6-2005

# Improving Map Accuracy of Soil Variables Using Soil Electrical Conductivity as a Covariate

Alison B. Tarr
*United States Department of Agriculture*

Kenneth J. Moore
*Iowa State University*, kjmoore@iastate.edu

Donald G. Bullock
*University of Illinois at Urbana-Champaign*

Philip M. Dixon
*Iowa State University*, pdixon@iastate.edu

C. Lee Burras
*Iowa State University*, lburras@iastate.edu

# Improving Map Accuracy of Soil Variables Using Soil Electrical Conductivity as a Covariate

**Abstract**

Accurate characterization of soil properties across a field can be difficult, especially when compounded with the diverse landscapes used for pastureland. Indirect methods of data collection have the advantage of being rapid, noninvasive, and dense; they may improve mapping accuracy of selected soil parameters. The objective of this study was to determine if the use of soil electrical conductivity (EC) as a covariate improved mapping accuracy of five soil variables across four sampling schemes and two sampling densities in a central Iowa, USA pasture. In this study, cokriging methods were compared to kriging methods for the measured soil properties of soil pH, available P and K, organic matter and moisture. Maps resulting from cokriging each of the soil variables with soil EC exhibited more local detail than the kriged maps of each soil variable. A small, but inconsistent, improvement occurred in kriging variance and prediction accuracy of non-sampled sites when cokriging was implemented. The improvement was generally greater for soil variables more highly correlated with soil EC. This work indicates that cokriging of EC with less densely and invasively collected soil parameters of P, K, pH, organic matter (OM) and moisture does not consistently and substantially improve the characterization accuracy of pasture soil variability.

**Keywords**

soil properties, pastureland, cokriging, kriging, soil parameters

**Disciplines**

Agriculture | Agronomy and Crop Sciences | Applied Statistics | Climate | Meteorology | Natural Resources and Conservation | Soil Science

**Comments**

This article is from *Precision Agriculture* 6 (2005): 255, doi:10.1007/s11119-005-1385-9.

Precision Agriculture, 6, 255–270, 2005


# Improving Map Accuracy of Soil Variables Using Soil Electrical Conductivity as a Covariate

ALISON B. TARR
*USDA-NASS, Des Moines, IA, USA*

KENNETH J. MOORE
C. LEE BURRAS
*Department of Agronomy, Iowa State University, Ames, IA, USA*

DONALD G. BULLOCK
*Department of Crop Sciences, University of Illinois, Urbana-Champaign, Urbana, IL, USA*

PHILIP M. DIXON
*Department of Statistics, Iowa State University, Ames, IA, USA*

**Abstract.** Accurate characterization of soil properties across a field can be difficult, especially when compounded with the diverse landscapes used for pastureland. Indirect methods of data collection have the advantage of being rapid, noninvasive, and dense; they may improve mapping accuracy of selected soil parameters. The objective of this study was to determine if the use of soil electrical conductivity (EC) as a covariate improved mapping accuracy of five soil variables across four sampling schemes and two sampling densities in a central Iowa, USA pasture. In this study, cokriging methods were compared to kriging methods for the measured soil properties of soil pH, available P and K, organic matter and moisture. Maps resulting from cokriging each of the soil variables with soil EC exhibited more local detail than the kriged maps of each soil variable. A small, but inconsistent, improvement occurred in kriging variance and prediction accuracy of non-sampled sites when cokriging was implemented. The improvement was generally greater for soil variables more highly correlated with soil EC. This work indicates that cokriging of EC with less densely and invasively collected soil parameters of P, K, pH, organic matter (OM) and moisture does not consistently and substantially improve the characterization accuracy of pasture soil variability.

## Introduction

Interest in management of field variability has increased with the increasing availability and adoption of precision agriculture tools and technology. Because of this increased interest, methods for better spatial characterization of a field are valuable. Spatial characterization may include the collection and mapping of such variables as plant health, field productivity, and soil nutrients. Remote sensing imagery can identify differences in vegetation greenness at pixel sizes of 10 m from thousands of kilometers away (Jensen, 1996), and georeferenced yield monitors allow measurement of productivity by the second. While these tools allow increased spatial information to be known about above ground productivity at a high resolution, the basis for most soil management decisions, soil tests, are measured at a relatively low resolution. Because of the high cost, soil tests are measured at a much lower density when compared to the sensors available for greenness measurements and productivity. Gaining more information from traditional soil sampling would aid in the creation of more accurate soil maps and thus, provide direction for more precise management of field inputs such as nutrients and lime.

The obvious way to gain more information from soil sampling is to take more samples. However, this option can be both cost- and labor-prohibitive. Another option is to find a cheaper and easier method that can measure variability of a secondary soil parameter that may help to explain the variability of the soil parameter(s) of interest. This relationship can subsequently be used to improve the prediction variance associated with the soil variable of interest. For example, Vauclin *et al.* (1983) incorporated secondary information on the spatial variability of sand content to help predict values of two primary variables of interest, available water content and water stored at 33 1/3 kPa. McBratney and Webster (1983a) estimated topsoil silt content utilizing its spatial interrelationship with subsoil sand and subsoil silt as secondary variables. Trangmar *et al.* (1986) explored the use of densely sampled 25% HCl-extractable P as a covariate for estimating the variable of interest, $NaHCO_3$-extractable P. Noninvasively collected measures have also served as useful secondary variables. Leenaers *et al.* (1990) improved the quality of interpolation maps of zinc concentration in the Geul floodplain in the southern Netherlands by incorporating elevation data. Jaynes (1996) found that electromagnetic induction (EMI) conductivity data reduced the estimation error of soil organic carbon fraction by 33% compared to kriging only soil organic carbon samples.

All of these examples incorporated a geostatistical method called cokriging. Geostatistics is a field of study that involves modeling and prediction of spatial variability (Journel and Huijbregts, 1978). Kriging is a method of interpolation used when a variable displays spatial dependence. Cokriging is also an interpolation method used where there are two or more spatially interdependent variables. Often, cokriging is used when one or more other properties have been extensively sampled in comparison to the variable of interest (Oliver, 1987). Ideally, the densely sampled variable, called a covariate, secondary variable, or subsidiary variable, is measured more cheaply and quickly than the property of interest, or target variable.

