ELSEVIER

Journal of Hydrology 272 (2003) 226–237

Journal of Hydrology

www.elsevier.com/locate/jhydrol

Sugarcane production evaluated by the state-space approach☆

L.C. Timm[a,d,*], K. Reichardt[b,d], J.C.M. Oliveira[c], F.A.M. Cassaro[d], T.T. Tominaga[d], O.O.S. Bacchi[d], D. Dourado-Neto[e]

[a]*Department of Rural Engineering, ESALQ/USP, CP 9, 13418-970 Piracicaba, SP, Brazil*
[b]*Department of Exact Sciences, ESALQ/USP, Piracicaba, SP, Brazil*
[c]*Municipal University of Piracicaba, EEP, Piracicaba, SP, Brazil*
[d]*Soil Physics Laboratory, CENA/USP, CP 96, 13400-970 Piracicaba, SP, Brazil*
[e]*Department of Crop Production, ESALQ/USP, Piracicaba, SP, Brazil*

Abstract

The effects of soil properties on crop growth and yield have traditionally been analyzed using classical statistics methodologies. These methodologies do, however, not consider sampling position coordinates and assume independence between samples. This study had the objective of using the state-space approach, which considers sampling position, to evaluate and to discuss a spatial process using variables related to the soil–plant system. For this, six data sets were collected in a sugarcane experiment carried out on a Dark Red Latosol (Rhodic Kandiudalf), at Piracicaba, State of São Paulo, Brazil. The sugarcane was planted on a 0.21 ha field, comprising 15 rows, 100 m long, spaced 1.4 m apart, with three treatments (mulching, bare soil and straw burning before harvest) and four replicates, forming a transect of 84 points. In this way, the relationships between the number of canes per meter of row and available soil P, Ca and Mg, clay content and aggregate stability were studied using a first order state-space model. Results show that all of the used state-space equations described the spatial distribution of number of canes better than the equivalent multiple regression equations. It was also identified that the soil clay content spatial series has an effective contribution to describe the number of canes in this study, because it is related to the best performance in each different scenario.

Keywords: State-space approach; Sugarcane harvest; Multiple regression analysis

1. Introduction

The evaluation of land management practices and their impact on environmental quality requires adequate analytical tools and experimental designs. A significant progress to understand crop production has already been made by measuring and analyzing on-site processes (van Kessel and Wendroth, 2001). Traditionally soil scientists have used random sampling techniques, assuming independence between samples, in order to analyze the effect of soil properties on crop growth and yield. Hence, crop development variables and soil attributes collected at spatially different locations relative to each other, in general manifest low correlation when classical statistical analysis is used (Nielsen et al., 1997). According to Coelho et al. (1998), the importance of spatial and temporal variability of soil chemical and physical properties and their relation to crop yield should not be underestimated in planning soil management. Recently, applied analytical techniques

☆ Research funded by FAPESP, CAPES and CNPq.
* Corresponding author. Soil Physics Laboratory, CENA/USP, CP 96, 13400-970 Piracicaba, SP, Brazil.
E-mail address: lctimm@carpa.ciagri.usp.br (L.C. Timm).

0022-1694/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0022-1694(02)00267-6

in agriculture, such as the state-space methodology, have shown to provide opportunities for an on-site analysis and for a suitable identification of spatial relations between crop and soil variables taking into account their spatial association. State-space modeling is a technique that can filter noise underlying crop and soil processes at various scales if the observation density supports the identification of the correlation length. Hence, this technique can be applied to identify landscape-scale processes and generate reliable predictions, having practical advantages, and it can be a more effective research tool in comparison to other approaches to understand and explain landscape-scale variation in agricultural systems (Morkoc et al., 1985; Wendroth et al., 1992; Nielsen et al., 1999; Dourado-Neto et al., 1999; Timm et al., 2000, 2001). Multidimensional spatial processes are frequently found, involving soil variables such as water, temperature, soil salinity, infiltration, and crop yield (Warrick et al., 1986; Shumway et al., 1988; Ahuja and Nielsen, 1990). Li and Lascano (1999) used state-space analysis to describe the spatial correlation between cotton lint yield, soil water, phosphorus, and site elevation. It has also been used to describe adequately the spatial association of wheat grain yield, soil base saturation, and water storage capacity on a newly land-shaped field in North Carolina (Cassel et al., 2000). Nielsen et al. (1999); Wendroth et al. (1999) tested state-space modeling to quantify localized variation and their findings indicated that these models provide an insight into the spatial patterns of crop and soil variables.

Other autoregression techniques, geostatistical analysis, and multiple regression use a single set of relationships among the variables to explain variation across a data series and require that the mean and variance ratio remain constant across space. However, the variation across fields typically tends to be localized, and therefore, the state-space modeling, compared to other more commonly used analyses, can accurately and precisely predict landscape-scale variations (Stevenson et al., 2001). Other reports (Morkoc et al., 1985; Wendroth et al., 1992; Li et al., 2001) have observed that the state-space modeling can be an effective research tool to explain landscape-scale variation in agricultural systems. According to Stevenson et al. (2001), most agricultural scientists would agree that landscape-scale variation is always present and that the ANOVA (analyses of variance) approach has to first detect if landscape-scale variation is present, using statistical significance and traditional experimental designs, before providing any further explanation.

In this study, we want to improve our understanding on the relationships among sugarcane yield parameters, such as the number of canes per meter of row, and physico-chemical soil properties such as available phosphorus, calcium, and magnesium, clay content and aggregate stability, using a state-space approach. With this approach, we want to identify as well as eliminate factors involved in field processes, searching for an optimal management of soil resources and crop yield for sugarcane production in Brazil.

2. Material and methods

The state-space analysis characterizes the state of a system (set of unobservable variables) at a location i to its state at a location $i-h$, $h = 1, 2, 3, \ldots, n$. For $h = 1$, Shumway et al. (1988); Wendroth et al. (1997) and Nielsen et al. (1999) described the state-space approach as follows:

$$Z_p(x_i) = \phi_{pp} Z_p(x_{i-1}) + w_{Zp}(x_i) \tag{1}$$

$Z_p(x_i)$ being the state vector (a set of p unobservable variables) at location x_i; ϕ_{pp} a $p \times p$ matrix of state coefficients, which indicates the measure of the regression; and $w(x_i)$ noises of the system for $i = 1, 2, 3, \ldots, n$. Noise values are assumed to have zero mean, not being autocorrelated and being normally distributed with constant variances. If these Z variables were observable, this would be the usual structure of a vector autoregressive model, in which the coefficients of the matrix ϕ_{pp} could be estimated by multiple regression techniques such as least squares, taking $Z_p(x_i)$ and $Z_p(x_{i-1})$ as the dependent and independent variables, respectively. In the case of the state-space model, however, the true state of the variables is considered 'embedded' in an observation equation:

$$Y_p(x_i) = A_{pp}(x_i) Z_p(x_i) + v_{Yp}(x_i) \tag{2}$$

the observation vector $Y_p(x_i)$ being related to the state vector $Z_p(x_i)$ by an observation matrix $A_{pp}(x_i)$ (usually known as, for instance, an identity matrix, $p \times p$) and an observation noise vector $v(x_i)$, also considered of zero mean, not autocorrelated and normally distributed. Also, the noises $w(x_i)$ and $v(x_i)$ are assumed to be independent of each other, the measurements need not be considered true, but can be seen as indirect measurements reflecting the true state of the variable added to a noise (Wendroth et al., 1997). The state coefficients of the matrix ϕ_{pp} and noise variances of Eq. (1) are estimated through a recursive procedure given by Shumway and Stoffer (1982). They are optimized using the Kalman filter (Kalman, 1960; Gelb, 1974), combined with an iterative algorithm and their magnitude reflects the importance of each variable in their ability to define the system (Stevenson et al., 2001). According to Wendroth et al. (2001), the state-space modeling integrates the influence of local effects since the meaning and impact of the transition matrix coefficients for the state vector can change across space or time, providing more flexibility than a technique, that is limited to a unique response function.

