THESIS

This is to certify that the

thesis entitled

Effect of Soil Properties and Landscape Features on a Michigan Corn and Soybean Cropping System

presented by

PINGPING JIANG

has been accepted towards fulfillment of the requirements for

M.S. degree in Crop & Soil Sciences

Kurt D. Thelen
Major professor

Date July 25, 2002

EFFECTS OF SOIL PROPERTIES AND LANDSCAPE FEATURES ON A MICHIGAN CORN AND SOYBEAN CROPPING SYSTEM

By

Pingping Jiang

A THESIS

Submitted to
Michigan State University
In partial fulfillment of the requirements
For the degree of

MASTER OF SCIENCE

Department of Crop and Soil Sciences

2002

ABSTRACT

EFFECTS OF SOIL PROPERTIES AND LANDSCAPE FEATURES ON A MICHIGAN CORN AND SOYBEAN CROPPING SYSTEM

By

Pingping Jiang

Field scale crop yield variability is partly caused by soil and topographic variation of the field. A 50 hectare corn-soybean field in Michigan was sampled. Two topographic features (elevation and slope) and twenty-three soil properties from the top two horizons were examined. Yield data were collected from 1996 through 2001. Correlation coefficients showed that topographic data were more frequently related to soil texture. The relationship of soil properties and field topography to crop yield varied by year and field. Elevation had no significant correlation with crop yield. The cumulative effect of both soil and topography explained approximately 30-60% of yield variability during the study period. However, a considerable amount of yield variability remained unexplained.

A sampling scheme that can ensure the desired accuracy in mapping soil properties at efficient cost is always of interests. A total number of 220 samples were taken, including 123 grid and 97 off-grid samples. A test data set of 31 samples was randomly selected for validation and the rest of the samples were used to design four sampling schemes. Geostatistical analysis was used to characterize spatial structure of soil P and K and to predict values to the test data locations. Sampling schemes with closely spaced samples performed better than the schemes with only grid samples. Less samples would be needed without compromising the prediction accuracy if a combination of grid samples and closely spaced off-grid samples was used.

ACKNOWLEDGEMTS

I would like to thank Dr. Fran Pierce for giving me this great and new opportunity of education and career at MSU. I sincerely appreciate Dr. Kurt Thelen, my major advisor, for his unreserved support and trust, which was a great help in ease of completing the degree. My appreciation is also sent to my committee members, Dr. Donald Christensen, Dr. Jiaguo Qi, for being like friends who always give sincere suggestions and encouragements, and Dr. Sasha Kravchenko, for her critical review on the Chapter 2 of this thesis.

I would like to thank Brian Long and Cal Bricker for their assistance in field.

My appreciation always flies to my parents, Jiang Lizheng and Li Jian, from the deepest corner of my heart, for their understanding of my pursuit, for being healthy and happy, and for being an endless source of strength and inspiration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 1

VARIABILITY OF SOIL AND LANDSCAPE PROPERTIES AND THEIR RELATIONSHIP TO CROP YIELD IN A MICHIGAN CORN-SOYBEAN FIELD

Abstract

Field scale crop yield variability is partly caused by soil and topographic variation. In certain years, soil properties and topographic features can explain as much as 60% of yield variability. The objectives of this study were to identify soil properties that cause most soil variation and investigate the complex interactions among crop yield, soil properties and field topography. A 50 hectare corn-soybean field in Michigan was sampled, and 23 soil properties from the top two horizons up to 50 cm deep were tested. Corn and soybean yield data were collected from 1996 through 2001 using a yield monitor mounted on the tractor. A multivariate statistical model, principle component analysis (PCA), was used to identify important soil properties. Eleven soil properties returned by PCA and two topographic variables, elevation and slope derived from elevation, were used to investigate their effect on crop yield. Correlation coefficients showed that topographic data were more frequently related to the textual components clay, silt and coarse sand contents. The relationship of soil properties and field topography to crop yield varied by year and field. Only clay content (B horizon), K (A and B horizons), NH_4-acetate base saturation (B horizon), pH (B horizon) had a significant and mostly consistent correlation with yield. Elevation had no significant correlation with yield. The cumulative effect of both soil and topography explained

approximately 30-60% of the yield variability during the study period. However, a considerable amount of yield variability remained unexplained.

Introduction

In agricultural fields, yield variability is partly caused by soil variability and varying topographic features of the field. Although yield is a function of a host of factors including climate, biological factors, soil properties, topography and management, in certain years as much as 60% or even more of the yield variability can be explained by a combination of soil properties and topographic features (Yang, et al., 1998; Kravchenko and Bullock, 2000). Soil properties affecting crop yield include physical and chemical properties such as soil available water, texture, bulk density, clay content (Stone, 1985, Wright et al, 1990), organic carbon (Ciha, 1984; Stone, 1985; Wright et al., 1990), pH (Kreznor et al., 1989; Moore et al., 1993), subsoil acidity (Wright et al., 1990), fertility and soil thickness (Kreznor et al., 1989).

Crop growth and yield are also affected by field landscape features such as elevation, slope steepness and aspect, and surface curvature (Ciha, 1984; Daniels et al., 1985; Stone et al., 1985; Simmons et al., 1989; Yang et al., 1998; Kravchenko and Bullock, 2000). Field topography can have a direct effect on crop growth and yield by redirecting and changing soil water availability, and an indirect effect through its influence on certain soil chemical and physical properties such as organic matter content, base saturation, soil temperature, and particle size distribution (Franzmeier et al., 1969; Bennett et al., 1972; Stone, 1985). Walker et al. (1968) investigated a series of soil properties such as A horizon thickness, distance from surface to mottles, and to carbonate and manganese segregations. They concluded that slope positions were most strongly related to these soil properties. In addition, the authors reported that different erosion classes caused by different slope positions could affect the speed of development and change of certain soil properties such as clay content and organic matter. Ovalles and Collins (1986) conducted a study on a broad selection of soil chemical and structural properties including pH, organic carbon, total P, coarse sand, medium sand, fine sand, very fine sand, total sand, silt and clay content from three topographic positions of summit, shoulder and backslope in north central Florida. They demonstrated that all these selected soil properties had a significant dependence on the topographic positions of the field. Daniels et al. (1985) reported a close relationship between topographic position and erosion classes, which both have been shown to influence soil properties, and hence soil productivity (Sinai, 1981; Ciha, 1984; Stone et al., 1985).

Field-scale yield variability is affected by these soil properties and the specific topography of the field. Previous studies have shown that soil properties and topographic features effect yield differently across locations and years. Fiez et al. (1994) pointed out that due to the inconsistent effect of topographic features on crop yield, research was needed to determine and investigate the soil factors that vary with field topography. This would result in a better understanding of yield variability under varying soil conditions and field topography. Wright et al. (1990) found that soil properties like texture, fertility, organic matter, and available water have a significant effect on corn yield. A study by Yang et al. (1998) showed that topographic features (elevation, slope and aspect) alone can explain 15-35% of wheat yield variability at the whole-field scale. In addition, they reported that topographic features account for 49-84% of the yield variability in some areas of the field. Higher wheat yields were generally found at lower elevation and gentle slope positions. Lower wheat yields were associated with higher elevation level and steep slope positions. Miller et al. (1988) found that although sand and clay content vary with slope position, percent sand had no correlation with yield, whereas percent clay was negatively correlated with total above ground biomass and positively correlated with harvest index ratio (ratio of grain wt/total above ground biomass × 100). They also found that the total above ground biomass increased from the knolls to the swales. It also has been reported that topographic influence on yield can be enhanced by extreme weather conditions. For example, severe early season drought could stress the effect of elevation on the yield (Simmons, 1989). Kravchenko and Bullock (2000) also found that in general, the greatest effect of topography was observed during extremely dry or wet years and at locations with extreme topography. For average weather conditions and moderate topographic position, this effect was relatively small.

Soil information allows one to investigate part (sometimes, a large part) of the yield variability. Due to the complex nature of soil and complicated interactions between soil components, a multivariate analysis model is necessary to unbiasedly assess variability within soils covariance structure. Principle component analysis (PCA) is one of the multivariate analyses models that has been employed to investigate the covariance structures of soil and identify important and differentiating soil variables. PCA uses linear combinations of soil variables to explain the covariance structure of a data set with a large number of soil variables. It reduces the dimension of the original data set without losing substantial information and often reveals relationships that were not previously suspected and thereby allows new interpretations or further analysis (Johnson and Wichern, 2000). Richardson and Bigler (1984) used this model and found that soil electrical conductivity (EC) and soluble Mg and Na were the most important variables in explaining observable differences in wetland soils among clay content, pH, organic carbon, CaCO3 equivalent (CCE), EC and Mg, Ca and Na. In another soil study using PCA, Ovalles and Collins (1988) found that the first five principle components explained more than 73% of the total variance. Total sand, fine sand, clay and OC contents were selected by the PCA as important properties.

The objectives of this study were to:

1) Determine the important soil properties in explaining the variation of a wide selection of soil properties;
2) examine the correlation between these important soil properties and field topographic data;
3) Investigate the relationship between crop yield variability and these soil properties under the specific field topography.

Materials and Methods

Study Site Description

The study site was a corn (*Zea mays* L.) – soybean [*Glycine max* (L.) Merr.] field located in Kalamazoo County, Michigan (42.22 N, 86.36 W). The entire study site is approximately 50 hectare in size and was divided into two sub-fields – Field 1(east side) and Field 2 (west side). The two sub-fields were planted to corn and soybean in an alternate year rotation (Fig.1.1). The elevation of the study site ranges from approximately 288 to 303 m with near level (0-2%), gentle (2-6%), and moderate (6-12%) slopes (Fig. 1.2). The soil texture class of the study site is predominately

Kalamazoo (fine-loamy, mixed, mesic Typic hapludalfs) with Oshtemo (coarse-loamy, mixed, mesic typic Hapludalfs) at the south-west corner of the study site (Soil Survey of Kalamazoo County, Michigan, 1979). The site was non-irrigated and a minimum tillage system was employed. During the study period, Field 1 was planted with soybean in odd numbered years (1997, and 1999), and with corn in even numbered years (1996, 1998, and 2000). This sequence was rotated in Field 2 except corn was planted in both sub-fields in 2001. Weather data were acquired from an adjacent weather station. Figure 1.3a shows the monthly precipitation during the1996-2001 growing seasons and Figure 1.3b depicts the departure of precipitation from normal levels.

Figure 1.1. Sample locations. Seventeen soil pedons from Field 1 and 16 from Field 2 were taken.