Measuring soil variability through the use of EMI can be performed quickly and inexpensively. Soil electrical conductivity (EC) can be measured on an extremely small grid in a rapid, easy, and nondestructive manner by the use of EMI. Soil EC is a measurement that is affected primarily by a combination of soil water content, dissolved salt content, clay content and mineralogy and soil temperature (McNeill, 1980a). Soil EC measured by EMI has been correlated with clay content (Williams and Hoey, 1987; Doolittle *et al.*, 1994), soil water content (Kachanoski *et al.*, 1988; Sheets and Hendrickx, 1995), sand deposition (Kitchen *et al.*, 1996), total soluble salts (Williams and Hoey, 1987), yield (Jaynes *et al.*, 1995), and soil available N (Eigenberg *et al.*, 2002). Therefore, soil EC data may serve as a covariate and noninvasively provide valuable and inexpensive information to aid in the production of more accurate soil maps for certain variables (Jaynes, 1996).

In this study, cokriging methods were compared to kriging methods for measured soil properties. The objective of this study was to determine if the use of soil EC as a covariate improved mapping accuracy of five soil variables across four sampling schemes and two sampling densities in a central Iowa, USA pasture.

## Materials and methods

### *Field methods*

Research was conducted in June 2001 at the Iowa State University Rhodes Research Farm (41°52′ N, 93°10′ W) in central Iowa, USA. The Wisconsin loess-covered landscape has an underlying Yarmouth-Sangamon paleosol. The soils are primarily slope and erosion phases of the Fayette (Fine-silty, mixed, superactive, mesic Typic Hapludalfs) and Clarinda (Fine, smectitic, mesic Vertic Argiaquolls) series (T.E. Fenton, personal communication, 2001). The pasture site of study included topographically distinct summit, sideslope, toeslope, backslope, and opposite summit landscape positions.

A dense sampling grid consisting of 116 points was devised for a 0.42-ha non-grazed, grass-legume pasture (Figure 1). Sampling points were arranged in a triangular grid with inter- and intra-row separation distances of 6 m. In order to obtain data from samples located closer than 6 m, an additional point was sampled within each row at randomly chosen 1 or 2 m separation distances. This short range variation in soil samples was investigated in order to obtain a more reliable experimental semivariogram model (Burgess and Webster, 1980; Kravchenko and Bullock, 2002).

Apparent soil EC was measured with the Geonics EM-38 (Ontario, Canada) as it was pulled behind a four-wheel drive vehicle in a nonconductive cart. The EM-38 was operated in the vertical dipole orientation at 0.20 m above the soil surface. The EM-38 integrates over an area approximately equal to its length of 1 m and over a depth of approximately 3 m (McNeill, 1980b). However, the measurement is primarily influenced by the 0–1.5 m depth increment (McNeill, 1980b). The vertical rather than horizontal dipole orientation was used so that any fluctuations in carrying height of the EM-38 above the soil surface would have little impact on EC measurements (McNeill, 1980b). Given the speed of travel and rate of EC data collection, an EC measurement was logged approximately every 4 m in 2 m transects throughout the pasture. Each soil EC measurement was georeferenced using a Trimble Global Positioning Systems (GPS) Pathfinder® Pro XR receiver (Trimble,

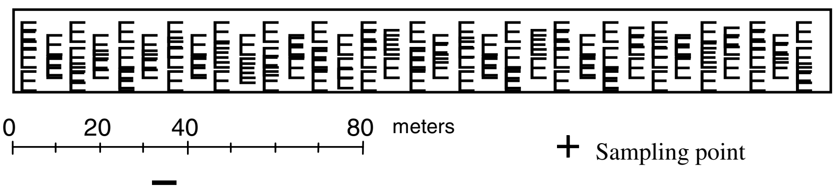


*Figure 1.* Initial dense sampling grid (116 sampling points).

Sunnyvale, CA, USA), and all GPS locations were differentially corrected (DGPS) to obtain 1–2 m accuracy. EM-38 measurements were recorded via a direct connection to Trimble System Controller (TSC) 1 Asset Surveyor software (Trimble, Sunnyvale, CA, USA) in the GPS datalogger. Positional data for the soil EC values were corrected for the lag distance between GPS receiver and the EM-38 instrument. Soil EC values at each of the 116 grid points were interpolated from the dense data set of the 834 georeferenced EC points.

Soil samples were collected via coring following the EM-38 measurements. At each of the 116 sampling sites, five subsample (25 mm $\times$ 150 mm) soil cores were collected and combined for soil pH and available P and K analysis in the Iowa State University Soils Testing Laboratory (Ames, IA, USA). Organic matter was analyzed from a single core composed of three depth increments to 510 mm, using a dry combustion method. Average organic matter percentage over all three depth increments was reported. To quantify soil moisture at each of the 116 sampling sites, seven subsamples (25 mm $\times$ 150 mm) soil cores were taken. The percent moisture content of each of the subsample cores was analyzed using the gravimetric moisture method (Buckman and Brady, 1971) and an average percentage soil moisture was reported for each of the 116 sampling sites.

Sub-meter accuracy elevation data were recorded using a Leica System 500 real time kinematic (RTK) system (Leica, Switzerland) and slope data were calculated from this using ArcView 3.2 Spatial Analyst (ESRI, Redlands, CA, USA). Geostatistical analyses were performed using ArcView 8.1 ArcGIS Geostatistical Analyst (ESRI, Redlands, CA, USA).

### *Sampling schemes*

Four different sampling patterns at two densities were created from the original sampling grid ($n$ = 116). The sampling schemes were a grid pattern, a triangular pattern, a multistage clustering scheme, and a random scheme. The $n$ = 30 and $n$ = 15 sampling schemes are shown in Figures 2 and 3, respectively. Because the sampling schemes were created from the original sampling grid (Figure 1), there were some restrictions on the arrangement of the patterns. The grid pattern was a rectangular grid with 6-m intra-row and 12-m inter-row separation distances for the $n$ = 30 scheme. Intra-row grid spacing was also 6 m for the $n$ = 15 scheme, but inter-row separation distance increased to 24 m.

The sampling scheme originated on the west end of the pasture and because of the specified sample size, sampling density on the east end of the pasture was less dense for both the $n$ = 15 and $n$ = 30 grid schemes. For similar restriction reasons, the $n$ = 30 triangular scheme was more dense on the east end of the pasture and triangular grid size varied for the $n$ = 15 scheme. The triangular pattern was not equilateral; the triangles were formed with a base length between points of 6 m and a side length of 19.0 m for nearly all of the pasture and 12–13.4 m on the extreme east end of the pasture for the $n$ = 30 scheme. Base length between points was also 6 m for the $n$ = 15 triangular scheme with side lengths of 30.6 m and 42.4 m.