Data used for this analysis belong to a sugarcane experiment started in October 1997, on a Dark Red Latosol (Rhodic Kandiudalf), called locally 'Terra Roxa Estruturada', at Piracicaba, State of Sao Paulo, Brazil (22° 42′ S; 47° 38′ W), 580 m above sea level, 250 km inside the continent. The sugarcane variety SP 80-3280 of medium to late cycle was planted on a 0.21 ha field, with 15 rows, each 100 m long, and a row distance of 1.4 m. Sugarcane was planted along leveled furrows due to a 7.4% slope of the experimental area, a practice that minimizes run-off. A completely randomized block design was used, with three treatments and four replicates distributed on lines 7, 8 and 9, separated by borders, in such a way that each plot had three cane rows of 4 m, totaling a cropped area of 16.8 m^2. Each replicate and each border were divided in strips of 1 m, forming a spatial transect of 84 points, as shown in Fig. 1.

During the first year (1997/1998) no treatments were imposed on the whole field, which was managed homogeneously according to traditional agricultural practices. After the October 1998 harvest, three management treatments were established: (i) mulching the ratoon crop with trash (cane tips and straw from harvest, T_1 and T_2);[1] (ii) bare soil between rows after harvest (T_3); and (iii) soil surface covered by residues left by the traditional practice of straw burning before harvest (T_4). More details about the field experiment can be found in Oliveira et al. (2001).

Sugarcane harvest (October 2000) was made by hand counting the total number of cane plants (NC) in the three central lines along each meter for the 84 point transect, and NC was used in the analysis of this study.

To determine soil available P, Ca and Mg, and clay content and aggregate stability, 1.0 kg samples (0–0.15 m layer) were collected (after harvest October 2000) at the center of each meter, hence, for the same 84 points along the transect. Soil samples were dried and sieved (2 mm) and sub-samples were used to perform laboratory analyses according to Gee and Bauder (1986); Kemper and Rosenau (1986); EMBRAPA (1997).

Data were analyzed using the state-space approach, with the aid of the software, Applied Statistical Time Series Analysis (ASTSA), developed by Shumway (1988). Data $Z_p(x_i)$ were normalized with respect to their mean m and standard deviation s, as follows:

$$z_p(x_i) = [Z_p(x_i) - (m - 2s)]/4s \quad (3)$$

where $z_p(x_i)$ are the normalized values, dimensionless and having a mean of 0.5 and standard deviation of 0.25 (Nielsen et al., 1997; Wendroth et al., 1999). This transformation allows state coefficients of ϕ_{pp} having magnitudes directly proportional to their contribution to each state variable used in the analysis (Hui et al., 1998).

3. Results and discussion

The sugarcane crop is semi-perennial, being renewed every 5 years. It belongs to the grass family, having a bulky rhizome, cane stalks can reach 3 m in length, and a root system that is mostly confined to the 0–0.50 m top layer, with some roots growing more than 1.0 m in depth. It is planted in rows and harvested after 1 year or more. Stalks are used to manufacture sugar and/or alcohol. After each harvest, rhizome sprouts renew the crop, now called ratoon. After 4–7 ratoons, the crop is renewed with new plants, using stalk pieces that germinate.

The data set used in this analysis is from the second ratoon crop (October 1999–October 2000).

[1] T_1 differs from T_2 only in terms of ^{15}N label, a study reported elsewhere.

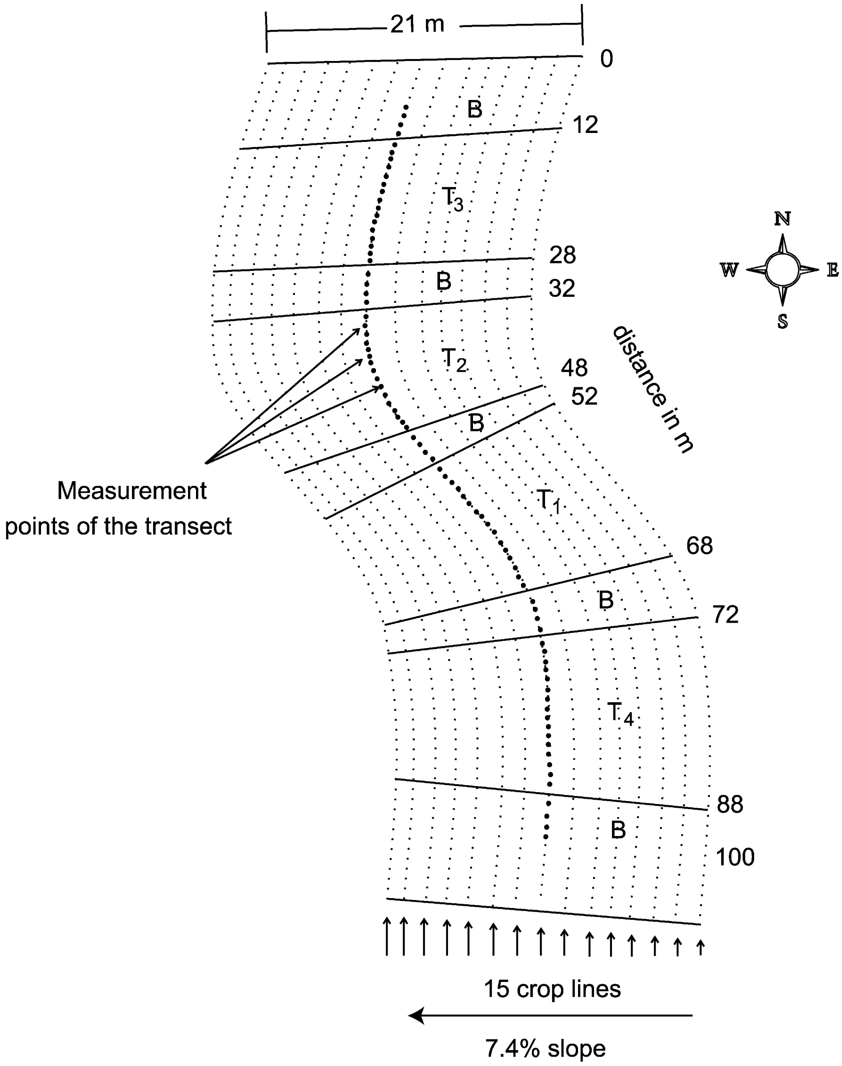

Fig. 1. Schematic experimental design showing the 15 cane lines, each 100 m long, indicating the 3 central lines (7, 8 and 9) used to measure physical and chemical soil properties. B = border; T = treatment; R = replicate.