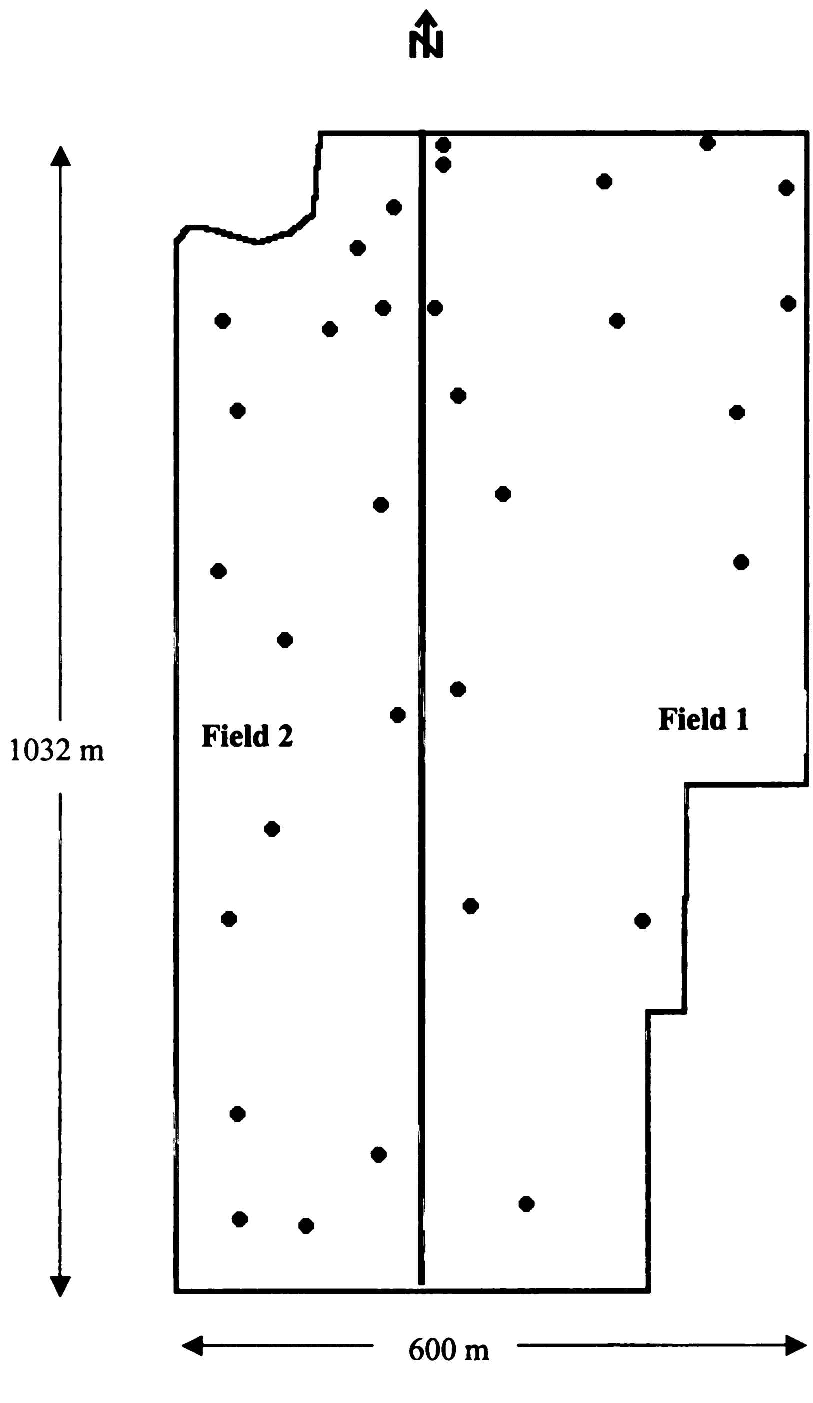
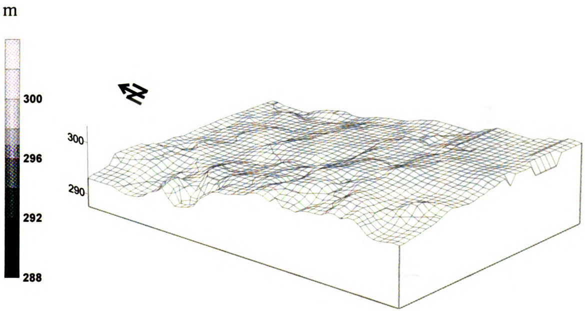


Figure 1.2. Topography of the study site.



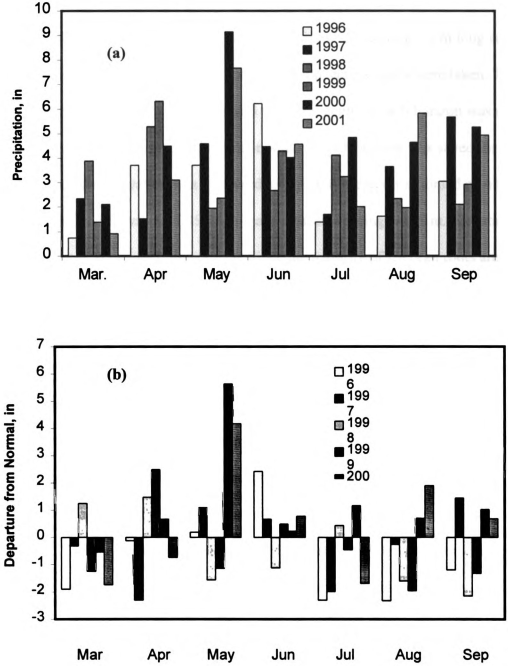


Figure 1.3a. Growing season precipitation during 1996-2001; Figure 1.3b, Departure from normal† growing season precipitation during 1996-2001.

† Average of 1971-2000.

Field Sample Collection and Laboratory Analyses

Thirty-three soil pedons were taken on May 1, 2001. In Field 1(east), 16 samples and in Field 2 (west) 17 samples were taken. The soil cores were 1.2 m long and 4 cm in diameter. Horizon designation was done as the soil sample cores were taken. The number of horizons of soil pedons ranged from 3 to 6. The depth of each horizon was recorded. A 5 to 10 cm undisturbed section from the middle of each horizon was saved for bulk density analysis. Elevation was measured using a GPS receiver mounted on a tractor. The speed of the tractor was 12.8-19.2 km/h and transect spacing was 9 m. Elevation data were collected at one second intervals. Since the majority of soybean roots are found in the upper 30 cm of soil, with a dominantly large proportion of the root mass in the topmost 16 cm, and corn roots are found in the upper 35-40 cm, only the top two horizons were analyzed (the top two horizons are referred to as A and B horizon for convenience, although sometimes the second horizon was designated as A2 or transitional BE soil taxonomy horizons). The depth to the lower border of the B horizon ranged from 35 to 85 cm with the average depth of 51 cm.

Laboratory analyses were conducted by the Soil Testing Lab at the University of Missouri. The parameters analyzed included soil texture by the pipette method; cations (K^+, Mg^{2+}, Ca^{2+}, Na^+) by ammonium acetate extraction methods; CEC by base summation + H ion; pH by soil solution ratio of 1:1; total C and organic C by emission method. More than 30 soil properties were determined. Since some of the properties represented different measures of the same parameter, only 23 soil variables were selected for analysis (Table 1.1).

Data Retrieved in GIS Environment

Yield Data

Six year yield data (from 1996 through 2001) were recorded using a commercial yield monitor mounted on a combine. Yield data were recorded every second as the combine proceeded along the transects with the swath width being the transect spacing. Latitude and longitude were recorded simultaneously by a GPS receiver for each yield data point. Yield data were thus incorporated as an information layer with other soil and topographic information, which enabled quantitative correlation between yield and field parameters in a GIS environment. Questionable and unrealistic yield data points possibly caused by significant positional errors, operating errors such as abrupt changes of speed, partial swath entering the combine, combine stops and starts, were removed from the dataset before any statistical analyses. In GIS Arcview (ESRI, 1996), the yield values for the sample locations were calculated by averaging the yield points that were contained in the circle area around each sample location with a diameter of 6.7 m. Due to the malfunction of the yield monitor or unstable GPS signals during the harvest, only corn yield data from 1996, 1998, 2000 and 2001, and soybean data from 1997 and 1999 were available in Field 1, and corn yield data from 1997,1999 and 2001 were available and no soybean data available in Field 2 during the study period.

Slope Derivation

Slope is the first order derivative of elevation and was derived using the Spatial Analyst of Arcview GIS. This process uses a 3 × 3 cell neighborhood around the processing or center cell and the average maximum technique to calculate slope values. It identifies the maximum rate of change in value from each cell to its neighbors (Manual of Spatial Analyst, 1996). The slopes were measured in degrees.

Data Analysis Procedures

Principle component analysis (PCA) was employed to investigate soil variance and identify important soil variables to be used as inputs for further analysis. Calculated PCs were in the following form:

$$\mathbf{Y}_i = \mathbf{a}_{i1}\mathbf{X}_1 + \mathbf{a}_{i2}\mathbf{X}_2 + \ldots. + \mathbf{a}_{ip}\mathbf{X}p,$$

Where Y_i is the ith principle component (PC), X_1, …. , Xp are the original variables, and $a_{i1,\ldots}$, a_{ip} are the coefficients of the ith PC and also an index of relative importance of the variable to that PC. The greater the absolute values of a_i are, the greater the importance of that soil variable to the PC. We termed this value as variable loading, as referred to by the statistical software package used (S-PLUS; Mathsoft, 1999). A sample correlation matrix was used instead of a covariance matrix in the PC calculation due to the differences of the order of magnitude between soil variables measured. The PCs that cumulatively explained more than 85% of the sample variance and individually explained more than 5% of the total variance were empirically retained for further processing. Variable loadings > 0.4 were considered to be important in the given PC and were hypothesized to have greater potential to explain yield variability. These variables were then used in further statistical analyses. Pearson correlation coefficients (r) were calculated between the important soil variables identified by PCA and topographic data, and yield data, and between topographic data and yield data. Multiple linear regression was employed to analyze the combined effect of both soil properties and topographic data on crop yield. A forward stepwise regression procedure was used. All statistical analyses were conducted using S-PLUS statistical package (Mathsoft Inc., 1999).

Results and Discussions

Soil Variation and Important Soil Variables

Principle component analysis was performed for the first (A) and second (B) horizons respectively. Table 1.1 shows the summary statistics of the 23 soil variables of each of the top two horizons which were evaluated by PCA. These soil variables included soil texture variables (total clay, fine silt, coarse silt, very find sand, medium sand, coarse sand, very coarse sand); soil chemistry/fertility (CEC, base saturation of NH4-Acetate, exchangeable Ca^{2+}, Mg^{2+}, K^{+}, sum of basis, pH (0.01 M $CaCl_2$), acidity, organic carbon and total N); and soil physical variables including electrical conductivity (EC), bulk density and horizon thickness.

Table 1.1. Selected variables analyzed by PCA and their summary statistics.

