Plant and Soil **258:** 269–280, 2004.

Relationships between soil bulk electrical conductivity and the principal component analysis of topography and soil fertility values

S.J. Officer[1], A. Kravchenko[2], G.A. Bollero[1,5], K.A. Sudduth[3], N.R. Kitchen[3], W.J. Wiebold[4], H.L. Palm[4] & D.G. Bullock[1]

[1]*Department of Crop Sciences, University of Illinois, Urbana, IL 61801 U.S.A.* [2]*Department of Crops and Soil Sciences Michigan State University East Lansing, MI 48824 U.S.A.* [3]*USDA Agricultural Research Service, Cropping Systems and Water Quality Research Unit University of Missouri, Columbia, MO 65211 U.S.A.* [4]*Agronomy Department, University of Missouri, Columbia, MO 65211 U.S.A.* [5]*Corresponding author**

Received 16 December 2002. Accepted in revised form 4 September 2003

Key words: principal components, site specific management, soil electrical conductivity

Abstract

Measures of soil electrical conductivity (EC) and elevation are relatively inexpensive to collect and result in dense data sets which allow for mapping with limited interpolation. Conversely, soil fertility information is expensive to collect so that relatively few samples are taken and mapping requires extensive interpolation with large estimation errors, resulting in limited usefulness for site-specific applications in precision agriculture. Principal component (PC) analysis and cokriging can be applied to create meaningful field scale summaries of groups of attributes and to decrease the estimation error of maps of the summarized attributes. Deep (0–90 cm) and shallow (0–30 cm) EC, elevation, and soil fertility attributes were measured in fields under corn (*Zea mays* L.) and soybean (*Glycine max* L.) rotations, at two sites in Illinois (IL) and two sites in Missouri (MO). Soil fertility and topography attributes were summarized by PC analysis. The first topography PC (TopoPC1) contrasted flow accumulation against elevation and curvature, to describe the main topographic pattern of the fields. The first soil fertility PC (SoilPC1) consistently grouped together cation exchange capacity (CEC), Ca, Mg, and organic matter (OM). SoilPC1 was well correlated to soil EC for all sites and cokriging with EC had higher r^2 in the crossvariogram models compared to ordinary kriging. The second and third soil fertility PCs (SoilPC2 and SoilPC3) were concerned with soil pH and P, and reflected historic land use patterns. Maps of SoilPC2 and SoilPC3 had little relationship to soil EC or topography and so could not be improved by cokriging.

Abbreviations: CEC – cation exchange capacity; EC – soil bulk electrical conductivity; OM – soil organic matter; PC – principal component; SoilPC – soil fertility principal component; Sph – spherical variogram function; TopoPC – topography principal component.

Introduction

Site specific management is a developing agriculture technology that, in part, combines global positioning systems (GPS) with variable application rate technology (VRT). The objective is to gather, analyze, and correctly interpret sufficient information to make accurate and economically viable site-specific management decisions. This is critical because the economic advantage and environmental benefits of the technology is dependent upon the quality of the information utilized (Bullock et al., 1998; Bullock and Bullock, 2000). Information on topography, soil chemical properties, and several years of yield are now typically available among farmers and practitioners in the US Midwest. A characteristic problem, however, is that the various types of information are rarely collected from the same locations in a field or at the same

*FAX No: 217-333-9474. E-mail: gbollero@uiuc.edu

sampling intensity. In particular, soil fertility information is usually discrete with much data collected from a small number of points, so that the information is likely to be insufficient for accurate construction of field-scale maps. However, if soil fertility information is correlated to a densely sampled variable and the pattern with which the two vary jointly with separation distance can be modeled, then the relationship can be used to more accurately describe soil fertility patterns. This process is termed cokriging and may significantly improve the quality of soil fertility information for precision agriculture (Goovaerts, 1998).

Thousands of measurements of *in situ* soil EC and elevation may be collected from a single field by mobile data collection systems, and very accurate information of the trends across a field can be obtained. Unlike many soil tests, *in situ* soil EC and elevation values are relatively inexpensive to collect and can be measured with very little or no disturbance of the soil. Measures of elevation can be made by land-based laser or via GPS systems. Mobile measuring systems for soil EC subject the soil to an electrical current and measure the EC of a relatively large volume of bulk soil. Systems using coulters attached to a traveling cart measure the change in EC between one set of coulters transmitting a current and a second set completing the circuit. The method integrates the soil EC for the distance between the coulters, and to a depth similar to the distance between the coulters. The measure is affected by soil moisture, clay content, soil temperature, and salinity, and has been found to be a useful measure of these properties as well as a surrogate measure of other soil properties with related spatial patterns (Jaynes, 1996; Sudduth et al., 1999; Johnson et al., 2001).

Cokriging of soil salinity with soil EC information has improved the accuracy of salinity maps (Pozdnyakova and Zhang, 1999; Triantafilis et al., 2001). Large temperature differentials and salinity effects are not expected in typical Midwestern cropping fields, so clay and water content are expected to be the main influences in spatial EC patterns (Sudduth et al., 1999).

A second characteristic of precision agriculture data is that groups of attributes are often related. For instance, P, K, Mg, Ca, pH, CEC, and OM are all measured from the same soil sample. In addition, topography data sets are generated when slope, curvature, and flow accumulation are derived from elevation maps. Therefore, it would be useful to summarize these related sets of attributes, preferably into one or two meaningful components before relating them to other attributes (Mallarino et al., 1996). The multivariate technique PC analysis can be applied to summarize soil fertility and topography data into new independent random variables.

Principal component analysis is a dimension reduction technique that takes correlated attributes, or variables, and identifies orthogonal linear recombinations (PCs) of the attributes that summarize the principal sources of variability in the data. The PC analysis method is established in soil science (Burrough and Webster, 1976) and has been used more recently to summarize large data sets gathered in soil quality research (Wander and Bollero, 1999; Maddonni et al., 1999). The value of a PC at a point can be mapped by kriging methods in the same way that an individual attribute such as soil P value can be mapped (Goovaerts, 1997). Cokriging can be applied if there is a correlation between a PC and a more densely sampled attribute, such as EC. We hypothesized that a combination of PC analysis and cokriging will create meaningful summaries of site-specific management information and improve the quality of the maps of the summarized attributes for site-specific applications.