The behavior of the number of canes (NC), phosphorus (P), calcium (Ca), magnesium (Mg), clay content (CC) and aggregate stability (AS) data, collected along the transect is shown in Fig. 2(A–F), respectively. In Fig. 2A and E, we can note the inverse behavior of the number of canes and the soil clay content along the transect.

In the Southern Hemisphere, the spring–summer period is important for the establishment of the ratoon sugarcane crops. Although being a relatively short period of the crop cycle (60–90 days), it is the period in which the crop rhizome is being renewed and is more sensitive to weather and soil conditions. During this period, high rain amounts were measured (52.1 mm: November 1999; 269.9 mm: December 1999; and 235.9 mm: January 2000). Rains associated to the higher clay content (poor natural drainage) and the effect of soil surface mulching on T_1 and T_2 treatments can explain the lower number of canes per meter in T_1 and T_2 when compared to T_3 and T_4. The establishment of a humid microclimate in the straw layer, which had, initially, a thickness of 0.20–0.30 m could have

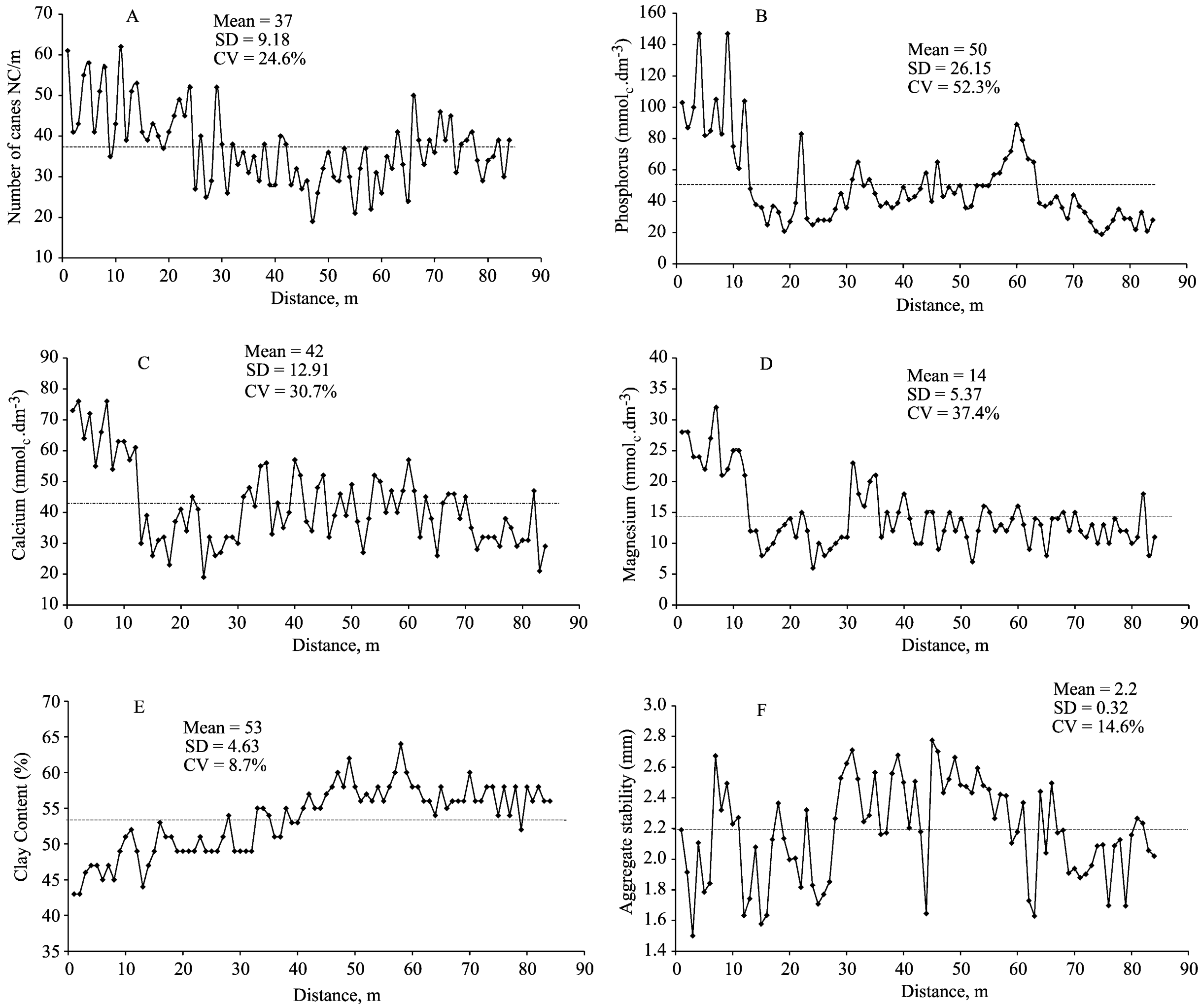

Fig. 2. Data distributions along the transect: (A) number of canes NC; (B) soil phosphorus P; (C) soil calcium Ca; (D) soil magnesium Mg; (E) soil clay content CC; and (F) soil aggregate stability AS.

promoted fungi and microorganism growth that affected rhizome sprouting and stalk development.

Each data set was first analyzed using classical statistics to obtain descriptive parameters such as means, standard deviations and coefficients of variation (CV). From Fig. 2 it can be seen that CV values for all soil and plant variables ranged from 8.7 to 52.3%. The smallest CV was for clay content (Fig. 2E) and the largest for P variation along the 84 point transect (Fig. 2B). All these variables exhibit point-to-point fluctuations due to soil heterogeneity (natural spatial variability), which presents local characteristics and, therefore, may be better represented by a local model (e.g. state-space model). In this case, the use of global or space-independent models (e.g. standard multiple regression) based on the assumption that each data set manifests a constant mean along the entire transect, ignoring its local spatial variation, fails to express the soil spatial variability. Under such conditions, according to Wendroth et al. (1999), local trends may be better considered with nearest neighbor analysis (e.g. state-space modeling).

