaert, P. & Christakos, G. (2001). Application of the BME approach to soil texture mapping. *Stochastic Environmental Research and Risk Assessment*, Vol. 15, No. 1, pp. 87–100, ISSN 1436-3240.

D'Or, D. & Bogaert, P. (2003). Continuous-valued map reconstruction with the Bayesian Maximum Entropy. *Geoderma*, Vol. 112, No. 3-4, pp. 169–178, ISSN 0016-7061.

Da Silva, A.P.; Nadler, A. & Kay, B.D. (2001). Factors contributing to temporal stability in spatial patterns of water content in the tillage zone. *Soil and Tillage Research*, Vol. 58, No. 3-4, pp. 207–218, ISSN 0167-1987.

De Iaco, S.; Myers, D.E. & Posa, D. (2003). The linear coregionalization model and the product sum space-time variogram. *Mathematical Geology*, Vol. 35, No. 1, pp. 25–38, ISSN 0882-8121.

De Iaco, S.; Palma, M. & Posa, D. (2004). Modelling and prediction of multivariate space-time random fields. *Computational Statistics and Data Analysis*, Vol. 48, No. 3, pp. 525–547, ISSN 0167-9473.

Delcourt, H.; Darius, P.L. & De Baerdemaker, J. (1996). The spatial variability of some aspects of topsoil fertility in two Belgian fields. *Computers and Electronics in Agriculture*, Vol. 14, No. 2-3, pp. 179–196, ISSN 0168-1699.

Douaik, A. (2005). *Evaluation of the space-time variability of soil salinity by statistical, geostatistical and Bayesian maximum entropy methods*. PhD Dissertation, Ghent University, Ghent, Belgium.

Douaik, A.; Van Meirvenne, M. & Tóth, T. (2004a). Spatio-temporal kriging of soil salinity rescaled from bulk soil electrical conductivity, In: *GeoEnv IV: Geostatistics for Environmental Applications*, Sanchez-Vila, X.; Carrera, J. & Gomez-Hernandez, J. (Eds.), pp. 413-424, Kluwer Academic Publisher, ISBN 978-1-4020-2007-0, Dordrecht, The Netherlands.

Douaik, A.; Van Meirvenne, M.; Tóth, T. & Serre, M. (2004b). Space-time mapping of soil salinity using probabilistic Bayesian maximum entropy. *Stochastic Environmental Research and Risk Assessment*, Vol. 18, No. 4, pp. 219–227, ISSN 1436-3240.

Douaik, A.; Van Meirvenne, M. & Tóth, T. (2005). Soil salinity mapping using spatio-temporal kriging and Bayesian maximum entropy with interval soft data. *Geoderma*, Vol. 128, No. 3-4, pp. 234–248, ISSN 0016-7061.

Douaik, A.; Van Meirvenne, M. & Tóth, T. (2006). Temporal stability of spatial patterns of soil salinity determined from laboratory and field electrolytic conductivity. *Arid Land Research and Management*, Vol. 20, No. 1, pp. 1–13, ISSN 1532-4982.

Douaik, A.; Van Meirvenne, M. & Tóth, T. (2007). Detecting and monitoring temporal changes of spatial patterns of soil salinity using electrode probes. *Soil Science Society of America Journal*, Vol. 71, No. 5, pp. 1629–1635, ISSN 0361-5995.

Douaik, A.; Van Meirvenne, M. & Tóth, T. (2009). Stochastic approaches for space–time modeling and interpolation of soil salinity. In: Metternicht, G. & Zinck, J.A. (eds). *Remote sensing of soil salinization: impact on land management*. CRC Press, Boca Raton, FL. pp: 273 – 290.

Egbert, G.D. & Lettenmaeir D.P. (1986). Stochastic modelling of the space-time structure of atmospheric chemical deposition. *Water Resources Research*, Vol. 22, No. 2, pp. 165–179, ISSN 0043–1397.

Ehrenfeld, J.G.; Han, X.; Parsons, W.F.J. & Zhu, W. (1997). On the nature of environmental gradients: temporal and spatial variability of soils and vegetation in the New Jersey pinelands. *Journal of Ecology*, Vol. 85, No. 6, pp. 785–798, ISSN 0022-0477.

Ettema, C.H.; Coleman, D.C.; Vellidis, G.; Lowrance, R. & Rathbun S.L. (1998). Spatiotemporal distributions of bacterivorous nematodes and soil resources in a restored riparian wetland. *Ecology*, Vol. 79, No. 8, pp. 2721–2734, ISSN 0012-9658.