Methods and materials

A survey of deep and shallow EC values was made using a Veris® technologies Model 3100 (Geoprobe® systems, Kansas)[1] in the fall of 1999 under unsaturated conditions (Sudduth et al., 1999). Between five and eleven thousand deep (0–90 cm) and shallow (0–30 cm) bulk soil EC values were collected from each of four fields in IL and MO. The IL sites were adjacent north and south 16-ha sections ('North Williams' and 'South Williams', respectively) within a 259-ha field at Williams Farm in McLean County, IL (location 40°18′11″ N, 88°32′36″ W). The MO sites were an approximately 14-ha field, 'Gvillo field' (39°14′2″ N, 92°8′44″ W) and an approximately 18-ha subsection of a field, 'Field One' (39°13′50″ N, 92°7′44″ W) near Centralia, MO.

The soils of the North and South Williams sites were predominantly fine-silty mixed mesic Typic Endoaquoll in the poorly drained depressions and fine illitic, mesic Mollic Hapludalf on the upper slopes. These soils formed in a relatively thin layer of loess over calcareous silty clay loam till (Windhorn, 1998). The MO soils were claypan soils, predominantly fine

Table 1. Descriptive statistics and variogram characteristics for electrical conductivity (EC) of the bulk soil for 0–300 mm (shallow) and 0–900 mm (deep) at four sites in two states

	South Williams, IL		North Williams, IL		Gvillo Field, MO		Field One, MO	
	Deep EC	Shallow EC	Deep EC	Shallow EC	Deep EC	Shallow EC	Deep EC	Shallow EC
Descriptive statistics								
Mean, mS m^{-1}	40.9	27.5	39.2	27.5	23.5	15.0	19.3	9.5
Std. dev., mS m^{-1}	11.3	6.4	11.8	7.8	11.3	7.1	2.7	8.1
Min, mS m^{-1}	7.7	12.3	10.7	12.0	2.6	1.5	2.3	1.6
Max. mS m^{-1}	74.5	48.5	74.9	51.9	68.6	44.9	45.1	19.2
Skewness	0.4	0.6	0.5	0.8	0.4	1.5	0.5	0.6
Variogram characteristics								
Model type†	Sph	Sph	Sph	Sph	Sph	Sph	Sph	Sph
Model fit (r^2)	0.94	0.91	0.92	0.98	0.99	0.98	0.93	0.93
Range (m)	68	76	84	69	350	428	162	133
Sill /(Nugget+Sill)	0.93	0.91	0.84	0.94	0.92	0.95	0.74	0.66
Cross-validation (r^2)‡	0.89	0.88	0.89	0.88	0.87	0.84	0.90	0.82
N	6823	6823	6612	6612	5076	5076	10806	10806

†Sph=Spherical function fitted to isotropic variogram.
‡Cross-validations were between predicted block kriged values and actual values at each point.

montmorillonitic, mesic Udollic Ochraqualf (Sudduth et al., 1999).

The soils were sampled in the fall of 1996 in IL in a modified grid pattern as described in Kravchenko et al. (1999) and Warrick et al. (1999) and in the fall of 1995 in MO in a grid pattern as described in Kitchen et al. (1999) and Sudduth et al. (1999). Soil sampling and extraction was carried out by commercial laboratories using standard measurement techniques (Kravchenko and Bullock, 2000; Sudduth et al., 1996). Phosphorus, K, Mg, Ca, CEC, pH, and OM were measured in all soil samples.

Regionalized variable theory was used to predict values of EC and elevation at unknown points (Goovaerts, 1997; Kravchenko and Bullock, 1999). Isotropic variograms were calculated for deep and shallow EC and elevation values at 15 m lags (Goovaerts, 1997). After a variogram model had been obtained, ordinary block kriging was applied to create a surface of mean estimates in 20-m^2 blocks.

Elevation was measured on a semi-regular grid of approximately 10-m or less using a Leica 500 RTK GPS system in IL and an Ashtech Z surveyor RTK GPS system attached to a four wheel ATV in MO. Ground slope, curvature and flow accumulation were calculated from the elevation map based on the 20-m^2 block size. Slope and curvature were calculated from elevation differences relative to the immediately surrounding block values. Flow accumulation was calculated as the sum of the total blocks in a field contributing overland water flow to a particular block (Kravchenko and Bullock, 2000).

The block values of EC and topography attributes were extracted that coincided with each soil fertility sampling point, creating multivariate subsets ranging from $n = 73$ at North Williams, to $n = 186$ from Gvillo. This process retained all the original soil fertility values, of which relatively few were available, and joined them with an appropriate average of the values of EC and topography that had been collected around each the soil sample point.

The subsets of topography and soil fertility data were summarized by PC analysis using the PRINCOMP procedure of SAS (SAS, 2000). Principal components were calculated based on the correlation matrix. Principal components with eigenvalues ≥ 1 were considered to have a significant contribution towards the explanation of total variation and thus retained, as suggested by Jolliffe (1986) and Khattree and Naik (2000). When the correlation matrix is used, all original variables have unit variances and thus it is not worth retaining any PC with variance (eigenvalue) < 1 because it contains less information than an original variable (Jolliffe, 1986). Each PC is a linear combination of the original variables. The eigenvector loading associated with each variable in this

Table 2. Descriptive statistics of the topography information estimated for each soil sample location from North and South Williams Fields, IL, and Gvillo Field and Field One, MO