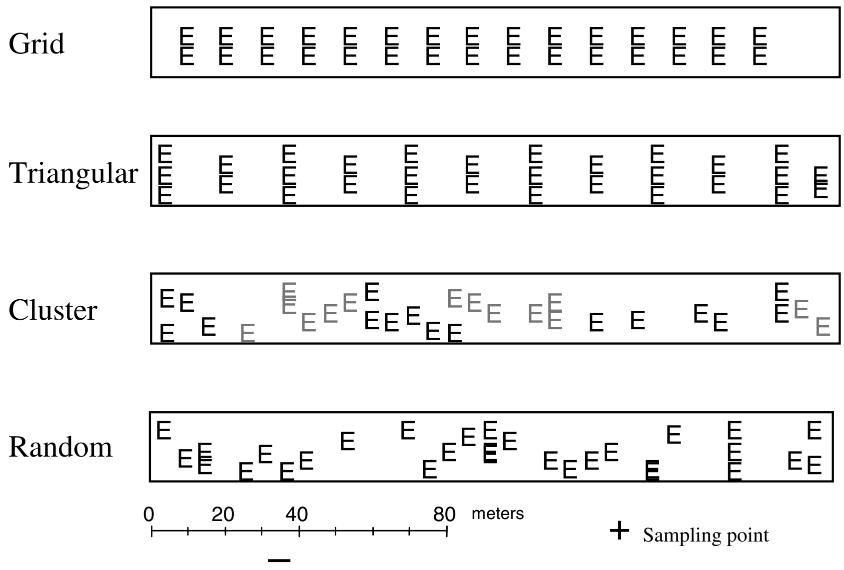


*Figure 2.* Four sampling schemes at $n = 30$ density.

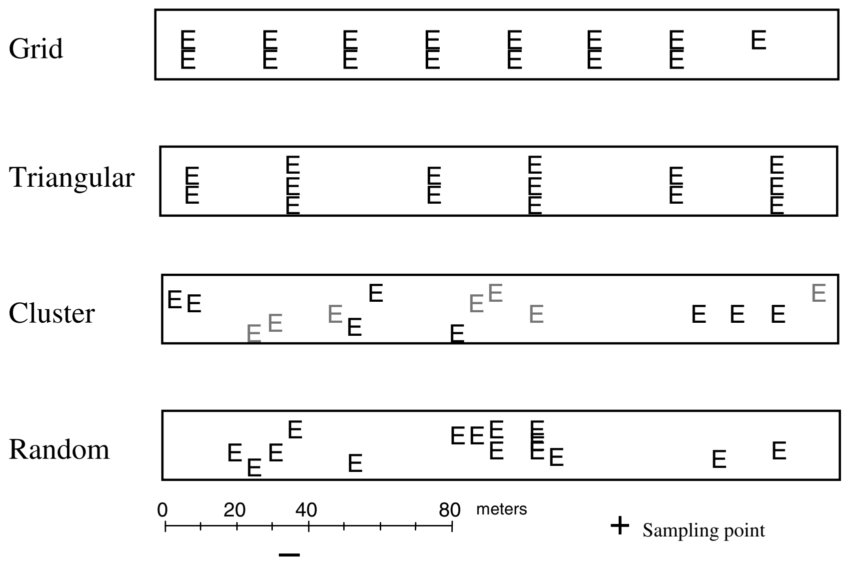


*Figure 3.* Four sampling schemes at $n = 15$ density.

The multistage clustering scheme was produced by first implementing a fuzzy k-means algorithm. The algorithm was used to stratify the field into relatively homogeneous zones based on densely collected soil parameters. The fuzzy k-means method has been utilized for classifying soil and landscape data when binary or strictly discrete groupings are not adequate to describe natural systems (Burrough, 1989; McBratney and de Gruijter, 1992; Odeh *et al.*, 1992; Irvin *et al.*, 1997). Given the continuous nature of soils, fuzzy set classification provides a suitable means of classifying areas of a field. A combination of topographic attributes and apparent soil EC were used to delineate zones using the software program Management Zone Analyst© (University of Missouri-Columbia and Agricultural Research Service, Columbia, MO, USA, Version 1.0). The clustering algorithm was iterated for 3–9 zones. Based on the fuzziness performance index (FPI) and normalized classification entropy (NCE), five zones appeared optimal for establishing strata of homogeneity in the pasture. The FPI and NCE performance indices were discussed and used to evaluate the varying number of climatic classes by McBratney and Moore (1985). Within these five strata, two intensities of a ranked set sampling scheme (McIntyre, 1952) were analyzed: $n = 30$ and $n = 15$ (Figures 2 and 3). Ranked set sampling was first described by McIntyre (1952) as a method for obtaining more precise and unbiased measurements of forage yield. The five strata delineated by clustering corresponded to the sets in ranked set sampling. The points in each of the five strata were ranked in order of EC value magnitude.

Soil EC was chosen as a concomitant variable because it was the variable most easily and accurately collected (Patil *et al.*, 1994), and it was correlated with the soil variables of interest. When constructing the $n = 30$ scheme, six points were selected in each of the five strata. With this scheme, the six points were chosen based on maximizing the within-zone variation of soil EC. The six points were selected based on choosing the minimum, maximum, and four in-between quantile values of soil EC within each zone. Maximum within-zone variation was sought in order to maintain the variability identified throughout the field. Similarly, when devising the $n = 15$ scheme, three points were selected in each of the five strata with the goal of maximizing the within-zone variation of the concomitant variable, soil EC. The three points were selected based on choosing the minimum, maximum, and median values of soil EC within each zone. Lastly, using a random number generator, sampling schemes with sizes $n = 30$ and $n = 15$ were produced from the original 116 sampling points for the random sampling schemes.

Because direct soil measurements were taken at each of the original 116 sampling points, relatively large validation sets were available. Eighty-six and 101 points, respectively, were used as independent validation sets for the $n = 30$ and $n = 15$ sampling densities. These validation sets were used to determine if cokriging improved the accuracy of the maps.