	HT†	Bulk Density	Total Clay	Total Silt	Total Sand	Fine Silt	Coarse Silt	VF‡ Sand	Fine Sand	Med Sand	VC§ Sand	Coarse Sand
	cm	g/cm^3	----------	----------	----------	----------	%	----------	----------	----------	----------	----------
						A Horizon						
Mean	24.0	1.4	13.6	47.7	38.6	29.2	18.6	10.0	17.6	7.0	2.9	1.2
S.D.	3.8	0.15	3.93	12.56	15.86	9.56	4.70	2.40	11.41	3.26	1.11	1.15
						B Horizon						
Mean	27.3	1.5	18.2	40.1	41.7	24.0	16.1	10.3	20.2	7.4	2.9	1.0
S.D.	10.2	0.16	5.87	15.66	18.23	12.67	4.94	3.00	13.15	4.01	1.76	1.09

Table 1.1 (cont'd).

	Ca	Mg	K	Sum Basis	Acidity	CEC	Base Sat.	OC¶	Total N	pH#	EC
	----------	----------	meq/100g	----------	----------	----------	----------	----------	%----------		mmho/cm
					A Horizon						
Mean	5.9	1.2	0.4	7.5	8.0	10.8	68.8	1.3	0.1	5.4	0.4
S.D.	2.33	0.45	0.22	2.69	3.05	3.24	14.84	0.39	0.05	0.54	0.15
					B Horizon						
Mean	5.7	1.2	0.3	7.2	8.1	12.0	61.7	0.6	0.1	4.9	0.5
S.D.	1.99	0.52	0.18	2.41	3.34	3.71	14.72	0.47	0.04	0.54	0.52

† Horizon thickness
‡ Very fine sand
§ Very coarse sand
¶ Organic carbon
# pH in 0.01M $CaCl_2$

The PCA results are presented in Table 1.2. In the A horizon, the first five principle components (PCs) accumulatively explained 86.2% of the total sample variance, and each of them explained more than 5% of the total sample variance (Table 1.3). PC1 explained 44.3% of the total sample variance of that horizon with most of the soil variables linearly combined except for horizon thickness, very coarse sand content, extractable Mg, base saturation, pH and soil EC. Variable loading was established at 0.4 to be considered significant. The results showed that the variable loadings were generally all less than 0.4 indicating that none of these soil variables were of absolute importance, or dominant in the PC. However, soil texture variables such as total sand (-0.304), total silt (0.295), medium sand (-0.287), fine silt (0.283) suggest that soil texture variables were relatively more important in this PC than soil variables from other categories. A distinct contrast was observed between the coarse materials of various sized sands and the fine materials such as clay and silts. This contrast was differentiated by the opposite sign of the loadings (Table 1.2). In PC1 of the B horizon, a very similar situation can be seen in terms of the relative importance of variable loadings (Tables 1.2 and 1.3). Again, most of the soil variables were linearly combined, except for extractable Mg, base saturation and pH. Similar soil texture variables (sand, fine sand, fine silt) were relatively high in the loadings. The contrasting relationship once again appeared between coarse materials and fine materials, and only the signs were opposite with those in the A horizon. Therefore, PC1 of both A and B horizons can be referred to as a "texture component". In PC2, both the A and B horizons had fewer soil variables linearly combined than in PC1. Again, all the variable loadings were below 0.4 with the only exception of base saturation (-0.424) in the B horizon. In PC2, higher loadings were represented by chemistry/fertility variables as opposed to primarily soil texture variables in PC1. For example, base saturation (0.389), sum of bases (0.370), extractable Ca (0.357), and pH (0.341) were carrying the highest loadings in the A horizon, and base saturation (0.424), Mg (0.393), sum of bases (0.366), pH (0.333) were the highest in the B horizon. In PC2, both the A and B horizons had very similar soil chemistry/fertility variables that appeared in first positions. Additionally, these fertility/chemistry variables all had a positive relationship in the PC. Therefore, we can refer to this PC as the "fertility component". The "fertility component" of the A horizon explained 23% of the total sample variance and that of the B horizon explained 16 % of the total sample variance. In sum, PC1 - "Texture component" and PC2 - "fertility component" have accumulatively accounted for 67% and 60% of the total sample variance of the A and B horizons, respectively. This implied that up to a depth of 50 cm (average depth of A and B horizons), 60% or more of the total sample variance can be explained by the first two PCs represented primarily by soil texture and fertility variables. Richardson and Bigler (1984) in their wetland soil study using PCA reported that PC1 was "chemical potentiality" dominated by EC and soluble Na and Mg, PC2 was dominated by organic carbon and PC3 was dominated by soil texture. The difference between theirs and this study showed that results of PC analysis on soil properties are data dependent and related to soil forming processes and field management history.

Table 1.2. PCA results of the first five PCs with variable loadings.

	PC1	PC2	PC3	PC4	PC5
			A Horizon		
HT†		-0.199	-0.155		-0.675
Bulk Density	-0.220				-0.482
Total Clay	0.252			-0.376	
Total Silt	0.295		-0.180		
Total Sand	-0.304		0.137		
Fine Silt	0.283	-0.129		0.142	
Coarse Silt	0.142	0.192	-0.469	-0.151	
VF Sand‡	-0.198	0.156	-0.263	-0.301	0.255
Fine Sand	-0.285		0.163	0.174	
Med.Sand	-0.287		0.194		
Coarse Sand	-0.152	0.222	0.222	-0.213	0.136
VC Sand§		0.297	0.262	-0.109	-0.166
Ca	0.156	0.357		0.109	-0.118
Mg		0.315	0.190	-0.300	
K	0.157		0.519	-0.133	
Sum Bases	0.154	0.370			-0.111
Acidity	0.237	-0.226	0.172		0.145
CEC	0.297			-0.114	
Base Sat.		0.389		0.150	
OC¶	0.276		0.113	0.325	
pH#		0.341	-0.193	0.225	-0.134
Total N	0.266		0.218	0.182	
EC		0.184		0.521	0.333
			B Horizon		
HT†	0.126	-0.201	-0.141	-0.340	
Bulk Density	0.199	0.171	-0.200	0.279	
Total Clay	-0.150		-0.481		
Total Silt	-0.287		0.184		
Total Sand	0.307				
Fine Silt	-0.263	-0.143	0.241		
Coarse Silt	-0.225	0.159		-0.222	0.369
VF Sand‡	0.109	0.250	-0.216	-0.361	0.146
Fine Sand	0.290			-0.231	
Med.Sand	0.284			0.204	
Coarse Sand	0.196			0.448	0.262
VC Sand§	0.141			0.410	0.397
Ca	-0.203	0.315	-0.119		0.137
Mg	0.393		0.240	-0.315	
K	-0.155	0.165			-0.607

† Horizon thickness
‡ Very fine sand
§ Very coarse sand
¶ Organic carbon
# pH in 0.01M $CaCl_2$

Table 1.2 (cont'd).

Sum Bases	-0.198	0.366	-0.107		
Acidity	-0.226	-0.268	-0.174	0.211	
CEC	-0.261		-0.274	0.152	
Base Sat.		0.424	0.253	-0.136	
OC¶	-0.239		0.305	0.116	
pH#		0.333	0.405	-0.122	
Total N	-0.249		0.280		
EC	0.188	-0.133	0.133	-0.148	-0.152

Table 1.3. Importance of first five PCs of A and B Horizons.

	PC1	PC2	PC3	PC4	PC5
		A Horizon			
Standard deviation	3.19	2.30	1.38	1.13	1.09
Proportion of Variance†	0.44	0.23	0.08	0.06	0.05
Cumulative Proportion‡	0.44	0.67	0.76	0.81	0.86
		B Horizon			
Standard deviation	3.17	1.91	1.74	1.33	1.04
Proportion of Variance†	0.44	0.16	0.13	0.08	0.05
Cumulative Proportion‡	0.44	0.60	0.73	0.81	0.85

† Proportion of variance explained by each PC.
‡ Cumulative proportion of variance explained.

The subsequent PCs were more difficult to characterize than the first two PCs. In PC3 of the A horizon, K had the highest loading (0.519), and the loading of coarse silt was the second highest of -0.469. These two loadings were both above 0.4, which, by our established criteria, were considered important. In addition, they were considerably higher than the variables following them, such as very fine sand (–0.263). Therefore K and coarse silt were considered the two dominant variables in the PC3 of the A horizon. In PC3 of the B horizon, clay content, and pH were the two variables with the highest loadings, -0.481 and 0.405, respectively. In PC4 of the A horizon, soil EC was the only important variable with a loading of 0.521. In PC4 of B horizon, two soil texture variables, coarse sand (0.448) and very coarse sand (0.410) were important. In the last PCs investigated, two soil physical properties, horizon thickness (-0.675) and bulk density (-0.482) were important in the A horizon and K (-0.607) was the only important variable in PC5 of the B horizon. In sum, most of the variables that were important in PC3, 4 and 5 were representative of soil texture and fertility, although some soil physical properties were also important in some cases. From PC3, however, soil variables with higher loadings in the A and B horizons began to differ. Although only a small portion of the total sample variance was explained in the lower PCs (the portion of total sample variance explained by PC3, PC4, PC5 was 8.3%, 5.5% and 5.1%, for A horizon and 13.2%, 7.7% and 4.7%, for B horizon, respectively), they provided insight on identifying important soil variables and changes of soil variable covariance structure at different depths of the soil profile. In characterizing soil variability, attention has been focused on the first couple of PCs that generally bear most of the total variance (Ovalles and Collins, 1988). In this study, however, the subsequent PCs provided more information in determining important soil properties.

The Relationship Between Soil Properties and Topographic Data

Eleven soil properties had loading > 0.4 from the first 5 PCs of both A and B horizons and were continuously used in the further analyses. These soil variables included coarse silt, extractable K, soil EC, horizon thickness, bulk density from the A horizon and NH_4-acetate base saturation, clay, pH, coarse sand content, very coarse sand content and extractable K from the B horizon.

Table 1.4. Summary statistics of elevation and slope at sample locations.

Field		Mean	Max.	Min.	S.D.	Range
1	Elevation (m)	296.3	303.1	292.4	3.1	10.7
1	Slope (Degree)	0.97	2.04	0.29	0.47	1.75
2	Elevation (m)	297.7	302.2	292.4	3.1	9.8
2	Slope (Degree)	1.11	4.25	0.16	1.18	4.09

Table 1.4 shows the summary statistics of elevation and slope at the sample locations in both sub-fields. The Pearson coefficient of correlation (r) was calculated between important soil variables identified by PCA and field elevation and slope. The results of significant r values ($P \leq 0.1$) are listed in Table 1.5.