	Elevation	Slope	Curvature	Flow accumulation
	M	degrees	10^{-2}zunits[†]	Σcells
	North Williams, IL			
Mean	237.4	1.10	0.00	9.1
Std.dev.	1.7	0.46	0.03	20.3
Min	233.5	0.24	−0.05	0.0
Max	240.5	2.15	0.09	105.0
Skewness	−0.7	0.2	0.3	3.5
	South Williams, IL			
Mean	233.2	1.23	0.01	7.7
Std.dev.	1.5	0.40	0.03	22.6
Min	230.5	0.52	−0.08	0.0
Max	237.0	2.27	0.08	133.0
Skewness	0.4	0.5	−0.04	4.8
	Gvillo Field, MO			
Mean	261.7	1.02	0.00	6.2
Std.dev.	2.1	0.58	0.03	15.9
Min	258.1	0.02	−0.08	0.0
Max	264.9	2.93	0.07	177.0
Skewness	0.1	0.70	−0.3	7.3
	Field One, MO			
Mean	263.5	0.39	−0.00	24.1
Std.dev.	0.6	0.15	0.01	68.9
Min	262.3	0.06	−0.03	0.0
Max	264.7	0.98	0.02	469.0
Skewness	0.2	0.70	−0.1	4.8

[†]The reasonably expected values of all curvature for a hilly area (moderate relief) may differ from about −0.5 to 0.5 (ArcView GIS 3.2 Help notes, Environmental Systems Research Institute Inc.)

linear combination represents the contribution of the original variable to the PC. In this study, variables with eigenvector loadings ≥ 0.3 were considered in the interpretation of the PC as suggested by Harris (2001). To prevent the undue influence of outliers, and place the emphasis on the median trend rather than the mean, soil fertility attributes with a skew greater than one were log transformed before analysis (Parkin and Robinson, 1992). The same rule was not applied to the topography data although flow accumulation was generally skewed, however the skew was too extensive to be corrected by a single transformation.

Ordinary point kriging and ordinary point cokriging were applied to estimate values of SoilPC1 on a 10 m grid at all four sites (Goovaerts, 1997; Deutsch and Journel, 1998). Ordinary cokriging was applied to the values of the first PC scores calculated at each soil fertility sampling position, and the estimated EC value at each soil sample point, and the raw EC data. The improvement in the quality of the soil fertility maps generated by cokriging as compared to kriging was measured using cross-validations between known SoilPC1 values and the predicted values.

Results and discussion

Electrical conductivity

At both North and South Williams the deep and shallow EC information was significantly correlated ($r = 0.82$, $P < 0.0001$) and both measures exhibited a very strong fit to a spherical variogram model with a range of approximately 70 m (Table 1). Cross-validations of the ordinary block kriging were very high, indicated excellent prediction of average values in each 20-m^2 block (Table 1).

At the MO sites, the mean EC values were lower than those found at North and South Williams. The linear correlation between deep and shallow EC was significant ($r = 0.72$ at Gvillo and $r = 0.73$ at Field One, $P < 0.0001$ (Table 1). Variogram model fits and cross-validations were again highly significant (Table 1). The strong spherical model fits and high cross-validations were in contrast to the high variability at short distances that is generally expected of soil exchangeable chemistry (Beckett and Webster, 1971).

Topography

The measures of elevation formed linear variograms and maps with very high cross-validations ($r^2 = 0.99$) for all sites (Table 2). The descriptive statistics of the topography subsets from the four sites were also similar, except that the Field One site was generally flatter, with a lower average slope and smaller range of elevation and curvature (Table 2).

For all sites, the topography PC analysis consistently formed two PCs, which together accounted for an average of 72% of the total variability (Table 3). Common elements in the first PC (TopoPC1) for all sites were that elevation and curvature were added together and contrasted by flow accumulation (loading had opposite sign), indicating that increasing elevation was usually associated with positive curvature (convex) and decreasing flow accumulation at the whole field scale.

Table 3. Principal component (PC) analysis of topography data extracted at each soil sample point. The results of the analysis at four sites are summarized as the variance (eigenvalue) accounted for by each PC (TopoPC1 and TopoPC2) for each site, the eigenvectorloadings of each variable in the PC, and correlations between each PC and soil electrical conductivity (EC)

	South Williams, IL		North Williams, IL		Gvillo Field, MO		Field One, MO	
	Topo-PC1	Topo-PC2	Topo-PC1	Topo-PC2	Topo-PC1	Topo-PC2	Topo-PC1	Topo-PC2
Eigenvalue	1.81	0.95	1.88	1.12	1.57	1.29	1.73	1.24
Prop. var.†	0.45	0.24	0.47	0.28	0.39	0.32	0.43	0.31
Cum. var.‡	–	0.69	–	0.75	–	0.71	–	0.74
PC composition								
Elevation	0.60	0.11	0.59	−0.28	0.66	−0.22	0.36	−0.70
Curvature	0.49	−0.64	0.56	0.20	0.44	0.52	0.57	−0.00
Slope	0.46	0.73	−0.16	0.88	−0.54	0.50	0.34	0.72
Flow accumulation	−0.43	0.22	−0.55	−0.34	−0.29	−0.65	−0.65	−0.02
Correlations between EC and topography PCs								
Shallow EC	−0.27*	0.15	−0.36**	−0.03	−0.13	0.001	0.22**	0.31
Deep EC	−0.37**	0.11	−0.51***	−0.17	0.22**	0.08	0.43***	0.30***

*, **, ***significance at the 0.05, 0.01 and 0.001 probability levels, respectively.
†Proportion of the total variance accounted for by each PC.
‡Cumulative proportion of the total variance accounted for.

Considering the results specifically by each site, at South Williams TopoPC1 had significant positive loadings for slope, elevation, and curvature and had a significant negative loading for flow accumulation. Therefore, TopoPC1 for this site could be described as an average of slope, elevation and curvature information, contrasted by flow accumulation. South Williams is characterized by a large central grassed drainage-way that bisects long slopes with small interfluves (Figure 1, Ia). Therefore, high values of slope were usually associated with high values of elevation and also positive curvature, while flow accumulation tended to be low at the same points. Similarly, low values of slope were associated with low elevation, negative curvature and high flow accumulation. TopoPC2 at South Williams was mainly concerned with curvature contrasted with slope. Specifically, Topopc2 contrasted areas of high slope and concavity, which were mainly on the edges of erosion rills, with areas of low slope and convexity, such as the interfluves.