## Results and discussion

### *Statistical data analysis*

Statistical summaries of topographic and soil attributes for the initial, dense grid ($n = 116$) sampling scheme are presented in Table 1. Large CVs and ranges for soil P, K, pH, OM, and moisture illustrated the magnitude of soil variability within the pasture (Table 1). From a sampling perspective, this known variability is of interest

*Table 1.* Descriptive statistics of measured topographic and soil attributes for 116 sampling points

| | Depth | Mean | SD | CV | Min. | Max. |
|---|---|---|---|---|---|---|
| Elevation, m | – | 301.2 | 2.1 | 0.7 | 296.6 | 303.0 |
| Slope, degrees | – | 5.4 | 1.6 | 29.6 | 1.7 | 8.8 |
| Available P, mg $kg^{-1}$ | 0.15 m avg | 14.9 | 7.1 | 47.7 | 4.0 | 51.5 |
| Available K, mg $kg^{-1}$ | 0.15 m avg | 159.8 | 59.1 | 37.0 | 56.0 | 369.0 |
| pH, 1:1 soil/water | 0.15 m avg | 6.0 | 0.4 | 7.3 | 5.4 | 7.3 |
| Organic matter, g $kg^{-1}$ | 0.51 m avg | 2.2 | 0.8 | 36.8 | 1.1 | 6.1 |
| Moisture, g $kg^{-1}$ | 1.07 m avg | 27.6 | 2.3 | 8.4 | 24.2 | 41.8 |
| Soil EC, mS/m | ≅1.5 m | 44.3 | 4.7 | 10.5 | 35.5 | 55.0 |

*Table 2.* Partial correlation matrix among chemical and physical soil attributes for 116 sampling points

| | Elevation | Slope | P | K | pH | OM | Moisture | Soil EC |
|---|---|---|---|---|---|---|---|---|
| Elevation | 1 | | | | | | | |
| Slope | −0.02 | 1 | | | | | | |
| Available P | −0.15 | −0.27 | 1 | | | | | |
| Available K | 0.65 | 0.08 | 0.14 | 1 | | | | |
| PH | −0.75 | −0.31 | 0.14 | −0.67 | 1 | | | |
| Organic matter | −0.42 | −0.21 | 0.44 | −0.26 | 0.41 | 1 | | |
| Moisture | −0.46 | −0.15 | 0.31 | −0.30 | 0.51 | 0.40 | 1 | |
| Soil EC | −0.62 | −0.10 | −0.06 | −0.56 | 0.70 | 0.24 | 0.17 | 1 |

because it is important that a sampling technique can identify this variability in its resulting map. Furthermore, management decisions are made based upon this map.

The degree to which soil EC is spatially correlated with the other soil variables of interest is called coregionalization. Although statistical correlation does not imply spatial correlation, the dense data set in this study does suggest a baseline for spatial relationships to exist among the soil variables. The relationships among the chemical and physical attributes for the initial, dense sampling scheme are shown in Table 2. The results of the large database indicated a strong correlation between soil EC and soil pH, elevation, and soil K ($r$ values greater than 0.5). Soil EC was moderately correlated with soil moisture and OM values ($0.10 < r < 0.50$) and weakly correlated with soil P and slope ($r \leq |0.10|$) (Table 2).

In the pasture of study, variability in soil EC appeared closely related to the variation in landscape position and depth to paleosol (T.E. Fenton, personal communication, 2002). Higher values of soil EC were measured in the toeslope positions. These positions have a higher moisture content and are underlain by a clay-textured paleosol (Clarinda series). Lower values of soil EC were measured upslope on the summit positions where the soils were developed entirely in loess. Middle values were measured on the backslope where the soils are formed in loess and the underlying paleosol. Because the soil was not dominated by carbonates, the variation in soil EC values was likely related to soil moisture content and textural properties (Brevik and Fenton, 2002). Textural properties influence soil parameters such as organic matter and ionic properties; thus, it was concluded that soil EC is measuring variation in soil properties related to moisture and texture.

### *Geostatistical data analysis*

Ordinary kriging is widely used in soil literature as a method for interpolation. Both the theory and application of ordinary kriging are described in depth by Journel and Huijbregts (1978) and McBratney and Webster (1983a). We investigated the value of using soil EC as a covariate for improving mapping accuracy by comparing kriging the soil parameters to cokriging the soil parameters with soil EC as a covariate.

Kriging and cokriging were performed using the Geostatistical Analyst extension in ArcView 8.1. Models considered included linear, exponential, circular, spherical, tetraspherical, pentaspherical, and Guassian. Final variogram model selection was based on cross-validation (Vauclin *et al.*, 1983; Warrick *et al.*, 1986). In a cross-validation, each point in the sampling scheme is removed singly and its value is predicted based on kriging the remaining data. The resulting root mean square error (RMSE) of the cross-validation process was examined, and the variogram model with the lowest RMSE was selected (Vauclin *et al.*, 1983; Heisel *et al.*, 1999). In all cases either exponential or spherical models were selected and final selected model $r^2$ values were as large as 0.95. There were no nugget models (i.e. lack of spatial correlation). Visual observations were also made on all selected semivariogram models and all appeared appropriate. Skewness results indicated that the data was not normally distributed. To improve normality, the data for soil organic matter, phosphorus, and potassium were log transformed. Data were reported on the non-transformed values.

The pasture of study was oriented mostly in one dimension, and there were insufficient sample pairs of the soil variables for the $n = 30$ and $n = 15$ sampling schemes to obtain well-structured directional semivariograms (Trangmar *et al.*, 1986). Therefore, it was assumed that all semivariograms were isotropic. Lag distances ranged from 2 m to 14 m with the majority of values being 8–9 m. Lag distances were autocalculated using the ArcView 8.1 Geostatistical Analyst extension. This method tries a series of lag values, with their size increasing in a geometric sequence. Geostatistical Analyst then examines all of the lags and finds the lag and set of variogram parameters that have the "best fit," or smallest weighted least squares (J.M. Ver Hoef, personal communication, 2002).

The five soil variables of interest, soil P, K, pH, OM, and moisture were less sampled compared to the soil EC readings. This is the situation where cokriging is most useful. Because of the ease of collecting dense, rapid and georeferenced soil EC data, its spatial relationship with the five soil parameters was explored. In this 0.42 ha pasture, 834 soil EC points were measured. Consequently, the ratio of sampling intensities of soil EC to the other soil variables was nearly 28 for the $n = 30$ schemes and nearly 56 for the $n = 15$ schemes (Figure 4).

To apply cokriging, it was necessary to model semivariograms for each soil variable separately as well as cross-variograms for all pairs of soil EC and soil variable measured at the same location (McBratney and Webster, 1983a; McBratney and Webster, 1983b; Vauclin *et al.*, 1983; Triantafilis *et al.*, 2001). The correlations

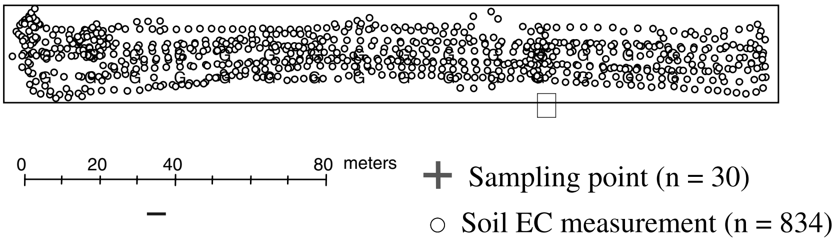


*Figure 4.* Map of soil sampling points and soil EC covariate points.

shown in Table 2 were also examined (Chien *et al.*, 1997). Soil variables that were more highly correlated with soil EC also produced better-structured cross-variograms (Journel and Huijbregts, 1978).