Table 1.5. Significant correlation coefficients (r) between topographic data (elevation, slope) and soil properties ($P \leq 0.1$)

	K (A†)	K (B‡)	pH§ (B)	Base Sat (B)	HT¶ (A)	Bulk Dens (A)	EC (A)	Coarse Silt (A)	VC# Sand (B)	Coarse Sand (B)	Total Clay (B)
					A Horizon						
Elevation	**-0.48**							**-0.43**			
Slope		**-0.47**			**-0.50**				**0.70**	**0.60**	
					B Horizon						
Elevation	**-0.69**	**-0.71**					**-0.47**	**-0.59**			**-0.63**
Slope				**0.56**						**0.59**	

† A horizon.
‡ B horizon.
§ pH in 0.01M $CaCl_2$.
¶ Horizon Thickness.
# Very coarse sand.

Elevation had a consistently negative correlation with all K measurements except for the B horizon in Field 1. Kravchenko and Bullock (2000) also reported a negative correlation between elevation and K. This was likely due to greater K leaching due to higher sand content observed at the higher elevation in both Field 1 and 2. Also, elevation was negatively correlated with the coarse silt content of the A horizon in both fields and clay of the B horizon in Field 2. Slope was positively correlated with very coarse sand and coarse sand contents of the B horizon in Field 1. The correlation coefficients were 0.70 and 0.60, respectively. Similarly, slope was also positively correlated with the coarse sand content ($r=0.59$) of the B horizon in Field 2. In the study by Moore et al. (1993), slope was also found to be positively correlated with the sand content of the surface soil (top 0.1 m), and this study extended this depth up to 0.5 m from the surface. K of the B horizon and A horizon thickness of Field 1 were negatively correlated to slope, and base saturation of the B horizon in Field 2 was positively correlated to slope.

In general, elevation and slope are the two topographic data that are most strongly related to soil properties (Walker et al., 1967). In this study, we found that elevation and slope were more frequently correlated to soil texture variables, such as percent clay, coarse silt, different sized sand contents, and A horizon thickness. In addition to the difference of parent material, this relationship was also related to the erosion class and severity. Elevation and slope positions can alter the distribution of soil particles through erosion, which can be increased by cultivation practices (Ciha, 1984). Both Stone et al. (1985) and Kreznor et al. (1989) reported that clay content increased as soil erosion class became more severe. This was also supported by our results, as indicated by the negative correlation between slope and A horizon thickness, which has been used as an indicator of erosion severity (Deniels et al., 1985; White et al., 1985).

The Relationship of Soil Properties and Topographic Data to Crop Yield

Table 1.6 shows the summary statistics of corn and soybean yield from1996 to 2001. Dense data collected by the yield monitor allowed for the calculation of population statistics. The numbers in the parentheses in Table 1.6 are the sample statistics. Contents of coarse sand and very coarse sand of B horizon were summed to represent the total amount of coarse material, and to avoid high correlation between the two in the multiple regression analysis. Two topographic parameters, elevation and slope, were also included in the analyses.

Correlation coefficients between yield and soil properties, and topographic data are shown in Table 1.7. The clay content of the B horizon seemed to be the most frequent soil property that consistently had a negative correlation to crop yield. A negative correlation between clay content in both A and B horizons and crop yield was also observed in previous studies (Stone, 1985; Wright et al., 1990). High water holding capacity and decreased drainage rate associated with high clay content. This may cause prolonged high moisture content in the rooting zone, and consequently has a negative influence on root respiration. Base saturation was significant in 3 site-years

Table 1.6. Summary statistics of available soybean and corn grain yield during 1996-2001.

Field	Year	Crop	Mean†	S. D. †
			--------------------Bu/acre--------------------	
1	1996	Corn	105.6 (106.5)	16.7 (11.9)
1	1997	Soybean	39.0 (38.0)	6.4 (6.8)
1	1998	Corn	161.1 (160.2)	15.1 (8.7)
1	1999	Soybean	33.9 (34.6)	4.9 (3.5)
1	2000	Corn	152.5 (150.7)	15.1 (15.0)
1	2001	Corn	123.5 (124.1)	15.1 (9.9)
2	1997	Corn	147.3 (152.2)	27.2 (20.1)
2	1999	Corn	145.0 (155.9)	27.5 (30.2)
2	2001	Corn	120.9 (125.5)	20.3 (20.3)

† Calculated from all yield points and yield points at sample locations (in parentheses).

out of 9 studied, but both positive and negative correlations were observed. The relatively high base saturation indicated adequate exchangeable nutrient availability and this may not be a limiting factor of yield in general. The K of both A and B horizons had consistent positive influence on corn yield in two consecutive years in Field 2, but was not significant in Field 1. Although the generally high soil testing level of K suggests that K level is less likely a yield limiting factor, our results showed that crop yield can benefit from the excessive K levels up to the B horizon in certain years. The pH ($CaCl_2$) of the B horizon and the coarse silt content of the A horizon had only one incidence of significant correlation to the corn yield in 1998 and 2000, respectively. The A horizon properties of thickness, bulk density, and EC were not significantly correlated to crop yield throughout the study period.

Between the two topographic variables, slope showed consistently negative correlation to crop yield in 3 site-years out of 9 studied. A negative relationship between slope and yield has been reported in many other previous studies (Ciha, 1984; Fahnestock, 1996; Changere and Lal, 1997; McConkay et al., 1997; Yang et al., 1998; Kravchenko and Bullock, 2000). Steep slope positions tend to associate with more severe erosion class which is characterized by thinner surface horizon, higher clay content, lower infiltration rate, and greater runoff resulting in poorer productivity of soil (Wright et al., 1990). Besides, the high coarse material content observed at steep slope positions (Table 1.5) indicated that lower water retention and fertility at these areas. Elevation had no significant correlation with either corn or soybean yield through the period of study. Kravchenko and Bullock (2000) found that elevation was the most contributing

Table 1.7. Correlation coefficients (r) between soil properties of A and B horizons, topographic data and yield, and the coefficients of determination (R^2) of multiple regression between yield and soil properties and topographic data at significant level of 0.05.

Year	Crop	Elev.	Slope	K (A)	HT† (A)	EC (A)	Coarse Silt (A)	Bulk Dens. (A)	PH (B)	Base Sat (B)	K (B)	Coarse Material (B) §	Total Clay (B)	R^2
Field 1														
96	**Corn**												**-0.55**	**0.31**
97	**Soybean**									**0.49**				**0.63**
98	**Corn**		**-0.52**						**0.52**				**-0.58**	**0.50**
99	**Soybean**												**-0.58**	**0.31**
00	**Corn**						**-0.49**							
01	**Corn**												**-0.44¶**	
Field 2														
97	**Corn**													**N/A‡**
99	**Corn**		**-0.52¶**	**0.60**						**-0.56**	**0.49¶**			**0.64**
01	**Corn**		**-0.64**	**0.48¶**						**0.52¶**	**0.52¶**			**0.61**

† Horizon Thickness.
‡ Not available due to insufficient degree of freedom.
§ Sum of coarse and very coarse sand of the B Horizon.
¶ Significant at 0.1 level.

topographic factor in their study and a fairly consistent negative correlation between elevation and crop yield was frequently observed in most of the fields studied. In most cases, the influence of elevation on yield is reflected through water availability and its effect is more readily observed in extreme weather conditions with extreme topography (Kravchenko and Bullock, 2000). The lack of correlation between elevation and yield in this study can be partly explained by greater influence of some other soil properties on the yield variability, less extended extreme precipitation during growing seasons, and bio-factors such as weeds, disease and insect injuries.

Coefficients of determination (R^2) are also shown in Table 1.7. All 11 soil properties and 2 topographic variables were included in the initial analysis, and only those that were significant (0.05) after stepwise regression were retained in the final equations. Clay content of the B horizon was the only significant soil property that explained approximately 30% of the corn yield variability in 1996 (Field 1) and the same amount of the soybean yield variability in 1999 (Field 1). Also in the year of 1999, the largest portion, 64%, of corn yield variability was explained by three soil and topographic variables, clay content of the B horizon, elevation and slope in Field 2. In Field 1, there was no significant terms remained in the multiple regression analysis in year 2000 and 2001.

Conclusion

PCA returned 5 PCs for each horizon investigated, which cumulatively explained more than 85% of the total sample variance of each horizon. The first PCs of both horizons were represented primarily by soil texture variables and the second PCs by chemistry/fertility variables. These first two PCs explained up to 60% or more of the total sample variance in their corresponding horizons. Subsequent PCs explained individually a relatively small portion of the total variance, but started to show a difference in the variance structure of the A and B horizon.

Pearson's coefficients of correlation showed that topographic data (elevation and slope) were more frequently correlated with soil texture properties such as clay content, silt content, and coarse materials. This suggested that field topography had an indirect effect on crop yield by influencing certain soil properties.

The relationship of soil properties and topographic data to crop yield varied by year and field. Both soil and topographic information were useful in explaining yield, however, there was no explicit pattern or consistency. Only clay content (B horizon), K (A and B horizon), NH4-acetate base saturation (B horizon), pH (B horizon) had a significant and mostly consistent correlation with yield. B horizon soil properties seemed to be more informative in explaining yield variability by having more incidence of significance. The combined effect of soil properties and topographic data explained approximately 30-60% of the corn/soybean yield variability during the study period. A considerable amount of yield variability remained unexplained, and this suggested that yield limiting factors other than soil and topography such as climate and bio-factors can dominantly control yield variability.

References:

Ciha, A. J.1984. Slope Position and Grain Yield of Soft White Winter Wheat. Agron. J. 76:193-196.

Bennett, O.L., E.L. Mathias, and P.R. Henderlong. 1972. Effects of North and South Facing Slopes on Yield of Kentucky Bluegrass (Poa pratensis L.) with variable rate and time of nitrogen Application. Agron. J. 64:630-635.

Daniels, R. B., J.W. Gilliam, D. K. Cassel, and L. A. Nelson. 1985. Soil Erosion Class and Landscape Position in the North Carolina Piedmont. Soil Sci. Soc. Am. J., Vol. 49: 991-995.

Environmental Systems Research Institute. 1996. ArcView Spatial Analyst. ESRI, Redlands, CA.

Fahnestock, P., R. Lal, and G.F. Hall. 1996. Land Use and Erosion Effects on Two Ohio Soils: II. Crop yields. J. Sustain. Agric. 7:85-100.

Fiez, T. E., B. C. Miller, and W. L. Pan. 1994. Winter Wheat Yield and Grain Protein across Varied Landscape Positions. Agron. J. 86:1026-1032.

Franzmeier, D.P., E.J. Pedersen, T.J. Longwell, J.G. Byrne, and C. Losche. 1969. Properties of Some Soils in the Cumberland Plateau as Related to Slope Aspect and Position. Soil Sci. Soc. Am. Proc. 33:755-761.