Farley, R.A. & Fitter A.H. (1999). Temporal and spatial variation in soil resources in a deciduous woodland. *Journal of Ecology*, Vol. 87, No. 4, pp. 688–696, ISSN 0022-0477.

Fasbender, D.; Brasseur, O. & Bogaert, P. (2009). Bayesian data fusion for space-time prediction of air pollutants: the case of NO_2 in Belgium. *Atmospheric Environment*, Vol. 43, No. 30, pp. 4632–4645, ISSN 0004-6981.

Figueira, R.; Sousa, A.J.; Pacheco, A.M.G. & Catarino, F. (2001). Use of secondary information in space-time statistics for biomonitoring studies of saline deposition. *Environmetrics*, Vol. 12, No. 3, pp. 203–217, ISSN 1180-4009.

Franzluebbers, K.; Franzluebbers, A.J. & Jawson, M.D. (2002). Environmental controls on soil and whole-ecosystem respiration from a tallgrass prairie. *Soil Science Society of America Journal*, Vol. 66, No. 1, pp. 254–262, ISSN 0361-5995

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*, Oxford University Press, ISBN 978-019-5115-38-3, New York, USA.

Goovaerts, P. & Chiang C.N. (1993). Temporal persistence of spatial patterns for mineralizable nitrogen and selected soil properties. *Soil Science Society of America Journal*, Vol. 57, No. 2, pp. 372–381, ISSN 0361-5995.

Gorres, J.H.; Dichiaro, M.J.; Lyons, J.B. & Amador, J.A. (1997). Spatial and temporal patterns of soil biological activity in a forest and an old field. *Soil Biology and Biochemistry*, Vol. 30, No. 2, pp. 219–230, ISSN 0038-0717.

Grant, L.; Seyfried, M. & McNamara, J. 2004. Spatial variation and temporal stability of soil water in a snow-dominated, mountain catchment. *Hydrological Processes*, Vol. 18, No. 18, pp. 3493–3511, ISSN 0885-6087.

Guo, D.; Mou, P.; Jones, R.H. & Mitchell, R.J. (2002). Temporal changes in spatial patterns of soil moisture following disturbance: an experimental approach. *Journal of Ecology*, Vol. 90, No. 2, pp. 338–347, ISSN 0022-0477.

Heathman, G.C.; Larose, M.; Cosh, M.H. & Bindlish, R. (2009). Surface and profile soil moisture spatio-temporal analysis during an excessive rainfall period in the Southern Great Plains, USA. *Catena*, Vol. 78, No. 2, pp. 159–169, ISSN 0341-8162.

Heuvelink, G.B.M. & Webster, R. (2001). Modelling soil variation: past, present, and future. *Geoderma*, Vol. 100, No. 3-4, pp. 269–301, ISSN 0016-7061.
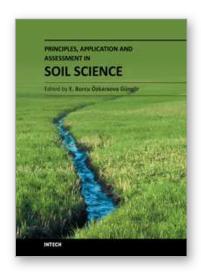
Heuvelink, G.B.M.; Schoorl, J.M.; Veldkamp, A. & Pennock, D.J. (2006). Space-time Kalman filtering of soil redistribution. *Geoderma*, Vol. 133, No. 1-2, pp. 124–137, ISSN 0016-7061.

Hu, W.; Shao, M.; Han, F. & Reichardt, K. (2011). Spatio-temporal variability behavior of land surface soil water content in shrub- and grass-land. *Geoderma*, Vol. 162, No. 3-4, pp. 260–272, ISSN 0016-7061.

James, S.E.; Partel, M.; Wilson, S.D. & Peltzer, D.A. (2003). Temporal heterogeneity of soil moisture in grassland and forest. *Journal of Ecology*, Vol. 91, No. 2, pp. 234–239, ISSN 0022-0477.

Jenny, H. (1980). *The soil resource: origin and behavior* (4th ed.), Springer-Verlag, ISBN 978-038-7905-43-3, New York, USA.

Kachanoski, R.G. & De Jong, E. (1988). Scale dependence and the temporal persistence of spatial patterns of soil water storage. *Water Resources Research*, Vol. 24, No. 1, pp. 85–91, ISSN 0043–1397.

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the American Society of Mechanical Engineers: Journal of Basic Engineering*, Vol. 82, No. D1, pp. 35–45, ISSN 0021-9223.