Traditionally, in most agronomic investigations statistical analysis such as ANOVA or regression, are done to identify yield response to different inputs aiming to describe the changes observed within and among plots. Using these traditional statistical tools the spatial coordinates of the observations are not considered, assuming that the observations are spatially independent of each other, being randomly distributed over the entire field (no correlation structure). According to Nielsen and Alemi (1989), classical statistics, with the applied ANOVA-based experimental design, assumes that observations within and between treatments are independent. Such analyses still lack proven concepts and strategies to interpret the cause of spatial patterns, mainly because response functions between crop yield and different variables are not constant nor consistent across an agricultural field (Nielsen et al., 1999). To understand, how soil physical and chemical properties can affect crop growth and yield, we need to examine their spatial relationships. Several statistical tools, like autocorrelation, semivariogram, crosscorrelation, kriging, cokriging and state-space analysis (Morkoc et al., 1985; Wendroth et al., 1992; Wendroth et al., 1997; Hui et al., 1998; Dourado-Neto et al., 1999) have been used to evaluate the correlation structure of spatial distributions. In our case, we calculated the autocorrelation function (ACF) for each data set, i.e. number of canes (Fig. 3A), P (Fig. 3B), Ca (Fig. 3C), Mg (Fig. 3D), clay content (Fig. 3E) and aggregate stability (Fig. 3F) with the objective of evaluating the spatial correlation of the observations, in other words if they had been monitored at a distance sufficient for identifying their spatial representativity. Using a t test at the 5% probability, the ACF of the number of canes data presented in Fig. 3A manifests significant spatial correlation up to 10 lags. Fig. 3B–D shows that the spatial dependence of the soil P, Ca and Mg data set is significant up to 6 lags. The strong spatial dependence between adjacent observations of clay content data is presented in Fig. 3E. From Fig. 2(A–E) it can be seen that the spatial distributions of number of canes, soil P, Ca, Mg and clay contents each manifest a trend along the transect. This trend causes a relatively strong spatial dependence of each variable as shown by the ACF in Fig. 3(A–E). Unlike the other observations, aggregate stability had no discernible trend, and manifests spatial dependence up to 3 lags (Fig. 3F). Stationarity (no trend) means that the series develops stable along space, with mean, variance and autocovariance constant along the domain of interest. Most of the series in practice present some sort of non-stationarity, such as trends and periodicity. Some models used in time series analysis (e.g. autoregressive AR, moving average MA, autoregressive integrated moving average ARIMA) impose the need of transforming the original data to remove trends and then use standard statistical procedures.

Soil spatial variability can occur at different levels, related to different factors, such as variation of the soil's parent material, climate, relief, organisms and time, i.e. related to the processes of soil formation and/or effects of management practices adopted for each of its agricultural use (McGraw, 1994). Since soil properties show spatial variability, it is important to take into account the uncertainty in model input parameters, when the behavior of soils is simulated. Another important aspect is related to the scale in which the data were collected. Such scale effect, in a statistical sense, may be referred to as the spatial correlation length or integral scale of the measurement, property or process (Dagan, 1986). In principle, it is possible to run the model in a large scale and then to evaluate its performance in a smaller scale. As pointed out by Cushman (1990), spatial patterns of soil properties, within and between scales, might be different from the organization of the soil hydrological processes across spatial scales. As the physical process moves towards a larger spatial scale, soil properties may change from deterministic to random, with the smaller-scale variations filtered out by the larger-scale process, thereby eliminating non-stationary trends at the smaller spatial scales.

According to Journel and Huijbregts (1991), the autocorrelation function is a tool that reflects the local variation between samples for different separation distances, being used to identify the range of the spatial correlation of the observations of a variable. If physical, chemical and biological phenomena are being observed in a form that leads to a numerical quantification in a sequence, distributed in space, and a set of observations present a spatial dependence such that a range of correlation does exist, we can take advantage of this spatial dependence making use of the location of each observation to better understand the different processes of our study.

Ignoring the locations of the observations, we perform a classical linear and multiple regression

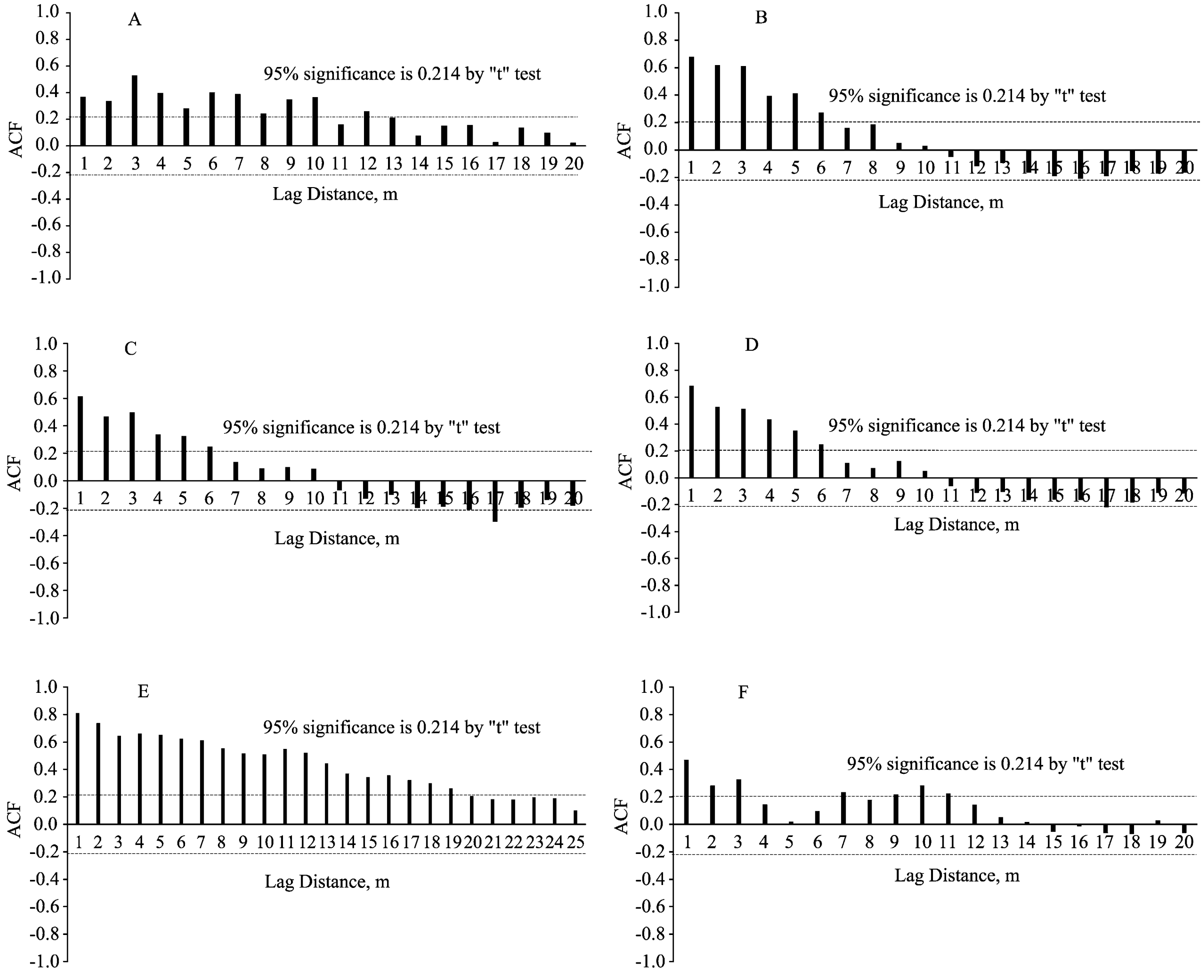

Fig. 3. Autocorrelation functions (ACF) for: (A) number of canes NC; (B) soil P content; (C) soil Ca content; (D) soil Mg content; (E) soil CC content; and (F) soil AS.

analyses using any combination of the sets of observations determining how well the set of number of canes measured across the transect are described by classical regression equations. We found that not more than 38% of the variance of the number of cane data is explained by such analyses using any combination of the sets of observations (Table 1). We verified that the best regression results from using the five dependent variables, and the poorest linear regression were obtained using aggregate stability only.