Johnson, R. A., and D. W. Wichern. 2002. Applied Multivariate Statistical Analysis. Fifth Edition. P.426-427 Prentice Hall.

Kravchenko, A. N., and D. G. Bullock. 2000. Correlation of Corn and Soybean Grain Yield with Topography and Soil Properties. Agron. J. 92:75-83

Kreznor, W. R., K. R. Olson, W. L. Banwart, and D. L. Johnson. 1989. Soil, Landscape, and Erosion Relationships in a Northwest Illinois Watershed. Soil Sci. Soc. Am. J. 53:1763-1771 (1989).

Miller, M. P., M. J. Singer, and D. R. Nielsen. 1988. Spatial Variability of Wheat Yield and Soil Properties on Complex Hill. Soil Sci. Soc. Am. J. 52:1133-1141 (1988).

Moore, I. D., P. E. Gessler, G. A. Nielsen, and G. A. Peterson. 1993. Soil Attribute Prediction Using Terrain Analysis. Soil Sci. Soc. Am. J. 57:443-452 (1993).

Ovalles, F. A., and M. E. Collins. 1986. Soil-landscape Relationships and Soil Variability in North Central Florida. Soil Sci. Soc. Am. J. 50:401-408

Ovalles, F. A., and M. E. Collins. 1988. Variability of Northwest Florida Soils by Principal Component Analysis. Soil Sci. Soc. Am. J. 52:1430-1435 (1988).

Richardson, J. L., and R. J. Bigler. 1984. Principal Component Analysis of Prairie Pothole Soils in North Dakota. Soil Sci. Soc. Am. J. 48:1350-1355.

Walker, P. H., G. F. Hall, and R. Protz. 1968. Relation Between Landform Parameters and Soil Properties. Soil Sci. Soc. Amer. Proc., Vol. 32, 101-104.

Walker, P. H., G. F. Hall, R. Protz. 1968. Soil Trends and Variability Across Selected Landscapes in Iowa. Soil Sci. Soc. Amer. Proc., Vol. 32, 97-101.

Wright, R. J., D. G. Boyer, W. M. Winant, and H. D. Perry. 1990. The Influence of Soil Factors on Yield Differences Among Landscape Positions in an Appalachian Cornfield. Soil Sci. Vol. 149, No. 6.

Simmons, F. W., D. K. Cassel, and R. B. Daniels. 1989. Landscape and Soil Property Effects on Corn Grain Yield Response to Tillage. Soil Sci. Soc. Am. J. 53:534-539.

Sinai, G., D. Zaslavsky, and P. Golany. 1981. The Effect of Soil Surface Curvature on Moisture and Yield – Beer Sheba Observations. Soil Sci. 131:367-375.

Stone, J. R., J. W. Gilliam, D. K. Cassel, R. B. Daniels, L. A. Nelson, And H. J. Kleiss. 1985. Effect of Erosion and Landscape Position on the Productivity of Piedmont Soils. Soil Sci. Soc. Am. J. 49:987-991.

Yang, C., C. L. Peterson, G. J. Shropshire, and T. Otawa. Spatial Variability of Field Topography and Wheat Yield in the Palouse Region of the Pacific Northwest. Transactions of the ASAE. Vol. 41(1):17-27.

CHAPTER 2

THE EFFECTS OF SAMPLE SIZE AND CONFIGURATION ON THE SPATIAL STRUCTURE AND ACCURACY OF PREDICTION OF SOIL P AND K

Abstract

Precise soil fertility map is an important component of successful site-specific management. Hence, a sampling scheme that can ensure the desired accuracy in mapping soil properties at efficient cost is always of interest. The objectives of this study were (i) to evaluate effect of sample sizes and configurations on accurate description of spatial correlation structure for soil P and K concentrations, and (ii) to determine optimal sampling scheme for accurate and cost efficient mapping of soil P and K concentrations. A total number of 220 samples were taken, including 123 grid and 97 off-grid samples. A test data set of 31 samples was randomly selected for validation and the rest of the samples were used to design four sampling schemes. Scheme I had all data points (N_1=189), Scheme II had all grid samples and 42 randomly selected off-grid samples (N_2=148), Scheme III had grid samples only (N_3=106), and Scheme IV had 59 grid samples and 50 randomly selected off-grid samples (N_4=109). Geostatistical analysis was used to characterize spatial structure of the data and to predict values of the test data locations. Criteria used to compare the performance of the four sampling schemes included goodness of prediction (G), root mean square error (RMSE), and correlation coefficient (r). The sample semivariograms of soil P of different sampling schemes varied considerably, whereas those of K were fairly consistent. Scheme I of both soil P and K performed the best. Scheme IV performed as well as Scheme II with approximately 40

samples less, and Scheme III performed noticeably less accurate than Scheme IV, with approximately the same number of samples.

Introduction

Spatial variability of soil properties has been studied intensively. In agricultural studies, especially in site-specific management, researchers wish to have the knowledge on the pattern that one or more soil properties change over field and manage the field more efficiently. Many soil properties have been studied for spatial dependency, including soil pH, P, K, N, C (Cambardella et al., 1994; Cahn et al., 1994; Sutherland et al., 1991; Campbell, 1978; Webster and Cuanalo, 1975), soil water properties (Hamlett et al., 1986), pesticide residuals (Rao and Wagenet, 1985), soil infiltration rate (Vieira et al., 1983), soil respiration (Rochett, et al., 1991), soil bulk density (Cambardella, et al., 1994), and soil texture (Webster and Cuanalo, 1975; Campbell, 1978). Once the spatial correlation is determined, predictions can be made to unsampled locations and a soil property map can be created. Geostatistical techniques, originating from the mining industry, have been widely used in soil spatial studies. The main motivation for using geostatistical tools in soil science is the design of optimal sampling schemes and the spatial prediction of soil properties (McBratmey and Webster, 1986). Semivariogram and kriging are two central components of geostatistics. A semivariogram describes the structure of spatially dependent soil properties, and kriging is the best linear unbiased estimator based on the spatial structure described by semivariograms computed from samples (Trangmar et al., 1985). The advantage of geostatistics is that it takes into account spatial structures of the data and hence improves the precision of prediction (McBratney and Webster, 1983).

Different prediction methods have been compared in previous studies. Laslett et al. (1987) compared 8 prediction techniques to predict soil pH. These 8 techniques were classified as global or local, interpolating or non-interpolating, and smooth or non-smooth. They found that kriging and Laplacian smoothing splines, both as non-interpolating methods, generally performed best. In a following study, Laslett and McBratney (1990) reported universal kriging method performed consistently the best among four methods compared. Weber and Englund (1994) reported that the choice between kriging and inverse-distance for mapping soil properties remained unclear, since the prediction accuracy of different methods were dependent on data characteristics, such as variance, skewness, kurtosis, etc. Gotway et al. (1996) reported that compared to inverse-distance method, predictions made by kriging were unaffected by sample coefficient of variation, and were relatively high for all of the sampling configurations considered in the study. Kravchenko and Bullock (1999) also showed that ordinary kriging performed better than inverse distance methods with different power in the majority of the 30 data sets used in the study.

Grid sampling is the most widely used sampling scheme in soil spatial analysis due to maximum coverage of the field area with a minimum number of samples. Semivariograms computed from small samples usually appear erratic and have wide confidence intervals (Webster and Oliver, 1992). Consequently a large amount of uncertainty can be introduced into predictions. The sample size and the minimum spacing between sample points are most important factors in designing a sampling scheme. Short spaced samples can provide important information on the spatial behavior of a variable at short distance. The difference of spacing of grids can lead to different prediction results (Kravchenko and Bullock, 1999). Sample size and configuration have greater effect on the accuracy of prediction than the methods of prediction method used (Wollenhaupt et al., 1994; Mohamed et al., 1996). Wollenhaupt et al. (1994) compared predictions using samples taken from grid points and from grid cells, and found that at a 318ft grid set, samples from grid points gave a substantial improvement in prediction accuracy over samples from grid cells. The prediction map sampled from grid-points at 212 ft and 106 ft further increased map accuracy. Wibawa et al. (1993) concluded that a 50 ft grid sampling could effectively evaluate the variability of soil fertility in the field. The larger and denser the sample set is, the more accurate will be the semivariogram estimates and subsequent prediction. However, due to the cost and time issues associated with soil sampling and lab analysis, the optimum sample size and configuration have been always of interest. Ferguson et al. (1996) suggested that additional selective samples taken several years after initial grid sampling can provide useful information on the variation of nitrogen management. Mulla (1993) suggested an optimum sampling interval of 50-100 m would be adequate for the mapping of most soil properties. This was similar to the results obtained by Mohamed et al. (1996). According to Webster and Oliver (1992), for a normally distributed isotropic soil variable, 150 samples might often be satisfactory and 225 samples will usually be reliable over a study region. In summary, the sampling scheme should vary based on the purpose of the study, scale of the field, and the prediction accuracy desired.

The objectives of this study were:

1) to examine the effect of different sample sizes and configurations on the spatial behavior of soil P and K.

2) to evaluate the effect of different sampling schemes on the accuracy of prediction at unsampled locations.

Theory Background

Semivariogram

Semivariogram is the essential tool to evaluate and describe the spatial dependence of soil properties over a study region. A semivariogram represents the semi-variance, whose numerical value equals half the expected squared difference between sample values separated by a given distance (Trangmar et al., 1985). This process assumes *intrinsic stationarity* over the study region, which is defined through a constant mean and a constant variance in the difference between values at locations separated by a given distance and direction (Bailey and Gatrell, 1995). That is,

$$\mathbf{E}\ (\mathbf{Y}(\boldsymbol{s}+\boldsymbol{h}) - \mathbf{Y}(\boldsymbol{s})) = 0 \qquad (1)$$

$$\mathbf{VAR}\ (\mathbf{Y}(\boldsymbol{s}+\boldsymbol{h}) - \mathbf{Y}(\boldsymbol{s})) = 2\gamma(\boldsymbol{h}) \qquad (2)$$

Where $\mathbf{Y}(\boldsymbol{s})$ and $\mathbf{Y}(\boldsymbol{s}+\boldsymbol{h})$ are the observed values at location $\boldsymbol{s}$ and the location separated by $\boldsymbol{h}$ from $\boldsymbol{s}$; the $\gamma(\boldsymbol{h})$ is the semivariogram function. The semivariance γ depends on the separation vector $\boldsymbol{h}$. Three parameters, namely nugget, sill and range are used to describe the shape of sample semivariograms. If the soil process is spatially correlated, the semivariogram value (γ) is minimal at $\boldsymbol{h} = 0$, and increases as $\boldsymbol{h}$ increases. The level where the semivariogram starts to be more or less constant, is referred to as *sill.* Separation distance, at which semivariogram reaches its sill is called *range*. At the distance beyond the range, the soil process is no longer spatially correlated (Trangmar et al., 1985). The semivariogram value at $\boldsymbol{h} = 0$ is called *nugget* variance or *nugget* effect, which represents random variation and cannot be evaluated at the sampling scale. Nugget effect is caused by soil property variation at distances shorter than the minimum distance between the samples and by measurement errors. Most soil properties show distinct nugget effects (Burgess and Webster, 1980). The nugget plays an important role in kriging, because it sets a limit to the precision of the interpolation (Burrough, 1983; Trangmar et al., 1985). The estimates of the semivariogram can be calculated using the following equation:

$$\gamma(\boldsymbol{h}) = \frac{1}{2n(\boldsymbol{h})} \sum_{i=1}^{n} \left(Y(s_i+\boldsymbol{h}) - Y(s_i) \right)^2 \tag{3}$$

Where $\mathbf{n}(\boldsymbol{h})$ is the number of samples separated by a distance of $\boldsymbol{h}$, $\mathbf{Y}(s_i)$ are the values of the measured property at location s_i, and $\mathbf{Y}(s_i+\boldsymbol{h})$ are the values of the measured property at location $s_i+\boldsymbol{h}$.