For TopoPC1 at North Williams, curvature was added to elevation, and contrasted by flow accumulation. Slope was separated into the second PC (TopoPC2) where it had a significant contribution (Table 3). This site had broad, relatively flat areas on the upper elevations, with sloped central areas, leading to lower areas associated with a grassed drainage-way (Figure 1, Ia). Therefore, many points that had high values of curvature also had high values of elevation and low values of flow accumulation, and vice versa. Slope was contrasted against flow accumulation in TopoPC2. High scores of TopoPC2 represent areas in the field with pronounced slopes and low flow accumulation.

For TopoPC1 at Gvillo, curvature and elevation had significant positive loadings and were contrasted against slope. In TopoPC2, slope was instead added to curvature and these two were contrasted against flow accumulation. This site had a broad, relatively flat area on the eastern side, with sloped central areas leading to lower areas associated with a grassed drainage-way. The elevation rose again in the SW corner of the site (Figure 1, IIa). Therefore, many points that had low values of elevation were concave and had high values of slope and vice versa, describing the average trend for the field. In TopoPC2, decreasing slope was associated with negative curvature and increasing flow accumulation, characterizing the landform associated with a drainage-way running diagonally across the field.

At Field One, TopoPC1 was similar to South Williams (Table 3). Flow accumulation was contrasted against curvature, slope and elevation. TopoPC2, however, was different to South Williams, contrasting slope with elevation rather than curvature. Field One was gently sloping with a central depression forming a drainage-way, so that increasing flow accumulation

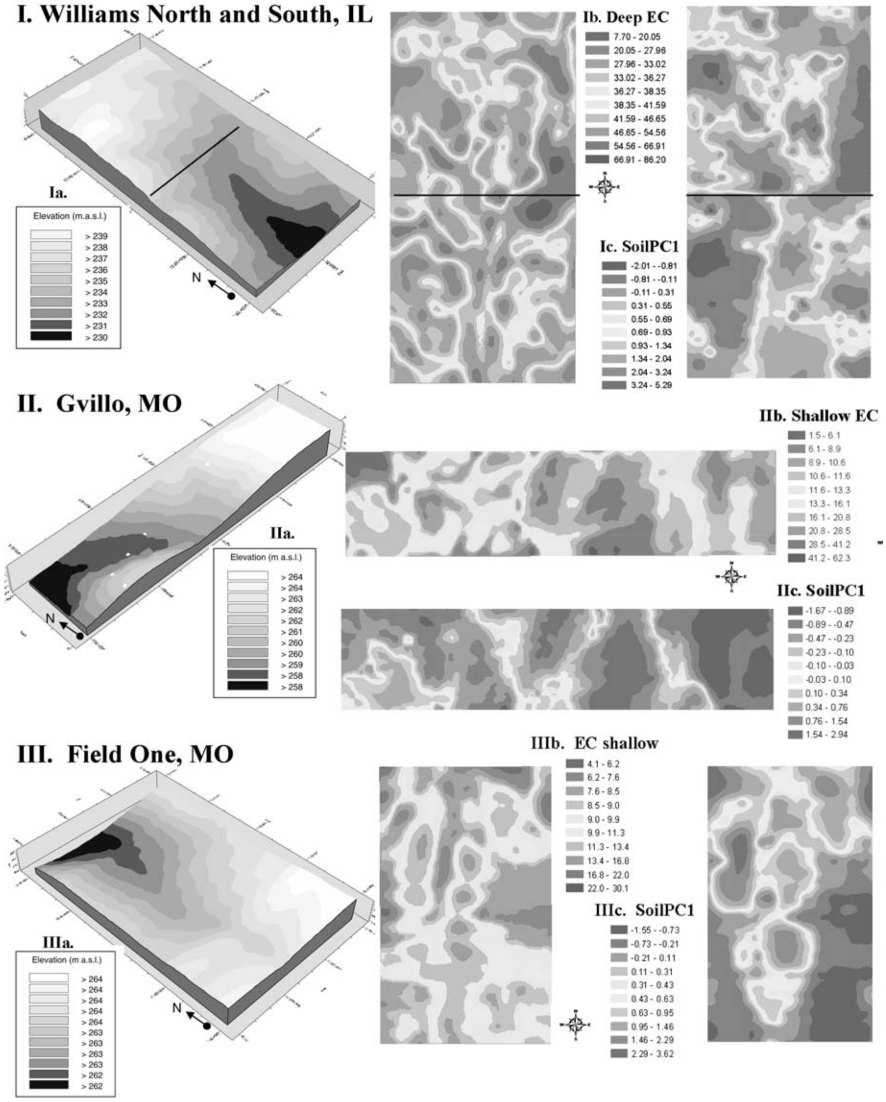

Figure 1. Elevation (a), electrical conductivity (EC) (b), and soil fertility principal component one (SoilPC1) (c) at North and South Williams (combined) (I), Gvillo (II) and Field One (III).

was particularly related to negative curvature, and to decreased slope and elevation to a lesser extent (Figure 1). The upper areas of the site were relatively flat, so that increasing elevation was related to decreasing slope and this secondary trend was sorted into TopoPC2.