Using the deterministic equations, the information that a variable carries from its neighborhood is often neglected (McBratney and Webster, 1983) and the additional information from the spatial variability of soil properties is ignored. Because, we note the obvious spatial trends in at least five of the six sets of observations (number of canes, soil P, Ca, Mg, and clay content), we would expect that these variables would be related to each other.

When we are interested in describing the spatial degree of linkage between two variables, the cross-correlation function (CCF) can be used (Wendroth et al., 1997). Shumway et al. (1988); Nielsen et al. (1999) and Cassel et al. (2000) documented the use of the crosscorrelation analysis to determine the spatial

Table 1
Linear and multiple regression analyses of the six sets of observations and values of R^2 coefficient

Equation	R^2
Multiple regression	
NC = 90.561 + 0.018P − 0.127Ca + 0.506Mg − 0.899CC − 3.667AS	0.373
NC = 88.971 + 0.023P − 0.135Ca + 0.442Mg − 0.999CC	0.360
NC = 47.850 + 0.054P − 0.414Ca + 1.401Mg − 7.319AS	0.253
NC = 96.298 + 0.010P + 0.064Ca − 1.026CC − 3.337AS	0.364
NC = 92.497 + 0.0012P+0.262Mg − 0.951CC − 3.722AS	0.370
NC = 91.184 − 0.092Ca + 0.485Mg − 0.913CC − 3.736AS	0.372
NC = 33.549 + 0.074P − 0.505Ca + 1.483Mg	0.192
NC = 94.170 + 0.015P + 0.035Ca − 1.104CC	0.352
NC = 46.647 + 0.045P + 0.141Ca − 8.018AS	0.153
NC = 91.011 + 0.0051P + 0.181Mg − 1.057CC	0.356
NC = 45.752 + 0.0003P + 0.677Mg − 8.332AS	0.214
NC = 97.852 + 0.034P − 1.043CC − 2.938AS	0.361
NC = 89.733 − 0.091Ca + 0.414Mg − 1.020CC	0.358
NC = 47.708 − 0.319Ca + 1.380Mg − 7.719AS	0.244
NC = 96.525 + 0.080Ca − 1.031CC − 3.385AS	0.363
NC = 92.506 + 0.266Mg − 0.951CC − 3.727AS	0.370
NC = 30.816 + 0.067P + 0.076Ca	0.080
NC = 28.421 + 0.010P + 0.587Mg	0.133
NC = 95.197 + 0.029P − 1.108CC	0.351
NC = 48.276 + 0.100P − 7.299AS	0.139
NC = 32.255 − 0.378Ca + 1.460Mg	0.175
NC = 94.474 + 0.058Ca − 1.114CC	0.352
NC = 46.543 + 0.214Ca − 8.344AS	0.147
NC = 91.039 + 0.198Mg − 1.057CC	0.356
NC = 45.754 + 0.678Mg − 8.333AS	0.214
NC = 102.602 − 1.116CC − 2.550AS	0.353
Linear regression	
NC = 32.537 + 0.096P	0.076
NC = 29.687 + 0.183Ca	0.066
NC = 28.420 + 0.622Mg	0.133
NC = 99.601 − 1.164CC	0.346
NC = 52.637 − 7.019AS	0.059

correlation structure of soil properties as a logically quantitative description of the spatial association between two soil attributes. The analysis of the crosscorrelation coefficient between variables that are sampled at neighboring locations with increasing distance also provides more insights on the spatial covariance structure of the two variables. This information can be more useful than the classical correlation coefficient. Also, the correlation over several lags is considered and provides a stronger basis for spatial interpolation (Nielsen et al., 1999).

In our case, CCF was calculated to analyze the spatial correlation structure between the number of canes and: (i) soil P; (ii) soil Ca; (iii) soil Mg; (iv) soil CC; and (v) soil AS. Using the t test at the 5% probability, the crosscorrelogram between number of canes and soil P (Fig. 4A) shows a weak spatial dependence between them. Similar results were also found for the crosscorrelation between number of canes and soil Ca (Fig. 4B).

The crosscorrelogram given in Fig. 4D shows the spatial dependence between number of canes and clay content, is stronger than that between number of canes and soil Mg content presented in Fig. 4C. Similar results were found for the crosscorrelation between number of canes and soil AS (Fig. 4E).