Semivariogram model is a continuous function fitted to a sample semivariogram, and is used to provide parameter estimates computed from samples at any separation distance, which is required in kriging (Isaaks and Srivastava, 1989). Most commonly used types of semivariogram models include, spherical model, expressed by,

$$\gamma(\boldsymbol{h}) \begin{cases} = C\,(3\boldsymbol{h}/2r - \boldsymbol{h}^3/2r^3) & \text{for } \boldsymbol{h} \le r \\ = C & \textbf{otherwise,} \end{cases} \tag{4}$$

and exponential model, expressed by:

$$\gamma(\boldsymbol{h}) = C\,(1 - e^{-3\boldsymbol{h}/r}) \tag{5}$$

where, **r** is the range and **C** is the sill (McBratney and Webster, 1986).

McBratney and Webster (1986) reviewed different methods of fitting a model to sample variograms, and noted that the least squares estimation procedure requires strict distribution assumptions on residuals; the generalized least squares with relaxed residual distribution assumptions and the maximum likelihood estimation are generally regarded as efficient means of estimation, but both methods involve a formidable amount of computation. A weighted least squares estimator was found to be reliable and computationally efficient, if the weights can be determined properly.

Ordinary Kriging

Kriging combines knowledge of both observed values and the spatial correlation between data values at sampled locations characterized by semivariogram. Kriging estimates at unsampled locations are optimal and unbiased (Trangmar et al., 1985). Under kriging paradigm, ordinary kriging is focused on optimal local spatial prediction. It forms the prediction of the process at unsampled location by using a weighted linear combination of the observed values $\mathbf{y}_i$ at the sample sites $\mathbf{s}_i$. Estimated values $\hat{\mathbf{Y}}(s)$ are obtained by calculating the following formula:

$$\hat{\mathbf{Y}}(s) = \sum_{i=1}^{n} \omega_i \mathbf{Y}(s_i) \tag{6}$$

where $\hat{\mathbf{Y}}(s)$ are the locations to be estimated, $\mathbf{Y}(s_i)$ are the neighboring samples and weights (ω) are a function of spatial structure described by semivariogram. The weights are chosen so that the estimated value is unbiased, i.e.

$$\mathbf{E}[\hat{\mathbf{Y}}(s)-\mathbf{Y}(s)] = 0 \tag{7}$$

This process assumes that the mean of $\hat{\mathbf{Y}}(s)$ and $\mathbf{Y}(s)$ are all μ and constrains that the sum of the $\mathbf{n}$ weights (ω) is unity. The expected mean square error between true values and observed value in kriging is minimized, i.e.

$$\hat{\sigma}^2 = \mathbf{VAR}\ [\mathbf{Y(s)\text{-}Y(s)}] = \mathbf{minimum} \tag{8}$$

The estimated value at any unsampled location is the most precise possible from the available data and one that can be used with known confidence (McBratney et al., 1982). The estimation variance depends only on the semivariogram and the configuration of the sample locations in relation to the predicted points and not on the observed values themselves (Burgess and Webster, 1980). The weights are determined by taking account of the known spatial dependence expressed in the semivariogram and geometric relationships among observed point locations. In general, near points carry more weight than distant points, points that occur in clusters carry less weight than lone points, and points lying between the point to be predicted and more distant points screen the distant points so that they have less weight (Burgess and Webster, 1980).

Materials and Methods

Study Site Description

The study site was a corn (*Zea mays* L.) – soybean [*Glycine max* (L.) Merr.] field located in Kalamazoo County, Michigan (42.22 N, 86.36 W). The field is approximately 50 hectares and was planted to corn and soybean in an alternate year rotation. The elevation of the field ranges from 288 to 303 m with near level (0-2%), gentle (2-6%), and moderate (6-12%) slopes. The soil of the study site is predominately Kalamazoo (fine-loamy, mixed, mesic Typic hapludalfs) with Oshtemo (coarse-loamy, mixed, mesic typic Hapludalfs) at the south-west corner of the field (Soil Survey of Kalamazoo County, Michigan, 1979). The field was non-irrigated and employed a minimum tillage system. The whole field was divided into two sub-fields – Field 1 (east side) and Field 2 (west side) by the rotation boundary of soybean and corn.

Field Data Collection and Laboratory Analyses

Soil sampling was conducted along 9 transects in April 2000. The total number of soil samples was 220. Two sets of grid samples, 123 one-acre samples and 25 five-acre samples, as well as 72 off-grid samples were taken. The smallest spacing between one-acre-grid samples was approximately 64 m. The off-grid samples were taken 7.5 m, 15 m, and 30 m away from three randomly selected one-acre-grid points in each of 9 transects in one of the four possible directions. The sample (topsoil) for each location was composited from one subsample taken from the center location and other four subsamples taken from 1.5 m away from the center in four directions. Soil P (Bray and Kurts P1, Knudsen and Beagle, 1988), and K (1N NH_4OAc at pH 7.0, Brown and Warncke, 1988) were determined by the Michigan State University soil testing lab.

Sampling Scheme Design

A subset of 31 samples, randomly selected from the total 220 samples, was used as a test dataset for validation purposes. The rest of the dataset (n=189) was used to create four sampling schemes with different sample sizes and configurations that were used for spatial analysis. These four sampling schemes were designed in the following ways and are graphed in Figure 2.1.

Scheme I: All data points were used, which included 106 one-acre grid points, and 83 off-grid points. Five-acre grid points were treated as off-grid samples. The total number of samples used in Scheme I (N_1) is equal to 189.

Scheme II: All one-acre grid points (106) were used along with 42 randomly selected off-grid samples. The number of the off-grid samples was equal to half of that used in Scheme I. N2=148. The total number of samples used in Scheme II (N_2) is equal to 148.

Scheme III: Used one-acre grid points only. The total number of samples used in Scheme III (N3) is equal to 106.

Scheme IV: 59 one-acre grid points, obtained by removing every other one-acre grid points, and 50 randomly selected off-grid samples. The total number of samples used in Scheme IV (N4) is equal to 109.

Figure 2.1a-d. Four sampling schemes†. a, Scheme I. All data points used, N_1=189; b, Scheme II. All one-acre-grid samples, with 42 randomly selected off-grid samples. N_2=149; c, Scheme III. One-acre-grid samples only. N_3=106; d, Scheme IV. 59 one-acre-grid samples, with 50 randomly selected off-grid samples. N_4=109.

a.

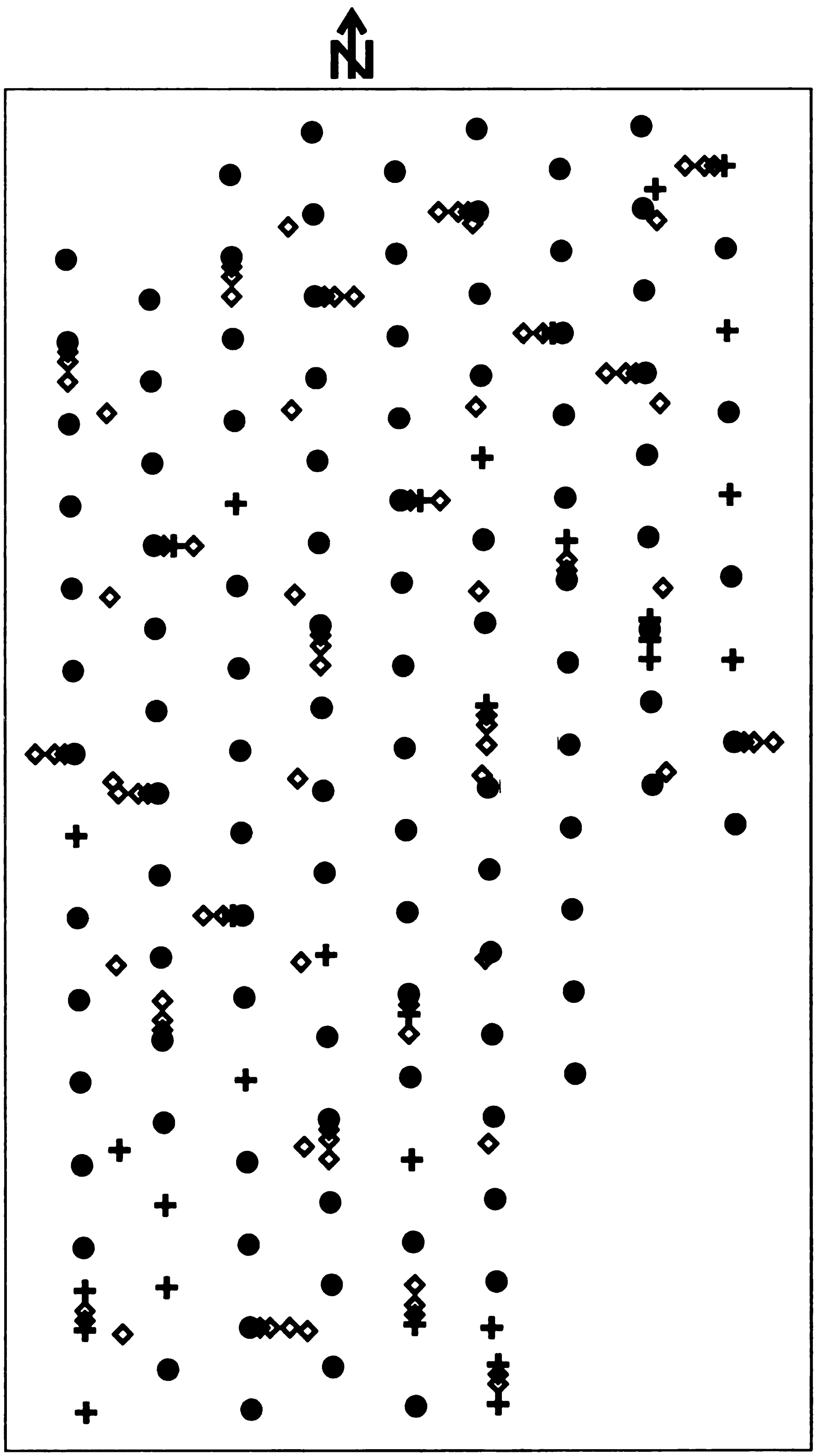


Figure 2.1. (cont'd).

b.