The deep and shallow EC trends were compared to the topography patterns, using a linear correlation of the kriged block averages of EC and the values of each topography PC score (Table 3). TopoPC1 at both North and South Williams had significant negative correlations to soil EC. The negative correlation accorded with a general increase in EC value in the lower elevations and large depressions at these sites. Top soil clay content had a similar pattern at these sites (Omonode, 2001). At Gvillo, high EC values were associated with one sloped area and an associated interfluve, but the effect was not consistent in a second sloped area at this site (Figure 1, IIa). Topsoil depth to the claypan has previously been found to be an important determinant of clay content near the surface and therefore EC, in the MO claypan soils (Kitchen et al., 1999). The claypan is expected to be closer to the surface in sloped areas where the topsoils have been eroded, and therefore EC values increase on the sloped areas. This effect only occurred in one of two sloped areas

Table 4. Descriptive statistics of the soil fertility information from North and South Williams Fields, IL, and Gvillo Field and Field One, MO

	pH	CEC†	OM‡	P	K	Ca	Mg
		$cmol_c kg^{-1}$	%	kg ha^{-1}	kg ha^{-1}	kg ha^{-1}	kg ha^{-1}
North Williams Field, IL (n=73)							
Mean	6.2	24.7	3.2	37.6	389.3	5216	754.6
Std.dev.	6.2	17.3	3.2	12.3	74.4	893.6	113.5
Min	5.5	13.3	2.4	10.1	237.4	3769.3	532.0
Max	7.0	24.7	3.8	67.2	618.2	8094.2	1097.6
Skewness	0.4	0.9	−0.5	0.3	0.8	1.3	0.9
South Williams Field, IL (n=76)							
Mean	6.6	18.1	3.2	49.6	363.3	5996.4	739.9
Std.dev.	0.4	2.7	0.3	24.5	83.4	1004.5	189.9
Min	5.7	13.5	2.4	12.3	232.9	4213.4	487.2
Max	7.7	24.4	3.8	107.5	595.8	8648.6	1285.8
Skewness	0.8	0.3	−0.7	0.8	0.9	0.2	1.1
Gvillo Field, MO (n=186)							
Mean	6.7	15.9	2.5	36.4	276.5	5775.1	548.0
Std.dev.	0.3	3.4	0.3	16.3	67.8	1019.2	286.4
Min	5.7	10.2	1.6	10.1	179.2	4129.4	170.2
Max	7.3	25.0	3.4	95.2	471.5	9022.7	1351.8
Skewness	−0.9	0.9	0.0	0.9	1.0	0.8	1.2
Field One, MO (n=175)							
Mean	6.4	10.6	2.3	16.9	197.0	3248.4	260.7
Std.dev.	0.5	1.8	0.3	6.1	31.8	406.4	106.8
Min	5.0	8.3	1.6	7.0	146.7	2287.0	118.0
Max	7.1	18.0	3.3	42.0	325.9	4411.0	866.0
Skewness	−0.7	1.8	0.4	1.2	1.4	0.3	2.3

†Cation exchange capacity.
‡Organic matter.

at Gvillo, so the average topography did not correlate well to EC (Table 3).

At Field One, EC values increased around a drainage-way where topsoils have eroded (Figure 1, IIIb), bringing the claypan closer to the surface. Both shallow and deep EC patterns had positive correlations to TopoPC1 (Table 3), indicating that on average EC increased as curvature increased (became convex) and flow accumulation decreased.

Soil fertility

The soil fertility data from IL and MO had skewed distributions and wide value ranges, which are typical of agricultural soil exchangeable chemistry (Becket and Webster, 1971; Kravchenko and Bullock, 1999). The mean soil fertility values from the MO sites were lower than the IL values, and values for Field One were lower than Gvillo field (Table 4).

Principal component analysis of typical soil fertility data sets from four sites in two states provided similar patterns of results. Between 78 and 87% of the total variability of the soil fertility data from each site was accounted for by two or three PCs with eigenvalues ≥ 1 (Table 5). A strong first PC (SoilPC1), with eigenvalues greater than 2.5, was generated in each case, which accounted for an average of 49% of the total variability. SoilPC1 was consistently associated with CEC, OM, and the cations (Table 5). At all the sites, the second PC (SoilPC2) accounted for approximately 25% of variability and was mainly concerned with pH. The effect of soil P was added to pH for the MO sites but was contrasted with pH in the IL sites. A third PC was significant at the IL sites, which was mainly concerned with K.

The analysis revealed some broad similarities between the various sites, but also differences that re-

Table 5. Principal component (PC) analysis of soil fertility data. The results of the analysis at four sites are summarized as the variance (eigenvalue) accounted for by each PC (SoilPC1, SoilPC2 etc for each site), the eigenvectorloadings of each variable in the PC, and correlations between each PC, the electrical conductivity (EC) and the two topography PCs (TopoPC1 and TopoPC2) for each site

	South Williams, IL			North Williams, IL			Gvillo Field, MO		Field One,MO	
	Soil-PC1	Soil-PC2	Soil-PC3	Soil-PC1	Soil-PC2	Soil-PC3	Soil-PC1	Soil-PC2	Soil-PC1	Soil-PC2
Eigenvalue	3.5	1.44	1.16	2.51	1.62	1.33	4.17	1.46	3.68	1.75
Prop. var.†	0.50	0.21	0.17	0.36	0.23	0.19	0.60	0.21	0.53	0.25
Cum. var.‡	–	0.71	0.88	–	0.59	0.78	–	0.81	–	0.78
PC composition										
pH	0.24	−0.53	0.47	−0.08	−0.60	0.27	−0.01	0.73	−0.31	0.53
OM	0.31	0.37	−0.45	0.41	0.32	−0.26	0.35	−0.21	0.38	0.01
P	0.33	0.50	0.31	0.00	0.47	0.47	−0.17	0.56	−0.21	0.56
Ca	0.49	−0.16	−0.17	0.59	−0.13	0.04	0.45	0.29	0.23	0.59
Mg	0.44	−0.37	0.03	0.33	−0.44	0.41	0.47	−0.05	0.47	0.04
K	0.26	0.41	0.58	−0.08	0.30	0.68	0.44	0.15	0.43	0.22
CEC	0.48	−0.04	−0.32	0.60	0.1	0.04	0.48	0.02	0.50	0.07
Correlations between soil PCs, EC and topography PCs										
Shallow EC	0.51***	−0.27*	0.09	0.36**	−0.25*	0.22	0.81***	0.16*	0.53***	0.17**
Deep EC	0.60***	−0.06	−0.05	0.60***	−0.06	0.00	0.53***	0.01	0.51***	0.03
TopoPC1	−0.62***	−0.21	0.14	−0.50***	−0.05	−0.03	−0.23**	0.25***	0.01	0.07
TopoPC2	−0.05	0.15	0.17	−0.14	0.14	0.24	0.14	−0.33***	0.68***	−0.28**