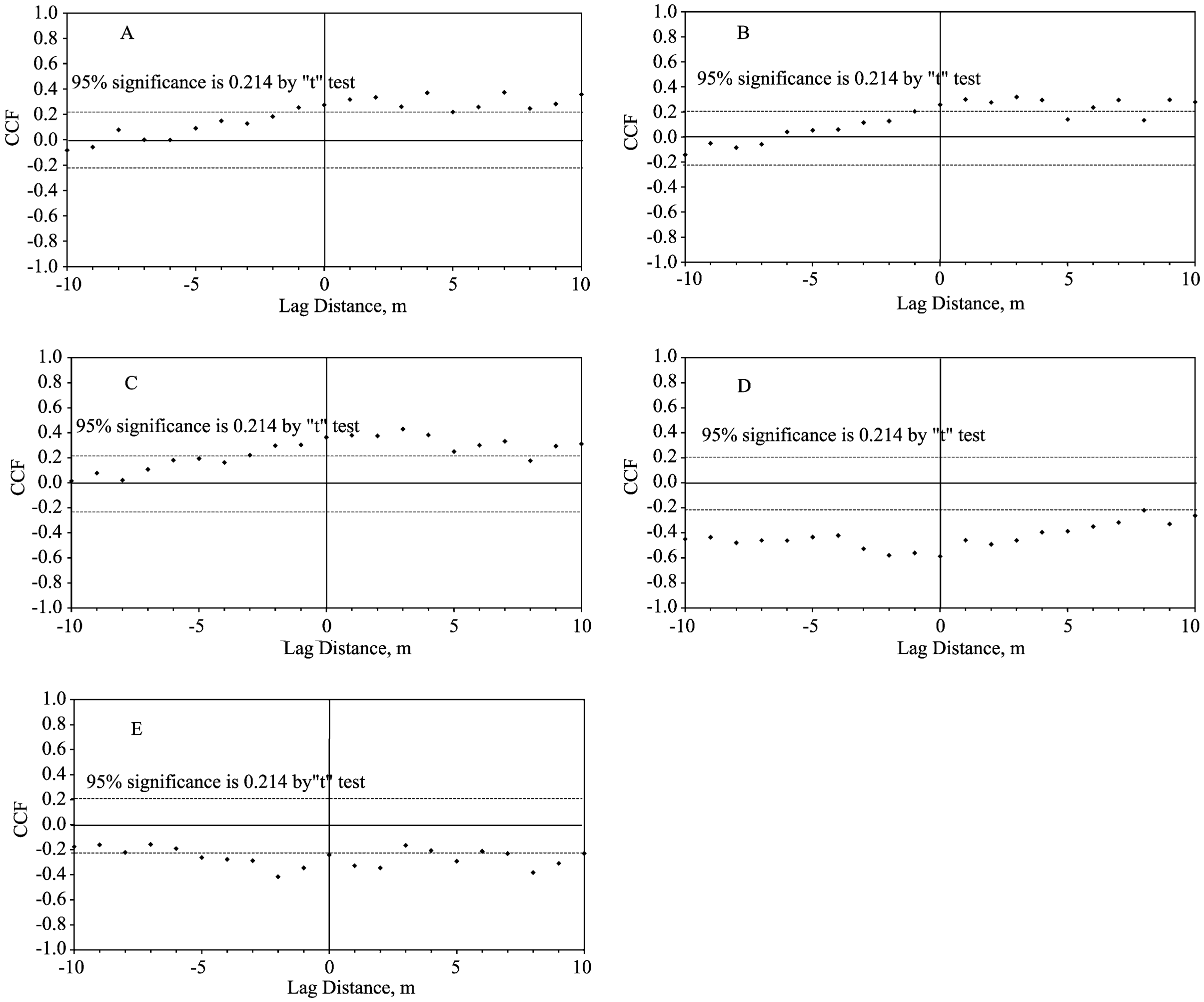

Fig. 4. Crosscorrelogram functions (CCF) between: (A) number of canes NC and soil P; (B) number of canes NC and soil Ca; (C) number of canes NC and soil Mg; (D) number of canes NC and soil CC; and (E) number of canes NC and soil AS.

From the magnitudes of the CCF, we recognize the potential for describing their distributions across the transect of observations with state-space analysis, verifying that its use leads to additional information on the spatial variability of our soil–plant system. Therefore, we are interested in evaluating how well the applied space series analysis can describe the number of canes series using various combinations of the soil P, Ca, Mg, CC and AS series to better understand how the number of canes is related to itself and to the other soil properties in the spatial neighborhood.

Data in Table 2 show all the state-space equations and values of their coefficients of determination (R^2) from linear regressions between estimated and measured values of number of canes NC (all observations have been scaled using Eq. (3)). The estimated values of NC and the ϕ coefficients were obtained using the ASTSA software, developed by Shumway (1988).

Examining the results given in Table 2, the best performance of all the state-space equations was that in which we used soil P, Mg, CC and AS, with the greatest R^2 coefficient of 0.579. In other words, the local and regional variations of soil P, Mg, CC and AS across the transect were the most important variations related to the spatial distribution of NC.

Table 2
State-space equations of number of canes (Fig. 2A) using the data presented in Fig. 2(B–F), and values of R^2 from linear regression between estimated and measured values of NC. All observations have been scaled using Eq. (3)

Equation	R^2
$NC_i = 0.857NC_{i-1} - 0.106P_{i-1} + 0.026Ca_{i-1} + 0.267Mg_{i-1} + 0.163CC_{i-1} - 0.221AS_{i-1} + w_{NC_i}$	0.502
$NC_i = 0.876NC_{i-1} + 0.098P_{i-1} - 0.144Ca_{i-1} + 0.133Mg_{i-1} + 0.0286CC_{i-1} + w_{NC_i}$	0.532
$NC_i = 0.971NC_{i-1} - 0.005P_{i-1} + 0.148Ca_{i-1} - 0.108Mg_{i-1} - 0.016AS_{i-1} + w_{NC_i}$	0.502
$NC_i = 0.898NC_{i-1} + 0.352P_{i-1} - 0.279Ca_{i-1} + 0.040CC_{i-1} - 0.028AS_{i-1} + w_{NC_i}$	0.521
$NC_i = 0.817NC_{i-1} - 0.094P_{i-1} + 0.314Mg_{i-1} + 0.165CC_{i-1} - 0.218AS_{i-1} + w_{NC_i}$	0.579
$NC_i = 0.785NC_{i-1} - 0.197Ca_{i-1} + 0.420Mg_{i-1} + 0.159CC_{i-1} - 0.182AS_{i-1} + w_{NC_i}$	0.521
$NC_i = 0.951NC_{i-1} + 0.068P_{i-1} + 0.048Ca_{i-1} - 0.079Mg_{i-1} + w_{NC_i}$	0.483
$NC_i = 0.920NC_{i-1} + 0.050P_{i-1} + 0.002Ca_{i-1} + 0.018CC_{i-1} + w_{NC_i}$	0.488
$NC_i = 0.902NC_{i-1} + 0.250P_{i-1} - 0.192Ca_{i-1} + 0.027AS_{i-1} + w_{NC_i}$	0.483
$NC_i = 0.926NC_{i-1} + 0.069P_{i-1} - 0.025Mg_{i-1} + 0.019CC_{i-1} + w_{NC_i}$	0.508
$NC_i = 0.968NC_{i-1} + 0.123P_{i-1} - 0.084Mg_{i-1} - 0.018AS_{i-1} + w_{NC_i}$	0.457
$NC_i = 0.925NC_{i-1} + 0.095P_{i-1} + 0.082CC_{i-1} - 0.115AS_{i-1} + w_{NC_i}$	0.503
$NC_i = 0.879NC_{i-1} - 0.012Ca_{i-1} + 0.097Mg_{i-1} + 0.027CC_{i-1} + w_{NC_i}$	0.503
$NC_i = 0.937NC_{i-1} + 0.160Ca_{i-1} - 0.094Mg_{i-1} - 0.018AS_{i-1} + w_{NC_i}$	0.489
$NC_i = 0.959NC_{i-1} + 0.056Ca_{i-1} + 0.062CC_{i-1} - 0.088AS_{i-1} + w_{NC_i}$	0.512
$NC_i = 0.918NC_{i-1} + 0.103Mg_{i-1} + 0.099CC_{i-1} - 0.134AS_{i-1} + w_{NC_i}$	0.494
$NC_i = 0.942NC_{i-1} - 0.039P_{i-1} + 0.087Ca_{i-1} + w_{NC_i}$	0.472
$NC_i = 0.946NC_{i-1} + 0.127P_{i-1} - 0.085Mg_{i-1} + w_{NC_i}$	0.489
$NC_i = 0.912NC_{i-1} + 0.058P_{i-1} + 0.019CC_{i-1} + w_{NC_i}$	0.526
$NC_i = 0.923NC_{i-1} + 0.062P_{i-1} + 0.004AS_{i-1} + w_{NC_i}$	0.477
$NC_i = 0.938NC_{i-1} + 0.148Ca_{i-1} - 0.097Mg_{i-1} + w_{NC_i}$	0.491
$NC_i = 0.923NC_{i-1} + 0.050Ca_{i-1} + 0.018CC_{i-1} + w_{NC_i}$	0.524
$NC_i = 0.929NC_{i-1} + 0.065Ca_{i-1} - 0.003AS_{i-1} + w_{NC_i}$	0.446
$NC_i = 0.904NC_{i-1} + 0.063Mg_{i-1} + 0.023CC_{i-1} + w_{NC_i}$	0.530
$NC_i = 0.914NC_{i-1} + 0.060Mg_{i-1} + 0.015AS_{i-1} + w_{NC_i}$	0.492
$NC_i = 0.972NC_{i-1} + 0.035CC_{i-1} - 0.018AS_{i-1} + w_{NC_i}$	0.478
$NC_i = 0.915NC_{i-1} + 0.073P_{i-1} + w_{Nci}$	0.510
$NC_i = 0.924NC_{i-1} + 0.065Ca_{i-1} + w_{Nci}$	0.501
$NC_i = 0.920NC_{i-1} + 0.068Mg_{i-1} + w_{Nci}$	0.501
$NC_i = 0.958NC_{i-1} + 0.030CC_{i-1} + w_{Nci}$	0.528
$NC_i = 0.962NC_{i-1} + 0.026AS_{i-1} + w_{Nci}$	0.528