Proof of some level of coregionalization between soil EC and the five soil parameters of interest was important for further cokriging steps.

### *Map results*

***Kriging vs. cokriging.*** In general, a visual discrepancy existed between the kriged and cokriged maps. The cokriged maps exhibited more local detail and less smoothness in their depiction of the variability of the five soil parameters. However, both methods generated similar trends in the variability of the soil parameters. Comparable results have been found for topsoil silt cokriged with subsoil silt and sand as covariates (McBratney and Webster, 1983a), $NaHCO_3$-extractable P cokriged with 25% HCl-extractable P (Trangmar *et al.*, 1986), $NO_3$ cokriged with soil EC, (Zhang *et al.*, 1992) and soil EC of soil paste extract cokriged with apparent soil EC (Vaughan *et al.*, 1995). Improved local detail of the cokriging maps was due to the finer sampling grid of the covariate, soil EC (McBratney and Webster, 1983a). Representative maps of these observations are shown in Figure 5.

Several exceptions to the trend of cokriging to increase local detail of variation were found in mapping soil moisture, OM, and P. However, these exceptions all exhibited similar spatial variability. In these anomalies, the autocorrelation of the soil variables revealed small lags and short range values; thus, the soil variables displayed short-range spatial variability. Small lags ranged from 2 to 4 m and range values were only 14–30 m for the semivariograms of the soil variables (data not shown). In addition, it is interesting to note that the three soil variables that exhibited less local detail with cokriging were also the soil variables least correlated with soil EC (Table 2). This result could be due to the relatively low cross-correlation between these variables (soil moisture, OM, and P) and soil EC; thus, modeling of the cross-variogram was made difficult (McBratney and Webster, 1983a). The result could also be due to the geometric arrangement of the soil sampling points themselves and how that arrangement affected the modeling of the semivariograms and cross-variograms (McBratney and Webster, 1983a). Whether or not the reduced local detail in the maps actually affected mapping accuracy will be discussed in the quantitative portion of the results.

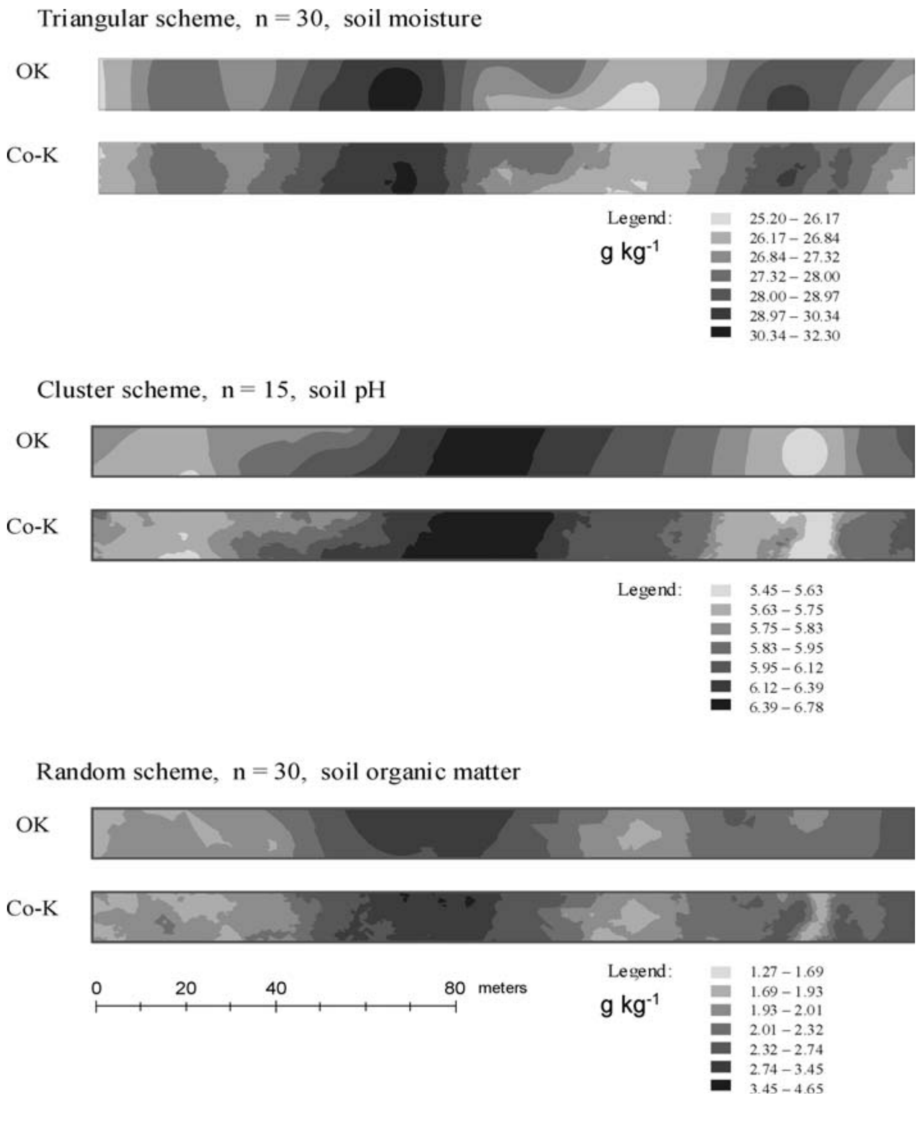


*Figure 5.* Representative comparisons of kriged (OK) and cokriged (Co-K) maps for (a) $n = 30$ triangular sampling scheme, soil moisture, (b) $n = 15$ cluster sampling scheme, soil pH and (c) $n = 30$ random sampling scheme, soil organic matter.

*Sampling scheme comparison*

Based on the design and results of this experiment, it cannot be concluded which of the four sampling schemes is optimal for producing accurate maps of soil variability. Each of the schemes used in this study was just one depiction of a grid scheme or one depiction of a random or clustered scheme in the pasture. Multiple scenarios of each scheme should be analyzed further in order to answer this question. General conclusions about the sampling schemes will be discussed after examining them quantitatively.