*,** ,***significance at the 0.05, 0.01 and 0.001 probability levels, respectively.
†Proportion of the total variance accounted for by each PC.
‡Cumulative proportion of the total variance accounted for.

quire a more detailed explanation to understand the effects generating these results. At South Williams, all the soil fertility attributes had some contribution to the first PC (SoilPC1) (Table 5). This PC indicated that all the soil fertility attributes at this site tended to increase and decrease at the same points, although only half of the total variability was accounted for by this average trend. The first component was significantly correlated to soil EC and had a significant negative correlation to the first PC of the topographic information, TopoPC1 at South Williams. These relationships indicated that a general increase in soil fertility was associated with the lower elevations around the drainage-way that bisected the site. This pattern was also well described by soil EC pattern, which reflected a mutual underlying relationship to clay content for both EC and CEC (Johnson et al., 2001; Omonode, 2001).

The second soil fertility PC (SoilPC2) at South Williams accounted for an additional 21% of the total variability. Soil pH was the main contributor to this component, with a significant negative loading (Table 5). Magnesium also had a significant negative loading for SoilPC2 and these two were held in contrast to OM, P, and K. Up until about 1970, there had been an active cattle-feeding operation on this farm which generated substantial amounts of manure which were dumped close to a house site. We proposed this manure moved down the drainage-way, via erosion, raising the P and K levels in one corner of the South Williams site (Kravchenko et al., 1999). The increase in OM may be an incidental association with wet conditions in the drainage-way. At the same site, soil pH values increased markedly in an area where three trans-continental gas pipelines had been buried, and the underlying glacial till had been exposed to the surface. A general increase in soil Mg levels was also associated with this line of soil disturbance. Phosphorus, K, and OM values tended to be low in the same area. All these patterns were effectively summarized into SoilPC2 which, unremarkably, had little correlation to the topography or EC, because these patterns were mainly the result of artificial disturbance.

At South Williams, a third soil fertility PC (SoilPC3) had an eigenvalue greater than one and accounted for an additional 16% of the variability. SoilPC3 was composed of significant soil K, P, and pH contributions contrasted to OM. At Williams Field, differential management across large areas of the farm has created bands of higher K, P, and pH values, including the upper part of the South Williams site. Organic matter levels also tended to drop in this area and these effects were sorted into a third PC. Again, there was little correlation to the topography or EC, because these patterns were mainly the result of artificial disturbance.

At North Williams, the same causal effects operating on a different topography produced a PC analysis that was clearer in some respects. CEC, Ca, Mg, and OM were grouped together into SoilPC1, which accounted for 36% of the variability. SoilPC1 again had a significant positive correlation to deep EC and a significant negative correlation to TopoPC1. At North Williams, CEC, Ca, OM, and Mg levels tended to be high in the sloped areas associated with rill erosion and also adjacent to a grassed drainage-way. Rill erosion marks concave areas on the slopes where overland flow is accumulating and where the topsoils have been eroded, decreasing the depth to the glacial till layer. Soil OM increased in these areas, possibly because the increased flow accumulation interacted with the impeded drainage of the shallow profile, resulting in generally wetter soils and increased rates of organic matter accumulation. Prior to drainage, the same effect was operating in the larger depressions, creating a characteristic change to organic-matter-rich Typic Haplaquolls in undrained areas (Malo and Worcester, 1975). Soil EC was also higher in these areas, consistent with shallow profiles on the slopes and an increased clay content in flow-accumulating areas.

SoilPC2 at North Williams was similar to the same component for South Williams, and mainly concerned with pH (Table 5). Manure dumping and ground disturbance again created distinctive pH, P, and Mg patterns that were not related to the topography or EC patterns. SoilPC3 at North Williams was also similar to the South Williams third component, emphasizing soil K, but this time associated with soil Mg and soil P. While previous management practices have raised K levels in a band across the farm, changing soil types in this area and topography have created erratic patterns of association to the other nutrients. Once again, this third component had little correlation to soil EC or topography.

At the MO sites, PC analysis of the soil fertility data gave results that were similar to the IL sites. At Gvillo, SoilPC1 was an average of CEC, K, Ca, and Mg and also had a significant contribution from OM. The CEC and cations all increased markedly in one sloped area of the site, but not in second sloped area. Soil OM increased on all the sloped areas, and also in the lower elevations of the site.

The soil CEC, Ca, K, and Mg patterns were very similar to the EC patterns, and SoilPC1 had a highly significant correlation to both shallow and deep EC (Table 5). Exposure of the claypan into the topsoil markedly increases the clay content, increasing the CEC, and creating the associated increases in exchangeable Ca, K, and Mg. Unlike the IL sites, there was little correlation between SoilPC1 and the topography because increases in shallow EC and SoilPC1 were specific to only one of the two sloped areas (Figure 1, IIb).

SoilPC2 for Gvillo once again was mainly concerned with the soil pH, and also had a significant positive contribution from soil P. Phosphorus and pH values were high near the eastern end of the field. Old photographs of the site indicated that this area had been used for animal production in the past, so that manmade disturbance of the fertility patterns was again sorted into SoilPC2.

At Field One, SoilPC1 consisted of an average of CEC, K, Mg, and OM contrasted against pH and P (Table 5). CEC, K, Mg, and OM values increased in a lower sloped area of the field where pH and P levels were low. Like all the other sites, soil EC had a significant correlation to SoilPC1. At Field One, the CEC, K, and Mg trends corresponded to the exposure of the claypan into the topsoil in the sloped areas immediately surrounding the main drainage-way. In contrast to the other sites, SoilPC1 at Field One also had a highly significant positive correlation to TopoPC2, which was a contrast between increasing slope and decreasing elevation that also described the SoilPC1 pattern well.