We were also interested in running the state-space model for two different scenarios, to evaluate the results when two (Fig. 5A) and three (Fig. 5B) out of four yield observations were not considered for the estimation, in order to identify how close the spatial process of number of canes could be described based on underlying processes manifested in the other variables. According to Wendroth et al. (2001), this implies that the weight of different variables that contribute to the estimation changed and depended on the available data. It can also be seen that the width of the confidence bands increased as compared to the case when all observations of number of canes were included in the estimation of the state-space model (Fig. 5C). For the case, when 50% of observations are available, the updating step becomes possible at every second location where number of canes is given. On the other hand, when 25% of observations are available, the updating step became possible at every fourth location, only. This is the reason why the width of the confidence band is greater than compared to the other cases.

Comparing the different results from the scenarios presented in Tables 1 and 2, we note that the state-space model using all of six data series describes the number of canes ($R^2 = 0.502$) better than the equivalent multiple regression equation ($R^2 = 0.373$). Using five of all series, the state-space model describes the number of canes with R^2 coefficient of 0.579 and the equivalent multiple

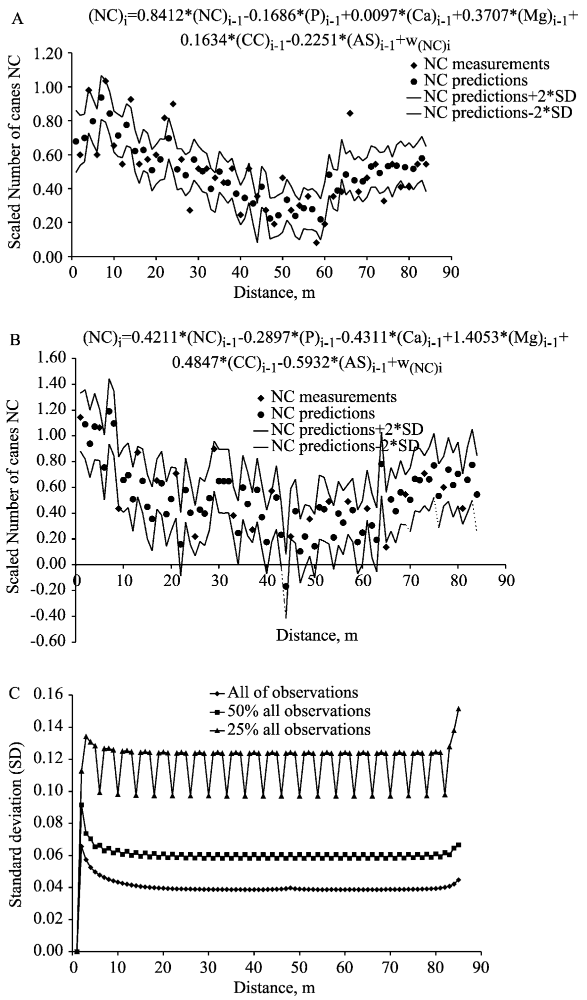

Fig. 5. (A) scaled NC estimated with a state-space model with one out of two observations considered in the estimation; (B) scaled NC estimated with a state-space model with only one out of four observations considered in the estimation; (C) standard deviation behavior along 84 point transect when all observations are included in the estimation of a state-space model; with 50% of all observations (data from A); and with 25% of all observations (data from B).

regression equation with R^2 of 0.372. Also, the state-space model described the number of canes better than the equivalent multiple regression using four, three and two of all series in this study. In summary, we concluded that all of the state-space equations described the number of canes better than the equivalent multiple regression equation. Also, we identified that the soil CC series has an effective contribution to describe the number of canes in this study because it is related to the best performance of each different scenario.

4. Conclusion

The relationships between the number of canes per meter of row and available soil P, Ca and Mg, clay content and aggregate stability of an area cultivated with sugarcane, analyzed through a first order state-space model and standard multiple regressions, show that all used state-space (stochastic model) equations described the spatial distribution of the number of canes better than the equivalent multiple regression equations. It was also identified that the soil clay content spatial series has an effective contribution in describing the number of canes of the crop, because it is related to the best performance in each different scenario. In summary, the adoption of alternative analytical tools, like the state-space approach, are adequate to describe the spatial association between different variables along space or time, and can be used to understand the complex relationship between yield, soil physical and chemical properties, since it is possible to underline an influence that causes a change in their relation, allowing for a management optimization of soil resources and sugarcane yield.

References

Ahuja, L.R., Nielsen, D.R., 1990. Field soil–water relations. In: Stewart, B.A., Nielsen, D.R. (Eds.), Irrigation of Agricultural Crops, ASA, Madison, WI, pp. 143–190.

Cassel, D.K., Wendroth, O., Nielsen, D.R., 2000. Assessing spatial variability in an agricultural experiment station field: opportunities arising from spatial dependence. Agron. J. 92, 706–714.

Coelho, A.M., Doran, J.W., Schepers, J.S., 1998. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), Irrigated Corn Yield as related to Spatial Variability of Selected Soil Properties, Proceedings of the Fourth International Conference on Precision Agriculture, July 19–22, ASA-CSSA-SSSA, St Paul, MN, USA, pp. 441–452.

Cushman, J.H., 1990. An introduction to hierarchical porous media. In: Cushman, J.H., (Ed.), Dynamics of Fluids in Hierarchical Porous Media, Academic Press, San Diego, CA, pp. 1–6.