*Sampling density comparison*

In general, maps of $n = 30$ and $n = 15$ schemes displayed similar spatial patterns for both kriging and cokriging. Visually, it appeared that cokriging the soil variables of interest with soil EC had a greater impact on the maps for the less dense sampling scheme. The cokriged maps for $n = 15$ exhibited greater local detail than the cokriged maps for $n = 30$, especially for the soil variables most highly correlated with soil EC (soil pH and K) (Table 2). Thus, when there are fewer points measured invasively, the variability of the noninvasive soil parameter (soil EC) appeared to have a greater influence on prediction of non-sampled areas of soil pH and K.

Few other observations could be made from visual assessment of the kriged and cokriged maps. Differences in mapping characteristics of the best-fit semivariogram models made it difficult to conclude "real" differences among maps. Maps are effective for identifying trends, but the actual differences between methods should be analyzed quantitatively. That will be done as a quantitative assessment of the results of kriging and cokriging in this work.

*Quantitative results*

***Kriging vs. cokriging.*** When comparing results estimated with kriging and cokriging, two parameters were analyzed. Root mean square error (RMSE) of prediction and the correlation between predicted values from kriging and the actual values taken from invasive soil measurements were evaluated. As previously mentioned, validation sets of 86 and 101 sample points were available for the $n = 30$ and $n = 15$ sampling schemes, respectively. The RMSE should be small for an unbiased and precise prediction. For both kriging and cokriging at each sampling density and for each sampling scheme, the root mean square error was calculated as Eq. (1):

$$RMSE = \left\{ 1/n \sum_{i=1}^{n} [z(s_i) - z^*(s_i)]^2 \right\}^{0.5} \tag{1}$$

where $n$ is the number of sample sites in the validation set, $z(s_i)$ are the observed soil values, and $z^*(s_i)$ are the predicted values. As shown in Table 3, RMSE for cokriging was not consistently lower than the RMSE for kriging as would be expected if cokriging helped improve the prediction of the validation sites. Correlation between each soil variable and soil EC was likely not the single factor affecting the success of cokriging. Relative reduction in RMSE was defined by Eq. (2):

$$100(RMSE_k - RMSE_{ck})/RMSE_k \tag{2}$$

*Table 3.* Validation set root mean square errors (RMSE) of soil data prediction for kriging and cokriging the n = *30 and n = 15 sampling densities for each of the four sampling schemes. Coefficient of determination (r*$^2$) between predicted and actual values for validation set

| | | | Grid | | Triangular | | Cluster | | Random | |
|---|---|---|---|---|---|---|---|---|---|---|
| Target variable | Sample size | | OK[a] | Co-K | OK | Co-K | OK | Co-K | OK | Co-K |
| P | 30 | RMSE[b] | 5.468 | 5.516 (+)[c] | 5.575 | 5.569 (−) | 4.903 | 6.787 (+) | 6.508 | 6.929 (+) |
| | | $r^2$ | 0.470 | 0.262 (−) | 0.332 | 0.340 (+) | 0.549 | 0.14 (−) | 0.206 | 0.088 (−) |
| | 15 | RMSE | 8.106 | 8.071 (−) | 7.693 | 7.625 (−) | 6.884 | 7.382 (+) | 8.008 | 7.952 (−) |
| | | $r^2$ | 0.031 | 0.057 (+) | 0.022 | 0.026 (+) | 0.265 | 0.046 (−) | 0.023 | 0.028 (+) |
| K | 30 | RMSE | 52.92 | 55.09 (+) | 43.08 | 39.85 (−) | 45.7 | 44.13 (−) | 45.88 | 45.71 (−) |
| | | $r^2$ | 0.389 | 0.371 (−) | 0.447 | 0.499 (+) | 0.518 | 0.500 (−) | 0.516 | 0.509 (−) |
| | 15 | RMSE | 47.27 | 49.03 (+) | 46.79 | 46.98 (+) | 44.16 | 45.28 (+) | 62.39 | 61.49 (−) |
| | | $r^2$ | 0.411 | 0.357 (−) | 0.491 | 0.479 (−) | 0.458 | 0.437 (−) | 0.099 | 0.111 (+) |
| PH | 30 | RMSE | 0.178 | 0.176 (−) | 0.185 | 0.179 (−) | 0.169 | 0.165 (−) | 0.175 | 0.168 (−) |
| | | $r^2$ | 0.828 | 0.835 (+) | 0.830 | 0.845 (+) | 0.858 | 0.854 (−) | 0.845 | 0.854 (+) |
| | 15 | RMSE | 0.184 | 0.186 (+) | 0.229 | 0.204 (−) | 0.306 | 0.286 (−) | 0.224 | 0.235 (+) |
| | | $r^2$ | 0.817 | 0.806 (−) | 0.795 | 0.802 (+) | 0.555 | 0.62 (+) | 0.795 | 0.731 (−) |
| OM | 30 | RMSE | 0.822 | 0.693 (−) | 0.746 | 0.744 (−) | 0.762 | 0.769 (+) | 0.767 | 0.763 (−) |
| | | $r^2$ | 0.116 | 0.191 (+) | 0.232 | 0.235 (+) | 0.187 | 0.166 (−) | 0.145 | 0.159 (+) |
| | 15 | RMSE | 0.918 | 0.915 (−) | 1.048 | 1.047 (−) | 0.733 | 0.727 (−) | 0.839 | 0.763 (−) |
| | | $r^2$ | 0.016 | 0.020 (+) | 0.036 | 0.037 (+) | 0.250 | 0.257 (+) | 0.012 | 0.129 (+) |
| Moisture | 30 | RMSE | 2.017 | 1.957 (−) | 2.004 | 1.991 (−) | 1.611 | 1.539 (−) | 2.536 | 2.559 (+) |
| | | $r^2$ | 0.379 | 0.395 (+) | 0.381 | 0.396 (+) | 0.501 | 0.369 (−) | 0.097 | 0.125 (+) |
| | 15 | RMSE | 2.081 | 2.098 (+) | 2.108 | 2.1 (−) | 2.414 | 2.392 (−) | 2.325 | 2.329 (+) |
| | | $r^2$ | 0.208 | 0.201 (−) | 0.213 | 0.214 (+) | 0.044 | 0.080 (+) | 0.131 | 0.109 (−) |

[a]OK is ordinary kriging, Co-K is cokriging.
[b]RMSE units are specific to each soil variable: mg kg$^{-1}$ soil P and K, units pH, g kg$^{-1}$ OM and moisture.
[c]Plus or minus sign in parentheses indicates an increase (+) or decrease (−) in RMSE or $r^2$ value between OK and Co-K.