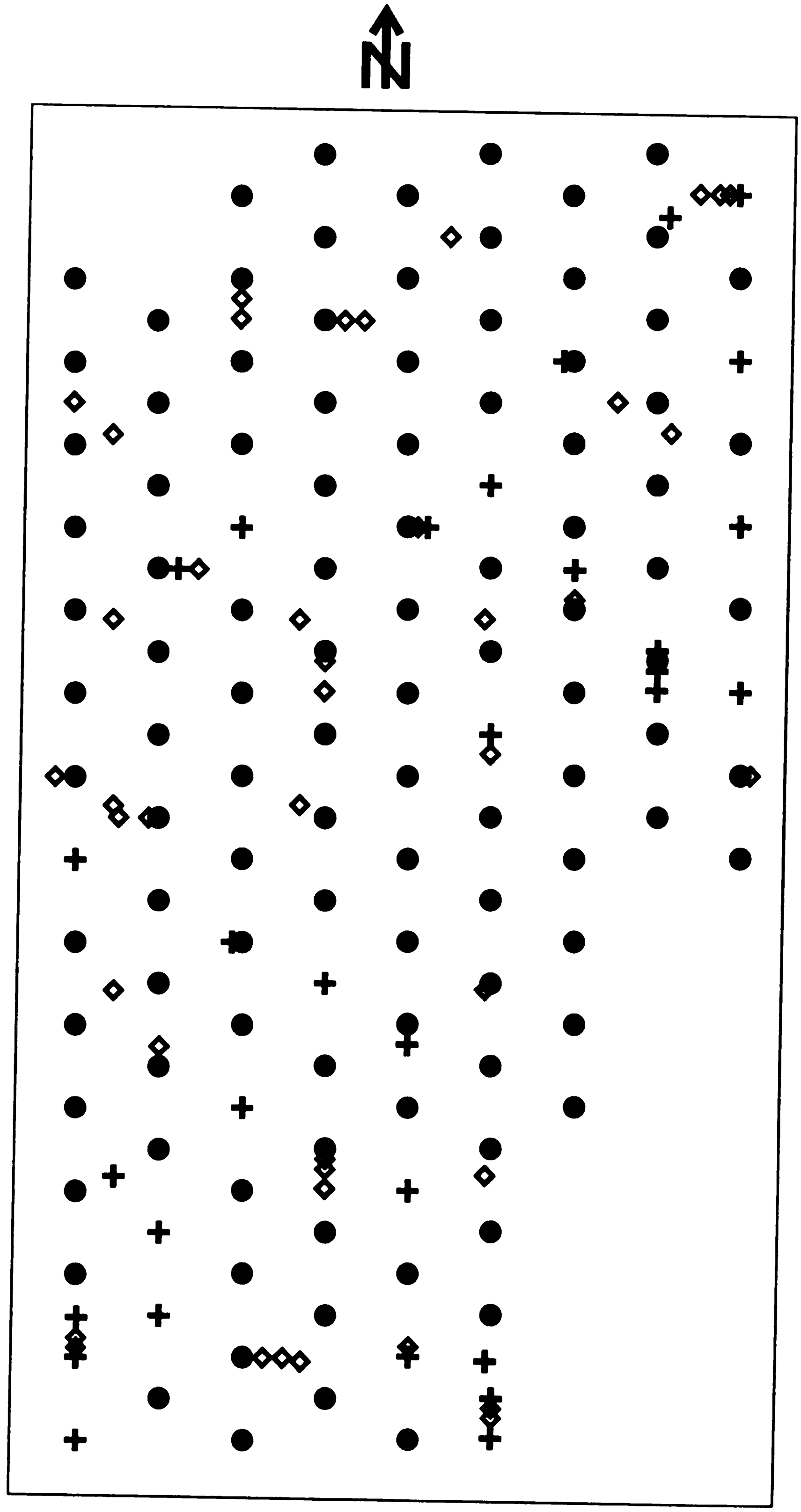


Figure 2.1. (cont'd).

c.

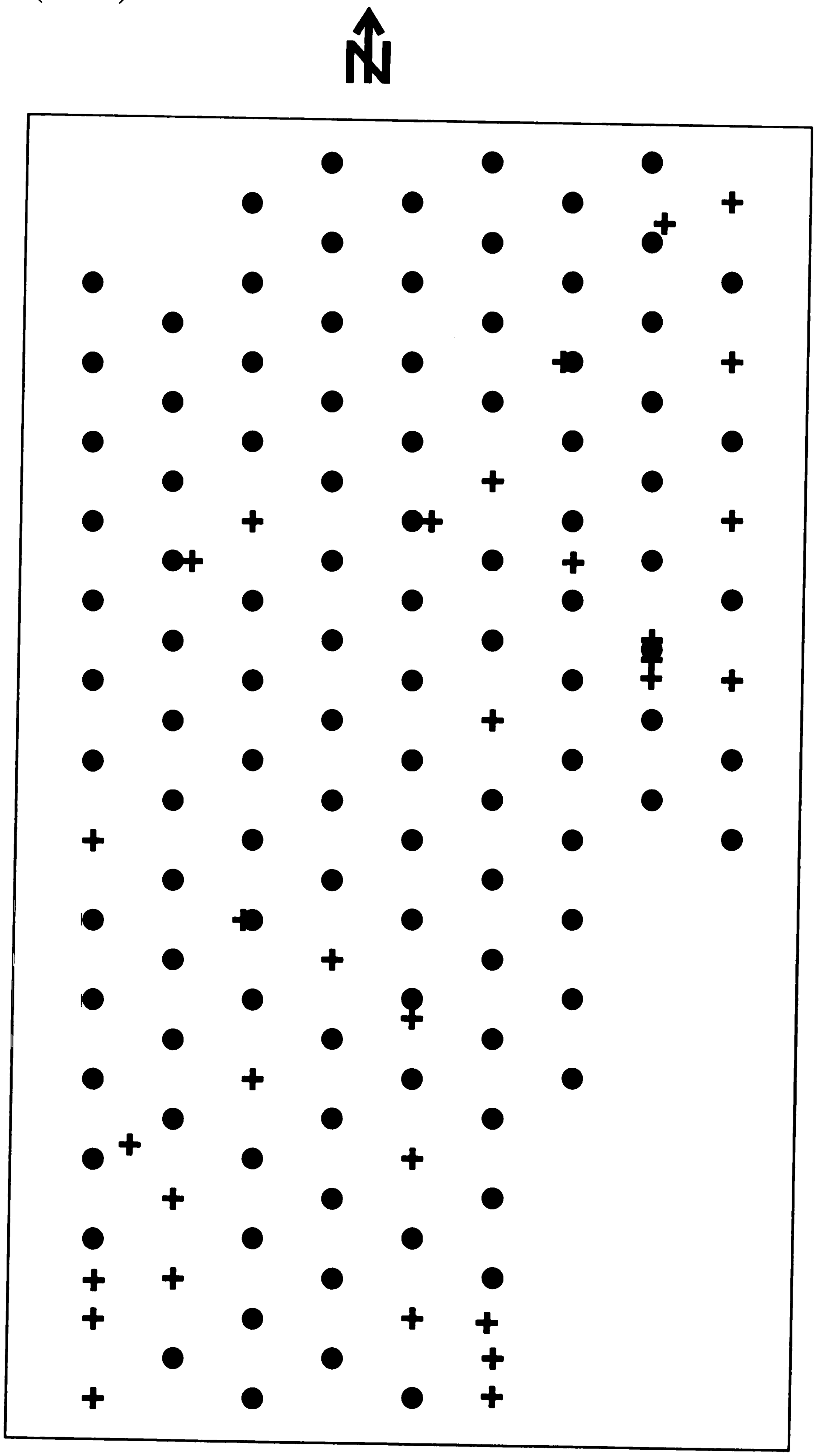
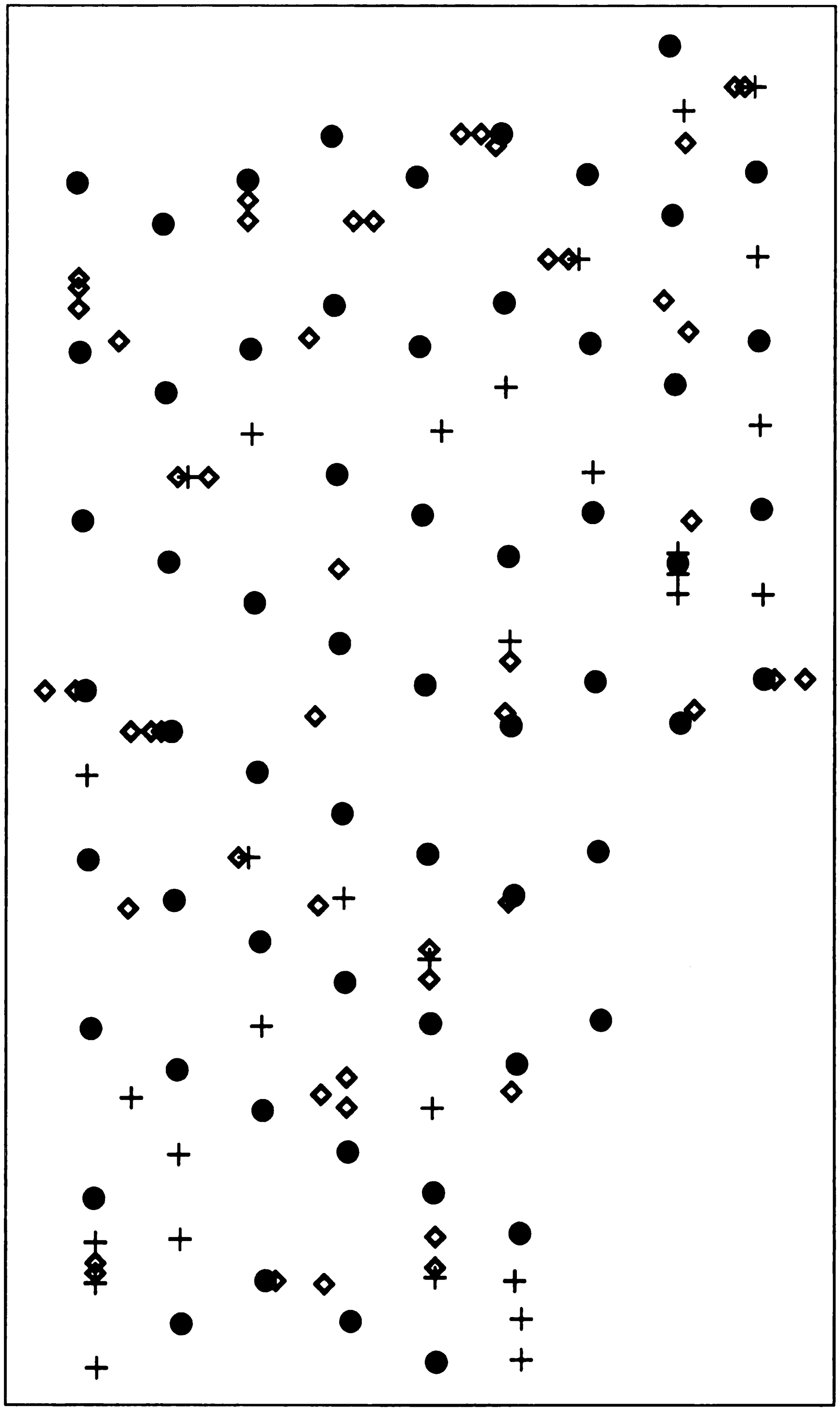


Figure 2.1. (cont'd).

d.