Kenny, E.A.; Hall, J.W. & Wang, C. 2002. Temporal trends in soil physical properties at a soil quality benchmark site in British Columbia, Canada. *Proceedings of the 17th World Congress of Soil Science*, [CD-ROM], Bangkok, Thailand, August 2002.

Kyriakidis, P.C. & Journel, A.G. (1999). Geostatistical space-time models: a review. *Mathematical Geology*, Vol. 31, No. 6, pp. 651–684, ISSN 0882-8121.

Lark, R.M. (2002). Robust estimation of the pseudo-cross variogram for cokriging soil properties. *European Journal of Soil Science*, Vol. 53, No. 2, pp. 253–270, ISSN 1351-0754.

Lesch, S.M.; Herrero, J. & Rhoades, J.D. (1998). Monitoring for temporal changes in soil salinity using electromagnetic induction techniques. *Soil Science Society of America Journal*, Vol. 62, No. 1, pp. 232–242, ISSN 0361-5995.

Lesch, S.M.; Strauss, D.J. & Rhoades, J.D. (1995). Spatial prediction of soil salinity using electromagnetic induction techniques.1 Statistical prediction models: a comparison of multiple linear regression and cokriging. *Water Resources Research*, Vol. 31, No. 2, pp. 373–386, ISSN 0043–1397.

Liu, X.; Zhang, W.; Zhang, M.; Ficklin, D. & Wang, F. (2009). Spatio-temporal variations of soil nutrients influenced by an altered land tenure system in China. *Geoderma*, Vol. 152, No. 1-2, pp. 23–34, ISSN 0016-7061.

Maas, S.; Scheifler, R.; Benslama, M.; Crini, N.; Lucot, E.; Brahmia, Z.; Benyacoub, S. & Giraudoux P. (2010). Spatial distribution of heavy metal concentrations in urban, suburban and agricultural soils in a Mediterranean city of Algeria. *Environmental Pollution*, Vol. 158, No. 6, pp. 2294–2301, ISSN 0269-7491.

McClave, J.T. & Sincich, T. (2006). *Statistics* (10th ed.), Prentice Hall, ISBN 978-013-1497-55-9, Upper Saddle River, USA.

Mostafa, M.M. & Yomota, A. (1998). Use of a covariance variogram to investigate influence of subsurface drainage on spatial variability of soil-water properties. *Agricultural Water Management*, Vol. 37, No. 1, pp. 1–19, ISSN 0378-3774.

Motaghian, H.R. & Mohammadi, J. (2011). Spatial estimation of saturated hydraulic conductivity from terrain attributes using regression, kriging, and artificial neural networks. *Pedosphere*, Vol. 21, No. 2, pp. 170–177, ISSN 1002-0160.

Papritz, A. & Fluhler, H. (1994). Temporal change of spatially autocorrelated soil properties: optimal estimation by kriging. *Geoderma*, Vol. 62, No. 1-3, pp. 29–43, ISSN 0016-7061.

Papritz, A.; Kunsch, H.R. & Webster, R. (1993). On the pseudo cross-variogram. *Mathematical Geology,* Vol. 25, No. 8, pp. 1015–1026, ISSN 0882-8121.

Petrone, R.M.; Price, J.S.; Carey, S.K. & Waddington, J.M. (2004). Statistical characterization of the spatial variability of soil moisture in a cutover peatland. *Hydrological Processes*, Vol. 18, No. 1, pp. 41–52, ISSN 0885-6087.

Reichardt, K.; Bacchi, O.O.S.; Villagra, M.D.L.M.; Turrati, A.L. & Pedrosa, Z.O. (1993). Hydraulic variability in space and time in a dark red latosol of the tropics. *Geoderma*, Vol. 60, No. 1-4, pp. 159–168, ISSN 0016-7061.

Rodriguez-Iturbe, I. & Eagelson, P.S. (1987). Mathematical models of rainstorm events in space and time. *Water Resources Research*, Vol. 23, No. 1, pp. 181–190, ISSN 0043–1397.

Rouhani, S.; Ebrahimpour, R.M.; Yaqub, I. & Gianella, E. (1992a). Multivariate geostatistical trend detection and network evaluation of space-time acid deposition data. – I. Methodology. *Atmospheric Environment*, Vol. 26, No. 14, pp. 2603–2614, ISSN 0004-6981.