Dagan, G., 1986. Statistical theory of groundwater flow and transport: pore to laboratory, laboratory to formation, and formation to regional scale. Water Resour. Res. 22, 120–134.

Dourado-Neto, D., Timm, L.C., Oliveira, J.C.M., Reichardt, K., Bacchi, O.O.S., Tominaga, T.T., Cassaro, F.A.M., 1999. State-space approach for the analysis of soil water content and temperature in a sugarcane crop. Scientia Agricola 56, 1215–1221.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA—EMBRAPA. Centro Nacional de pesquisa de Solo, 1997. Manual de métodos de análise de solos, second ed. Rio de janeiro, p. 212.

Gee, G.W., Bauder, J.W., 1986. Particle-size analysis. In: Klute, A., (Ed.), second ed, Methods of Soil Analysis, 9. American Society of Agronomy and Soil Science Society of America, Madison, pp. 383–423.

Gelb, A., 1974. Applied Optimal Estimation, Massachusetts Institute of Technology Press, Cambridge, MA.

Hui, S., Wendroth, O., Parlange, M.B., Nielsen, D.R., 1998. Soil variability—Infiltration relationships of agroecosystems. J. Balkan Ecol. 1, 21–40.

Journel, A.G., Huijbregts, C.J., 1991. Mining Geostatistics, Academic Press, London, p. 600.

Kalman, R.E., 1960. A new approach to linear filtering and prediction theory. Trans. ASME J. Basic Engng 8, 35–45.

Kemper, W.D., Rosenau, R.C., 1986. In: Klute, A., (Ed.), Aggregate Stability and Size Distribution, second ed, Methods of Soil Analysis, vol. 9. American Society of Agronomy and Soil Science Society of America, Madison, pp. 425–442.

van Kessel, C., Wendroth, O., 2001. Landscape research—exploring ecosystem processes and their relationships at different scales in space and time. Soil Till. Res. 58, 97–98.

Li, H., Lascano, R.J., 1999. State-Space Approach for Management of Field Heterogeneity in Cotton, Agronomy Abstracts, American Society of Agronomy, Madison, WI, p. 202.

Li, H., Lascano, R.J., Booker, J., Wilson, L.T., Bronson, K.F., 2001. Cotton lint yield variability in a heterogeneous soil at a landscape scale. Soil Till. Res 58, 245–258.

McBratney, A.B., Webster, R., 1983. How many observations are needed for regional estimation of soil properties. Soil Sci. 135, 177–183.

McGraw, T., 1994. Soil test level variability in Southern Minnesota. Better Crops Potash Phosphate Inst., Norcross 78 (4), 24–25.

Morkoc, F., Biggar, J.W., Nielsen, D.R., Rolston, D.E., 1985. Analysis of soil water content and temperature using state-space approach. Soil Sci. Soc. Am. J. 49, 798–803.

Nielsen, D.R., Alemi, M.H., 1989. Statistical opportunities for analyzing spatial and temporal heterogeneity of field soils. Plant Soil 115, 285–296.

Nielsen, D.R., Wendroth, O., Jürschik, P., Kühn, G., Hopmans, J.W., 1997. Precision agriculture: challenges and opportunities of instrumentation and field measurements. In: Cruvinel, P.E., Crestana, S., Neto, L.M., Colnago, L.A., Mattoso, L.H.C. (Eds.), Simposio Nacional de Instrumentação Agropecuaria—SIAGRO, EMBRAPA, CNPDIA, São Carlos, SP, Brazil, pp. 65–80.

Nielsen, D.R., Wendroth, O., Pierce, F.J., 1999. In: Roberts, P.C., Rust, R.H., Larson, W.E. (Eds.), Emerging Concepts for Solving the Enigma of Precision Farm Research, Proceedings of the Fourth International Conference on Precision Agriculture, 19–22 July, St Paul, Minnesota, ASA-CSSA-SSSA, Madison, WI, pp. 303–318.

Oliveira, J.C.M., Timm, L.C., Tominaga, T.T., Cassaro, F.A.M., Reichardt, K., Bacchi, O.O.S., Dourado-Neto, D., Camara, G.M.S., 2001. Soil temperature in a sugar-cane crop as a function of the management system. Plant Soil 230, 63–68.

Shumway, R.H., 1988. Applied Statistical Time Series Analyses, Prentice Hall, Englewood Cliffs, New York.

Shumway, R.H., Stoffer, D.S., 1982. An approach to time series smoothing and forecasting using the EM algorithm. J. Time Ser. Anal. 3, 253–264.

Shumway, R.H., Biggar, J.W., Morkoc, B.F.M., Nielsen, D.R., 1988. Time and frequency-domain analyses of field observations. Soil Sci. 147, 286–298.

Stevenson, F.C., Knight, J.D., Wendroth, O., van Kessel, C., Nielsen, D.R., 2001. A comparison of two methods to predict the landscape-scale variation of crop yield. Soil Till. Res. 58, 163–181.

Timm, L.C., Fante, L. Jr., Barbosa, E.P., Reichardt, K., Bacchi, O.O.S., 2000. A study of the interaction soil - plant using state-space approach. Scientia Agricola 57, 751–760.

Timm, L.C., Reichardt, K., Oliveira, J.C.M., Cassaro, F.A.M., Tominaga, T.T., Bacchi, O.O.S., Dourado-Neto, D., Nielsen, D.R., 2001. State-Space Approach to Evaluate the Relation between Soil Physical and Chemical Properties, Joint Meeting of the Czech Society of Soil Science and the Soil Science Society of America and International Conference of the Czech Society of Soil Science, Prague, Czech Republic, September, 16–20,.

Warrick, A.W., Myers, D.E., Nielsen, D.R., 1986. Geostatistical methods applied to soil science. In: Klute, A., (Ed.), Methods of Soil Analysis. Part 1. Physical and Mineralogical Methods, ASA-SSSA, Madison, WI, pp. 53–82.

Wendroth, O., Al Omran, A.M., Kirda, K., Reichardt, K., Nielsen, D.R., 1992. State-space approach to spatial variability of crop yield. Soil Sci. Am. J. 56, 801–807.

Wendroth, O., Reynolds, W.D., Vieira, S.R., Reichardt, K., Wirth, S., 1997. Statistical approaches to the analysis of soil quality data. In: Gregorich, E.G., Carter, M.R. (Eds.), Soil Quality for Crop Production and Ecosystem Health, p. 448.

Wendroth, O., Jürschik, P., Giebel, A., Nielsen, D.R., 1999. In: Roberts, P.C., Rust, R.H., Larson, W.E. (Eds.), Spatial Statistical Analysis of On-Site-Crop Yield and Soil Observations for Site-Specific Management, Proceedings of the Fourth International Conference on Precision Agriculture, ASA-CSSA-SSSA, Madison, WI, pp. 159–170.

Wendroth, O., Jürschik, P., Kersebaum, K.C., Reuter, H., van Kessel, C., Nielsen, D.R., 2001. Identifying, understanding, and describing spatial processes in agricultural landscapes—four case studies. Soil Till. Res. 58, 113–127.