Calcium had little contribution to SoilPC1 at Field One and instead was averaged with P and pH to form the second component. Differential lime application and historic animal production locations have created areas of high Ca, pH, and P values at the southern end of the field. Once again, there was little correlation between SoilPC2 and EC or topography at this site.

Co-kriging

The correlations between the SoilPC1 and EC indicated that the data intensive EC values at each site were suitable for co-kriging to improve the accuracy of maps of SoilPC1, as suggested by Goovaerts (1998). TopoPC1 was also correlated to SoilPC1 but not as consistently across all sites, so that cokriging with

Table 6. Variogram and crossvariogram models, and cross-validations for cokriging of the first soil fertility principal component (SoilPC1) and electrical conductivity (EC) at four sites. Variograms for SoilPC1, EC, and the crossvariogram, were modeled for four sites using linear combinations of nugget and spherical (Sph) functions where h is distance (m)

	South Williams, IL	North Williams, IL	Gvillo Field, MO	Field One, MO
Variogram for SoilPC1, γ_{PC1}	0.3+0.1Sph(h/67)+ 0.74Sph(h/182)	0.18+ 0.81Sph(h/91)	0.01+ 1.35Sph(h/334)	0.01+0.3Sph(h/67) +0.52Sph(h/152)
Variogram for EC, γ_{EC}	10+108Sph(h/67)+ 11Sph(h/182)	23+ 108Sph(h/91)	6+ 80Sph(h/334)	1.5+3.2Sph(h/67)+ 2.7Sph(h/152)
Cross-variogram, γ_{PC+Ec}	1.7+1Sph(h/67)+ 2.8Sph(h/182)	−2+ 6.5Sph(h/91)	0.1+ 9.2Sph(h/334)	0.01+ 1.12Sph(h/152)
Cross-validations between actual and predicted values (r^2)†				
Kriging	0.21	0.19	0.69	0.66
Cokriging	0.38	0.37	0.73	0.72

†Cross-validations are the regression between actual and predicted values for kriging alone, or cokriging with EC, for SoilPC1 at each site.

TopoPC1 or elevation alone was not as effective as using EC. A linear combination of nugget and spherical models was applied to model the SoilPC1 variogram, the EC variogram (deep EC for the IL sites and shallow EC for the MO sites), and the crossvariogram within the required constraints (Table 6) (Goovaerts, 1997). Values of SoilPC1 were then estimated on a 10 m grid at each site (Figure 1, Ic IIc IIIc).

The cross-validations indicated that cokriging improved the accuracy of the prediction of the SoilPC1 values in all cases. Similar effects were found by Pozdnyakova and Zhang (1999), where cokriging of the soil sodium adsorption ratio using soil EC values markedly improved the prediction of the adsorption ratio values. Cross-validations between known and estimated SoilPC1 values nearly doubled at the IL sites, where less than 72 SoilPC1 values were known over 16 ha (Table 6). The improvement in prediction was more limited in the MO fields where more than 173 SoilPC1 values were known over areas of 14 and 18 ha. Intensively measured EC values are a useful tool for improving the quality of maps of SoiPC1 where few values are known, although the meaning of SoilPC1 and the exact fit to the EC pattern will be specific to the site and probably also the soil moisture conditions at the time of EC measurement (Nugteren, 2000). Cokriging with intensively measured Soil EC values will also be useful to improve the accuracy of maps of the individual soil fertility attributes that were summarized in SoilPC1, such as maps of soil organic matter.

Interpretative summary

Principal component analysis was applied to the subsets of topography and soil fertility attributes, and successfully formed summary components that had a consistently meaningful interpretation. The average topography pattern of a typical field in a gently sloping glacial till landscape was broken into two PCs that captured 73% of the variability. Increasing elevation and positive curvature (convexity) were associated with decreasing flow accumulation at most sites to form the first topography PC (TopoPC1). The second topography PC (TopoPC2) was often concerned with slope, sorted as a secondary pattern because areas of greater slope tended to occur in areas of medium elevation, flow accumulation, and curvature.

Principal component analysis of the soil fertility data grouped the data into two or three components that explained between 78 and 87% of the total variability. The CEC, Ca, Mg, and OM fertility values were usually grouped together to form the first soil fertility PC (SoilPC1). Soil EC was a highly correlated to values of SoilPC1, and more consistently related to SoilPC1 than the topography PCs. At the IL sites, rill erosion has exposed the underlying glacial till into the topsoil in concave areas on the slopes, creating areas of high SoilPC1 and EC values. There were also high values of SoilPC1 in the larger depressions around drainage-ways, reflecting a historic build up of soil fertility in undrained depressions. At the MO sites, exposure of the claypan in eroded areas on the slopes was the dominant effect, causing sharp increases in SoilPC1 and EC. Therefore, soil EC was significantly correlated to SoilPC1 at all sites, although the factors generating the EC patterns and relationships between soil fertility, EC, and topography were different for the two states.

The second soil fertility PC (SoilPC2) was concerned with pH at all sites and usually also P. A third component (SoilPC3) was associated with K at the IL sites. SoilPC2 and SoilPC3 reflected the historic man-made disturbance of the natural patterns, creating soil fertility trends that were not related to the topography or EC patterns. Land use history, particularly amalgamation of the current field from smaller areas that have had different uses, was the main determinant of the present day pH and P patterns. Land use history was also an important determinant of the K pattern for the IL sites, but not the MO sites, reflecting the change in K retention properties between the soils in the two states.

Principal component analysis was concluded to be a useful way of summarizing soil fertility and topography information from precision agriculture systems. The method was very useful to sort information and identify sets of attributes with similar trends. Traditional soil fertility information is expensive to collect. Soil EC information is relatively inexpensive to collect and can be gathered in great intensity from a field, generating accurate maps. When a correlation was found between the soil fertility PCs and EC, cokriging with EC improved the accuracy of maps of the soil fertility patterns. The topography information was less useful. A combination of PC analysis and cokriging was concluded to be an effective method of summarizing precision agriculture information and increasing the quality of maps of traditional measures of soil fertility. When land management history determined the main patterns of nutrient, such as P or pH, there was little correlation to topography or EC so that the accuracy of maps of such attributes could not be improved by cokriging.