where $RMSE_k$ and $RMSE_{ck}$ are the root mean square errors of kriging and cokriging, respectively (adapted from Zhang *et al.*, 1992). These findings were comparable to the observations by Yates and Warrick (1987), which found that a reduction in kriging variance is observed as the correlation between the covariate and target variable increased. The findings of the current study were more volatile that those found by Yates and Warrick (1987), but Figure 6 shows that the lower correlation between soil EC and P resulted in the lowest reductions (largest gains) of RMSE. While OM exhibited the largest reduction in RMSE, soil pH showed more consistent reductions in RMSE for the sampling scheme and density combinations. General trends can be drawn from this figure, but it is important to note that the efficacy of cokriging is more than a function of correlation between a covariate and target variable. The efficacy also includes the strength of the spatial cross-correlation between the covariate and target variable, the geometric sampling pattern, and the ratio of the sampling intensities of the covariate and target variable (McBratney and Webster, 1983a).

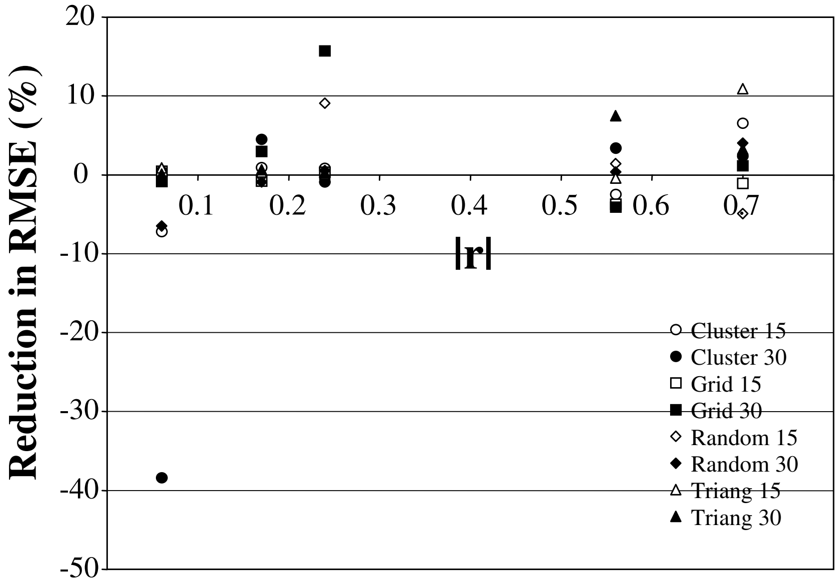


*Figure 6.* Reduction in root mean square error (RMSE) of prediction due to cokriging as a function of the absolute correlation ($r$) between soil EC and the five selected soil variables (P, moisture, OM, K and pH). Data is plotted for each soil sampling pattern and density.

In addition, $r^2$ values did not increase due to cokriging for every sampling scheme and density combination (Table 3). Soil OM, pH, and moisture displayed improved $r^2$ values most consistently among the soil variables.

*Sampling density comparison*

By decreasing the number of sampling sites, Vauclin *et al.* (1983) demonstrated the advantage of cokriging over kriging by comparing the estimation variances at interpolated points. When the effect of sampling density was compared to cokriging efficacy, no clear relationship was found (Table 3). For example, cokriging with soil EC provided better prediction accuracy for soil P at the lower sampling density ($n = 15$) but better prediction accuracy for soil moisture at the higher sampling density ($n = 30$) (Figure 5). It is hypothesized that the benefit of cokriging was not observed because of the random choice of fifteen points or because semivariogram and cross-variogram models were not well defined with such few sample pairs at a low density ($n = 15$). Absolute values of RMSE were generally higher and correlation between predicted and actual values were generally lower with the lower sampling density, as also shown by Zhang *et al.* (1992) (Table 3).

*Sampling scheme comparison*

The mapping accuracy from cokriging was also affected by the sampling pattern that was used. The triangular set of points showed the largest benefit from cokriging based on the reduction in RMSE (Figure 6) and improvement in $r^2$ values (Table 3). The triangular scheme also showed the strongest response to correlation between soil EC and the soil variables. As shown in Figure 6, the reduction in RMSE increased favorably as the correlation between the covariate (soil EC) and the soil variables of interest increased. The cluster and random sampling schemes may not have benefited as much from the cokriging with soil EC as a covariate because of the geometric pattern of the sampling sites. As seen from the maps of the cluster and random schemes (Figures 2 and 3), coverage of the sampling is somewhat uneven. Because of these gaps in coverage, estimation of the non-sampled sites may be difficult because information that could have been obtained is lacking (Webster and Oliver, 2001). The estimation error increases the farther an interpolated point is from the sampling points (Burgess *et al.*, 1981). In addition, because the soil variables were spatially autocorrelated to some degree, the closely neighboring points found in the random scheme were likely quite similar; thus, duplicated information was collected (Webster and Oliver, 2001). Despite gaps in sampling coverage, the benefit of soil EC as a covariate is that it is collected on a much denser grid than the soil variables of interest. However, because only one example of each sampling scheme was used, generalizations about each of the schemes cannot be made.

## Conclusions

Soil EC's spatial relationship with the five soil variables was exploited through the use of cokriging. Maps resulting from cokriging soil EC with each of the soil variables exhibited more local detail than the kriged maps of each soil variable. The increased detail was accompanied by an overall small but inconsistent, improvement in kriging variance and prediction accuracy of non-sampled sites when cokriging was implemented. As would be expected, the improvement was generally greater for soil variables more highly correlated with soil EC. A clear improvement from cokriging with a lower density target variable was not observed. The triangular grid appeared to benefit most from cokriging, but the results are inconclusive because only one triangular scheme (not multiple) was analyzed. These results lead us to the conclusion that cokriging of EC with less densely and invasively collected soil parameters of P, K, pH, organic matter (OM) and moisture does not consistently and substantially improve the characterization accuracy of pasture soil variability.

## References

Brevik, E. C. and Fenton, T. E. 2002. Influence of soil water content, clay, temperature, and carbonate minerals on electrical conductivity readings taken with an EM-38. Soil Survey Horizons **43**, 9–13.

Buckman H. O. and Brady N. C. 1971. *The nature and properties of soils*. 7th edn. (Macmillan, New York, USA).

Burgess, T. M. and Webster, R. 1980. Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. Journal of Soil Science **31**, 315–331.