Rouhani, S.; Ebrahimpour, R.M.; Yaqub, I. & Gianella, E. (1992b). Multivariate geostatistical trend detection and network evaluation of space-time acid deposition data. –II. Application to NADP/NTN data. *Atmospheric Environment*, Vol. 26, no. 14, pp. 2615–2626, ISSN 0004-6981.

Rouhani, S. & Myers, D.E. (1990). Problems in space-time kriging of geohydrological data. *Mathematical Geology,* Vol. 22, No. 5, pp. 611–623, ISSN 0882-8121.

Rouhani, S. & Wackernagel, H. (1990). Multivariate geostatistical approach to space-time data analysis. *Water Resources Research*, Vol. 26, No. 4, pp. 585–591, ISSN 0043–1397.

Schume, H.; Jost, G. & Katzensteiner, K. (2003). Spatiotemporal analysis of the soil water content in a mixed Norway spruce (Picea abies (L) Karst.)-European beech (Fagus sylvatica L.) stand. *Geoderma*, Vol. 112, No. 3-4, pp. 273–287, ISSN 0016-7061.

Serre, M.L.; Christakos, G.; Howes, J. & Abdel-Rehiem, A.G. (2001). Powering an Egyptian air quality information system with the BME space-time analysis toolbox: results from the Cairo baseline year study. In: *GeoEnvIII: Geostatistics for Enironmental Applications*, Monestiez, P.; Allard, D. & Froidevaux, R. (eds), pp. 91-101, Kluwer Academic Publisher, ISBN 978-079-2371-06-9, Dordrecht, The Netherlands.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, Vol. 27, No. 2, pp. 379–423, ISSN 0005-8580.

Shi, Z.; Wang, K.; Bailey, J.S.; Jordan, C. & Higgins, A.H. (2002). Temporal changes in the spatial distributions of some soil properties on a temperate grassland site. *Soil Use and Management*, Vol. 18, No. 4, pp. 353–362, ISSN 1475-2743.

Si, B.C. (2002). Spatial and statistical similarities of local soil water fluxes. *Soil Science Society of America Journal*, Vol. 66, No. 3, pp. 753–759, ISSN 0361-5995.

Snepvangers, J.J.J.C.; Heuvelink, G.B.M. & Huisman, J.A. (2003). Soil water content interpolation using spatiotemporal kriging with external drift. *Geoderma*, Vol. 112, No. 3-4, pp. 253–271, ISSN 0016-7061.

Solow, A.R. & Gorelick, S.M. (1986). Estimating monthly streamflow values by cokriging. *Mathematical Geology,* Vol. 18, No. 8, pp. 785–810, ISSN 0882-8121.

Stein, A. (1998). Analysis of space-time variability in agriculture and the environment with geostatistics. *Statistica Neerlandica*, Vol. 52, No. 1, pp. 18–41, ISSN 1467-9574.

Stein, A.; Kocks, C.G.; Zadocks, J.C.; Frinking, H.D.; Ruissen, M.A. & Myers, D.E. (1994). A geostatistical analysis of the spatiotemporal development of downy mildew epidemics in cabbage. *Phytopathology*, Vol. 84, No. 10, pp. 1227–1239, ISSN 0031-949X.

Stein, A. & Sterk. G. (1999). Modelling space and time dependence in environmental studies. *International Journal of Applied Earth Observation and Geoinformation*, Vol. 1, No. 2, pp. 109–121, ISSN 0303-2434.

Stein, A.; Van Groenigen, J.W.; Jeger, M.J. & Hoosbeek, M.R. (1998). Space-time statistics for environmental and agricultural related phenomena. *Environmental and Ecological Statistics*, Vol. 5, No. 2, pp. 155–172, ISSN 1352-8505.

Sterk, G. & Stein, A. (1997). Mapping wind-blown mass transport by modelling variability in space and time. *Soil Science Society of America Journal*, Vol. 61, No. 1, pp. 232–239, ISSN 0361-5995.

Swinbank, R.; Shutyaev, V. & Lahoz, W.A. (Eds.). (2003). *Data Assimilation for the Earth System*, Kluwer Academic Publisher, ISBN 978-140-2015-93-9, Dordrecht, The Netherlands.

Tóth, T.; Kuti, L.; Fórizs, I.; Kabos, S. & Douaik, A. (2002). Spatial and temporal aspects of soil salinization in a sodic grassland. In: *Proceedings of the Interntional Symposium on Sustainable use and management of soils in arid and semiarid regions*, Fez Cano, A.; Ortiz Silla, R. & Mermut, A.R. (eds.), vol. 1, pp. 276–288, Cartagena, Spain.