Acknowledgement

The authors thank the North Central Soybean Research Program and the United Soybean Board for funding support on this project.

Endnotes

[1]Mention of a trademark does not constitute a guarantee of, or warranty, of the product, nor does it imply an approval to the exclusion of other products that may be suitable.

References

Beckett P H T and Webster R 1971 Soil variability: A review. Soils Fert. 34, 1–15.

Bullock D S and Bullock D G 2000 From agronomic research to farm management guidelines: A primer on the economics of information and precision technology. Prec. Agric. 2, 71–101.

Bullock D G, Bullock D S, Nafziger E D, Doerge T A, Paszkiewicz S R, Carter P R and Peterson T A 1998 Does variable rate seeding of corn pay? Agron. J. 90, 830–836.

Burrough P A and Webster R 1976 Improving a reconnaissance soil classification by multivariate methods. J. Soil Sc. 27, 554–571.

Deutsch C V and Journel A G 1998 GSLIB, Geostatistical software library and users guide, 2nd edition. Oxford University Press, New York.

Goovaerts P 1997 Geostatistics for natural resources evaluation. Oxford University Press, New York.

Goovaerts P 1998 Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. Biol. Fertil. Soils 27, 315–334.

Harris R J 2001 A primer of multivariate statistics, 3rd ed. Lawrence Erlbaum Associates Inc., Mahwah, NJ.

Jaynes D B 1996 Improved soil mapping using electromagnetic induction surveys. *In* Precision agriculture. Proc. third int. conf. on precision agriculture, Minneapolis, MN. pp. 169–179.

Johnson C K, Doran J W, Duke H R, Wienhold B J, Eskridge K M and Shanahan J F 2001 Field-scale electrical conductivity mapping for delineating soil condition. Soil Sci. Soc. Am. J. 65, 1829–1837.

Johnson R A and Wichern D W 1998 Applied multivariate statistical analysis, 4th edition. Prentice Hall, Upper Saddle River, NJ.

Jolliffe I T 1986 Principal component analysis. Springer-Verlag, New York.

Khattree R and Naik D N 2000 Multivariate data reduction and discrimination with SAS software. SAS Institute Inc. Cary, NC.

Kitchen N R, Sudduth K A and Drummond S T 1999 Soil electrical conductivity as a crop productivity measure for claypan soils. J. Prod. Agric. 12, 607–617.

Kravchenko A N and Bullock D G 1999 A comparative study of interpolation methods for mapping soil properties. Agron. J. 91, 393–400.

Kravchenko A N and Bullock D G 2000 Correlation of corn and soybean grain yield with topography and soil properties. Agron. J. 92, 75–83.

Kravchenko A N, Boast C W and Bullock D G 1999 Multifractal analysis of soil spatial variability. Agron. J. 91, 1033–1041.

Maddonni G A, Urricariet S, Ghersa C M and Lavado R S 1999 Assessing soil quality in the rolling pampa, using soil properties and maize characteristics. Agron. J. 91, 280–287.

Mallarino A P, Oyarzabal E S and Hinz P N 1999 Interpreting within-field relationships between crop yields and soil and plant variables using factor analysis. Precision Agric. 1, 15–25.

Malo D D and Worcester B K 1975 Soil fertility and crop responses at selected landscape positions. Agron. J. 67, 397–401.

Nugteren W A, Malo D D, Schumacher T E, Schumacher J A, Carlson C G, Clay D E, Clay S A, Dalsted K J and Ellsbury M M 2000 Hillslope chronosequence of electromagnetic induction readings as influenced by selected soil properties. Agron. Abstr. 82 pp.

Omonode R A 2001 Soil spatial variability: structures, models, and their effects on crop yield variability in central Illinois. Ph.D. Thesis, University of Illinois at Urbana-Champaign.

Parkin T B and Robinson J A 1992 Analysis of lognormal data. Adv. Soil Sci. 20, 193–235.

Pozdnyakova L and Zhang R 1999 Geostatistical analyses of soil salinity in a large field. Precision Agric. 1, 152–165.

Harris R J 2001 A Primer of Multivariate Statistics, 3rd edition. Lawrence Erlbaum Associates Inc. Publishers, Mahwah, NJ.

SAS Institute 2000 SAS User's Guide. Statistical Analysis System Institute Inc. Cary, NC.

Sudduth K A, Drummond S T, Birrell S J and Kitchen N R 1996 Analysis of spatial factors influencing crop yield. *In* Proc. 3rd int. conf. on precision agric. ASA, CSSA, SSSA. Madison, WI. pp. 129–140.

Sudduth K A, Kitchen N R and Drummond S T 1999 Soil conductivity sensing on claypan soils: Comparison of electromagnetic induction and direct methods. *In* Proc. 4rd int. conf. on precision agric. ASA, CSSA, SSSA. St Paul, MN. pp. 129–140.

Triantafilis J I, Odeh O A and McBratney A B 2001 Five geostatistical models to predict soil salinity from electromagnetic induction data across irrigated cotton. Soil Sci. Soc. Am. J. 65, 869–878.

Wander M M and Bollero G A 1999 Soil quality assessment of tillage impacts in Illinois. Soil Sci. Soc. Am. J. 63, 691–971.

Warrick A W, Vargas-Guzman J A, Mayer S, Restrepo M and Ellsworth T R 1999 Strategies for improving spatial variability assessments. *In* Assessment of non-point source pollution in the vadose zone. Geophysical Monogr. 108, 93–105.

Windhorn R D 1998 Soil Survey of McLean County, Illinois. Illinois Agric. Experimental Stn. Soil Rep. 159.

Section editor: B.E. Clothier