Burgess, T. M., Webster, R. and McBratney, A. B. 1981. Optimal interpolation and isarithmic mapping of soil properties. IV. Sampling strategy. Journal of Soil Science **32**, 643–659.

Burrough, P. A. 1989. Fuzzy mathematical methods for soil survey and land evaluation. Journal of Soil Science **40**, 477–492.

Chien, Y., Lee, D., Guo, H. and Houng, K. 1997. Geostatistical analysis of soil properties of mid-west Taiwan soils. Soil Science **162**, 291–298.

Doolittle, J. A., Sudduth, K. A., Kitchen, N. R. and Indorante, S. J. 1994. Estimating depths to claypans using electromagnetic induction methods. Journal of Soil and Water Conservation **49**, 572–575.

Eigenberg, R. A., Doran, J. W., Nienaber, J. A., Ferguson, R. B. and Woodbury, B. L. 2002. Electrical conductivity monitoring of soil condition and available N with animal manure and a cover crop. Agriculture, Ecosystems and Environment **88**, 183–193.

Heisel, T., Ersboll, A. K. and Andreasen, C. 1999. Weed mapping with co-kriging using soil properties. Precision Agriculture **1**, 39–52.

Irvin, B. J., Ventura, S. J. and Slater, B. K. 1997. Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin. Geoderma **77**, 137–154.

Jaynes, D. B. 1996. Mapping the areal distribution of soil parameters with geophysical techniques. In: Application of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone SSSA Special Publication, edited by D. L. K. Corwin Loague (48 SSSA Madison WI, USA), p. 205–216.

Jaynes, D. B., Colvin, T. S. and Ambuel, J. 1995. Yield mapping by electromagnetic induction. In: Larson Proceedings of Site-Specific Management for Agricultural Systems: Second International Conference, edited by P. C. R. H. W. E. Robert Rust (ASA-CSSA-SSSA, Madison, WI, USA), p. 383–394.

Jensen J. R. 1996. *Introductory digital image processing: a remote sensing perspective*. 2nd edn. (Prentice Hall, Upper Saddle River, N.J., USA).

Journel A. G. and Huijbregts C.J. 1978. *Mining geostatistics*. (Academic Press, New York, USA).

Kachanoski, R. G., Gregorich, E. G. and Wesenbeeck, I. J.Van 1988. Estimating spatial variations of soil water content using noncontacting electromagnetic inductive methods. Canadian Journal of Soil Science **68**, 715–722.

Kitchen, N. R., Sudduth, K. A. and Drummond, S. T. 1996. Mapping of sand deposition from 1993 midwest floods with electromagnetic induction measurements. Journal of Soil and Water Conservation **51**, 336–340.

Kravchenko, A. N. and Bullock, D. G. 2002. Spatial variability of soybean quality data as a function of field topography: I. Spatial Data Analysis. Crop Science **42**, 804–815.

Leenaers, H., Okx, J. P. and Burrough, P. A. 1990. Employing elevation data for efficient mapping of soil pollution on floodplains. Soil Use and Management **6**, 105–114.

McBratney, A. B. and Webster, R. 1983a. Optimal interpolation and isarithmic mapping of soil properties. V. Co-regionalization and multiple sampling strategy. Journal of Soil Science **34**, 137–162.

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

McBratney, A. B. and Moore, A. W. 1985. Application of fuzzy sets to climatic classification. Agricultural and Forest Meteorology **35**, 165–185.

McBratney, A. B. and de Gruijter, J. J. 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. Journal of Soil Science **43**, 159–175.

McIntyre, G. A. 1952. A method for unbiased selective sampling, using ranked sets. Australian Journal of Agricultural Research **3**, 385–390.

McNeill J. D. 1980a. *Electrical conductivity of soils and rocks*. Technical Note TN-5. (Geonics, Ltd., Mississauga, Ontario, Canada).

McNeill J. D. 1980b. *Electrical conductivity of soils and rocks*. Technical Note TN-6. (Geonics, Ltd., Mississauga, Ontario, Canada).

Odeh, I. O. A., McBratney, A. B. and Chittleborough, D. J. 1992. Soil pattern recognition with fuzzy-c-means: Application to classification and soil-landform interrelationships. Soil Science Society of America Journal **56**, 505–516.

Oliver, M. A. 1987. Geostatistics and its application to soil science. Soil Use and Management **3**, 8–20.

Patil, G. P., Sinha, A. K. and Taillie, C. 1994. Ranked set sampling. In: Environmental Statistics, Handbook of Statistics, edited by G. P. C. R. Patil Rao (Elsevier Science, Amsterdam Netherlands), p. 167–200.

Sheets, K. R. and Hendrickx, J. M. H. 1995. Noninvasive soil water content measurement using electromagnetic induction. Water Resources Research **31**, 2401–2409.

Trangmar, B. B., Yost, R. S. and Uehara, G. 1986. Spatial dependence and interpolation of soil properties in West Sumatra, Indonesia: II. Co-regionalization and co-kriging. Soil Science Society of America Journal **50**, 1396–1400.

Triantafilis, J., Odeh, I. O. A. and McBratney, A. B. 2001. Five geostatistical models to predict soil salinity from electromagnetic induction data across irrigated cotton. Soil Science Society of America Journal **65**, 869–878.

Vauclin, M., Vieira, S. R., Vachaud, G. and Nielsen, D. R. 1983. The use of cokriging with limited field soil observations. Soil Science Society of America Journal **47**, 175–184.

Vaughan, P. J., Lesch, S. M., Corwin, D. L. and Cone, D. G. 1995. Water content effect on soil salinity prediction: a geostatistical study using cokriging. Soil Science Society of America Journal **59**, 1146–1156.

Warrick A. W., Myers D. E. and Nielsen D. R. 1986. Geostatistical methods applied to soil science, In *Methods of soil analysis, Part 1*. edited by A. Klute, 2nd edn. Vol. 9. Agronomy p. 53–82.

Webster R. and Oliver M. A., 2001. *Geostatistics for environmental scientists*. (John Wiley and Sons, Ltd., Chichester England).

Williams, B. G. and Hoey, D. 1987. The use of electromagnetic induction to detect the spatial variability of the salt and clay contents of soils. Australian Journal of Soil Research **25**, 21–27.

Yates, S. R. and Warrick, A. W. 1987. Estimating soil water content using cokriging. Soil Science Society of America Journal **51**, 23–30.

Zhang, R., Myers, D. E. and Warrick, A. W. 1992. Estimation of the spatial distribution of soil chemicals using pseudo-cross-variograms. Soil Science Society of America Journal **56**, 1444–1452.