Vachaud, G.; Passerat De Silans, A. ; Balabanis, P. & Vauclin, M. (1985). Temporal stability of spatially measured soil water probability density function. *Soil Science Society of America Journal*, Vol. 49, No. 4, pp. 822–828, ISSN 0361-5995.

Van Es, H.M. (1993). Evaluation of temporal, spatial, and tillage-induced variability for parametrization of soil infiltration. *Geoderma*, Vol. 60, No. 1-4, pp. 187–199, ISSN 0016-7061.

Van Es, H.M.; Ogden, C.B.; Hill, R.L.; Schindelbeck, R.R. & Tsegaye, T. (1999). Integrated assessment of space, time, and management-related variability of soil hydraulic properties. *Soil Science Society of America Journal*, Vol. 63, No. 6, pp. 1599–1608, ISSN 0361-5995.

Van Meirvenne, M. & Goovaerts, P. (2002). Accounting for spatial dependence in the processing of multi-temporal SAR images using factorial kriging. *International Journal of Remote Sensing*, Vol. 23, No. 2, pp. 371–387, ISSN 0143-1161.

Van Meirvenne, M.; Pannier, J.; Hofman, G. & Louwagie, G. (1996). Regional characterization of the long-term change in soil organic carbon under intensive agriculture. *Soil Use and Management*, Vol. 12, No. 2, pp. 86–94, ISSN 1475-2743.

Van Wesenbeeck, I.J. & Kachanoski, R.G. (1988). Spatial and temporal distribution of soil water in the tilled layer under a corn crop. *Soil Science Society of America Journal*, Vol. 52, No. 2, pp. 363–368, ISSN 0361-5995.

Van Wesenbeeck, I.J.; Kachanoski, R.G. & Rolston, D.E. (1988). Temporal persistence of spatial patterns of soil water content in the tilled layer under a corn crop. *Soil Science Society of America Journal*, Vol. 52, No. 4, pp. 934–941, ISSN 0361-5995.

Vanderlinden, K.; Ordonez, R.; Polo, M.J. & Giraldez, J.V. (2006). Mapping residual pyrite after a mine spill using non co-located spatiotemporal observations. *Journal of Environmental Quality*, Vol. 35, No. 1, pp. 21–36, ISSN 0047-2425.

Wendroth, O.; Pohl, W.; Koszinski, S.; Rogasik, H.; Ritsema, C.J. & Nielsen, D.R. (1999). Spatiotemporal patterns and covariance structures of soil water status in two northeast German field sites. *Journal of Hydrology,* Vol. 215, No. 1-4, pp. 38–58, ISSN 0022-1694.

Zhang, R.; Shouse, P. & Yates, S. (1999). Estimates of soil nitrate distributions using cokriging with pseudo cross-variograms. *Journal of Environmental Quality*, Vol. 28, No. 2, pp. 424–428, ISSN 0047-2425.

Principles, Application and Assessment in Soil Science

Edited by Dr. Burcu E. Ozkaraova Gungor

Our dependence on soil, and our curiosity about it, is leading to the investigation of changes within soil processes. Furthermore, the diversity and dynamics of soil are enabling new discoveries and insights, which help us to understand the variations in soil processes. Consequently, this permits us to take the necessary measures for soil protection, thus promoting soil health. This book aims to provide an up-to-date account of the current state of knowledge in recent practices and assessments in soil science. Moreover, it presents a comprehensive evaluation of the effect of residue/waste application on soil properties and, further, on the mechanism of plant adaptation and plant growth. Interesting examples of simulation using various models dealing with carbon sequestration, ecosystem respiration, and soil landscape, etc. are demonstrated. The book also includes chapters on the analysis of areal data and geostatistics using different assessment methods. More recent developments in analytical techniques used to obtain answers to the various physical mechanisms, chemical, and biological processes in soil are also present.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ahmed Douaik, Marc van Meirvenne and Tibor Tóth (2011). Statistical Methods for the Analysis of Soil Spatial and Temporal Variability, Principles, Application and Assessment in Soil Science, Dr. Burcu E. Ozkaraova Gungor (Ed.), ISBN: 978-953-307-740-6, InTech, Available from: http://www.intechopen.com/books/principles-application-and-assessment-in-soil-science/statistical-methods-for-the-analysis-of-soil-spatial-and-temporal-variability

INTECH

open science | open minds