† The dots represent one-acre-grid samples, the diamonds represent off-grid samples, and the crosses are the locations where predictions were to be made.

Data Analysis

Sample omnidirectional semivariograms were calculated for both soil P and K for four sampling schemes using Eq.3. A semivariogram model was fitted to each calculated sample semivariogram. Then, ordinary kriging (Eq.6) was performed to predict values for the 31 locations of the test dataset. The predicted values and the actual values at these locations were compared using goodness-of-prediction criteria, G (Gotway et al., 1996; Kravchenko and Bullock, 1999), root mean square error (RMSE), and coefficient of correlation (r). The goodness-of-prediction, G, was calculated using the following equation:

$$\mathbf{G = \{1 - \sum_{i=1}^{n} [Z^*(x_i) - Z(x_i)]^2 / \sum_{i=1}^{n} [Z_m - Z(x_i)]^2\} \times 100} \qquad [9]$$

where $\mathbf{Z^*(x_i)}$ are the predicted values at location **i**, $\mathbf{Z(x_i)}$ are the observed values at location **i**, $\mathbf{Z_m}$ is the mean of the observed values, and **n** is the number of samples used for calculation. The root mean square error (RSME) was calculated by the following equation:

$$\mathbf{RMSE = \sqrt{\left(\sum_{i=1}^{n} [Z^*(x_i) - Z(x_i)]^2 / n\right)}} \qquad [10]$$

The statistical software package SPLUS (Mathsoft, 1999) was used to perform the statistical analyses. Ordinary kriging was performed by Gstat (Pebesma, 1999), a geostatistical package.

Results and Discussion

Table 2.1 shows the summary statistics of soil P, and K concentrations calculated based on the data from the four sampling schemes. The minima and maxima were equal to 15.5 and 270.5 mg/kg for soil P and 75.0 and 435.0 mg/kg for K, respectively. Several outliers in both P and K were observed and caused the histograms of the two soil properties to be severely skewed (P>140 mg/Kg, and K>350 mg/Kg were considered as outliers). Since sample semivariograms are ideally calculated based on a normal distribution (Trangmar et al., 1985; Bailey and Gatrell, 1995), and the estimator of the semivariogram can be sensitive to departures from normality of the data (McBratney and Webster, 1986), the outliers were removed from the training datasets before sample semivariograms were calculated. Both skewness and kurtosis of the two soil properties improved after the outliers were removed. (Table 2.1. The numbers in parentheses show the improved skewness and kurtosis). Figure 2.2 presents sample histograms of P and K at Scheme I with outliers removed.

Figure 2.2. Example sample histograms of P and K after outliers were removed (Scheme I).

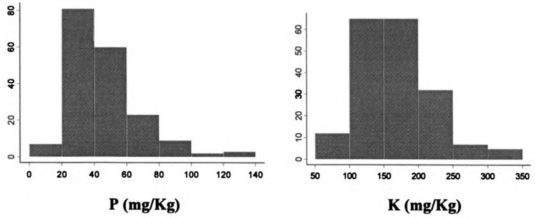


Table 2.1. Summary statistics of soil P, and K of four sampling schemes.

Scheme (N)	Mean	S.D.	Skewness	Kurtosis	Mean	S.D.	Skewness	Kurtosis
	---Mg/Kg---				---Mg/Kg---			
			P				K	
I (189)	50.0	31.3	3.2 (1.5)	16.0(3.0)	172.3	59.1	1.3 (0.8)	2.7 (0.6)
II (148)	49.7	31.8	3.7 (1.4)	19.3(3.1)	173.9	59.6	1.4 (0.7)	3.3 (0.7)
III(106)	51.4	34.5	3.9 (1.5)	19.4(2.8)	178.4	58.9	1.5 (0.6)	3.5(0.1)
IV(109)	50.1	34.4	3.4 (1.5)	16.9(2.8)	173.7	63.3	0.7 (0.8)	0.4(0.8)

Figures 2.3 and 2.4 show the semivariograms of soil P and K calculated at each sampling scheme. The lags used to calculate the semivariograms were adjusted to produce a better recognizable semivariogram and help visual fitting of a semivariogram model. This resulted in different numbers of data pairs in semi-variance estimators of the sample semivariograms. The lag values used in semivariogram calculations ranged from 20 to 45 m. An empirical semivariogram model was fitted "by eye" to each calculated sample semivariogram. Soil P was best fitted with an exponential model and soil K was best fitted with a spherical model for all four sampling schemes. The parameters of each semivariogram model fitted were summarized in Table 2.2.

Table 2.2. The parameters of semivariogram models fitted to each sampling scheme of soil P and K.

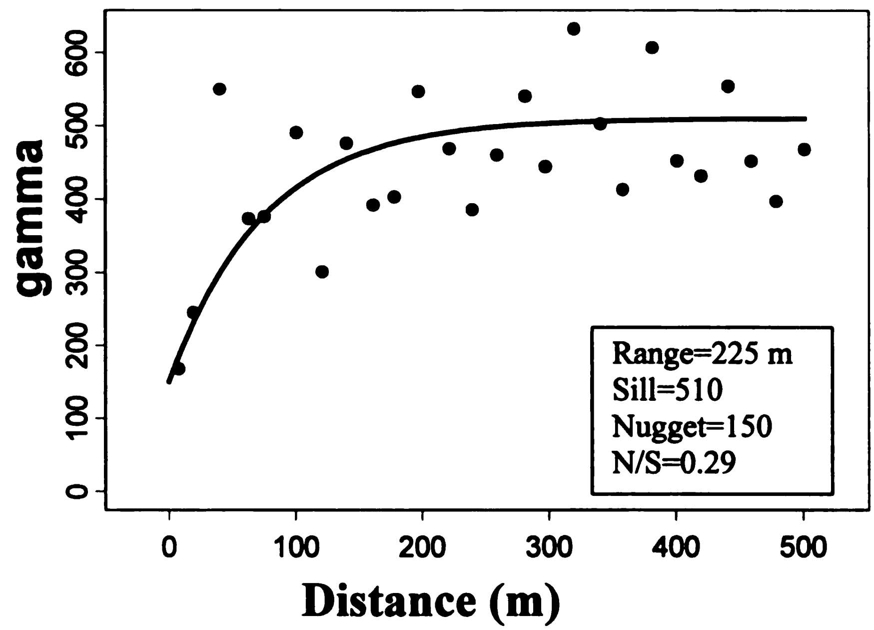
Scheme (N)	Range† ---m---	Sill	Nugget	Nugget/Sill
		P (Exponential)		
I (189)	225	510	150	0.29
II (148)	750	430	230	0.53
III (106)	1110	310	200	0.65
IV (109)	240	560	240	0.43
		K (Spherical)		
I (189)	280	2900	900	0.31
II (148)	300	2600	1000	0.38
III (106)	320	2900	1000	0.34
IV (109)	300	2400	1100	0.32

† The ranges of P are actual ranges, which were obtained by multiplying by 3 the ranges from the fitted exponential models.

Among the four models fitted to soil K, Scheme I (N_1=189) had the shortest range of 280m, and Scheme III (N_2=106) had the longest range of 320 m. Scheme II and IV had the same range of 300m. The shortest and longest ranges occurred at the same schemes in case of soil P with value of 75m (Scheme I) and 370m (Scheme III), respectively. Nugget effect of soil K increased slightly from 900 (Scheme I) to 1100 (Scheme IV), whereas the nugget effects of P varied much more noticeably from 150 (Scheme I) to 240 (Scheme IV). Nugget effect arises due to the combination of variation over distances smaller than the closest sample spacing and random measurement errors (McBratney and Webster, 1986). Both P and K had appreciable nugget effect in the study site at all sampling schemes. As expected, the shorter the sample spacing, the smaller the nugget effect (Trangmar et al, 1985). In both P and K, Scheme I (N_1=189) had the smallest nugget effect, and Scheme IV (N_4=109), with no short spaced samples, had the largest nugget (Table 2.2). Commonly, the ratio of nugget to sill is used to express the relative amount of the nugget effect. (Trangamar et al., 1985; Bailey and Gatrell, 1995). Cambardella et al, (1994) used this measure to examine the strength of spatial correlation of soil properties. They determined that if the ratio is ≤ 25%, the variable has strong spatial dependence, if the ratio is between 25-75%, the variable has moderate spatial dependence, and if the ratio is >75%, the variable has weak spatial dependence. By these criteria, both soil P and K had moderate spatial dependency (Table 2.2). However, the spatial dependency of K was reflected quite similarly by all sampling schemes, whereas that of P was noticeably different at different sampling schemes. The shapes of all semivariograms of K were fairly consistent (Fig. 2.4), and the nugget/sill ratios were low. Schemes II and III of P showed very weak spatial structure by having a large nugget to sill ratio and an extremely long range compared to Schemes I and IV (Fig 2.3). This implied that the short distance variation of soil K was rather small in the study field, and the sampling schemes didn't influence the three parameters considerably. For soil P, however, there may be appreciable noise in the samples besides the short distance variation. Webster and Oliver (1992) noted that semivariograms depend more on the shortest spacing of samples, and less on the number of samples. In case of P, an insufficient number of short spaced samples in Scheme III caused the loss of a considerable amount of spatial structure (Fig 2.3c).

Figure 2.3. Empirical semivariograms of soil P with the semivariogram models fitted. a. Scheme I. b. Scheme II. c. Scheme III. d. Scheme IV.

a



b

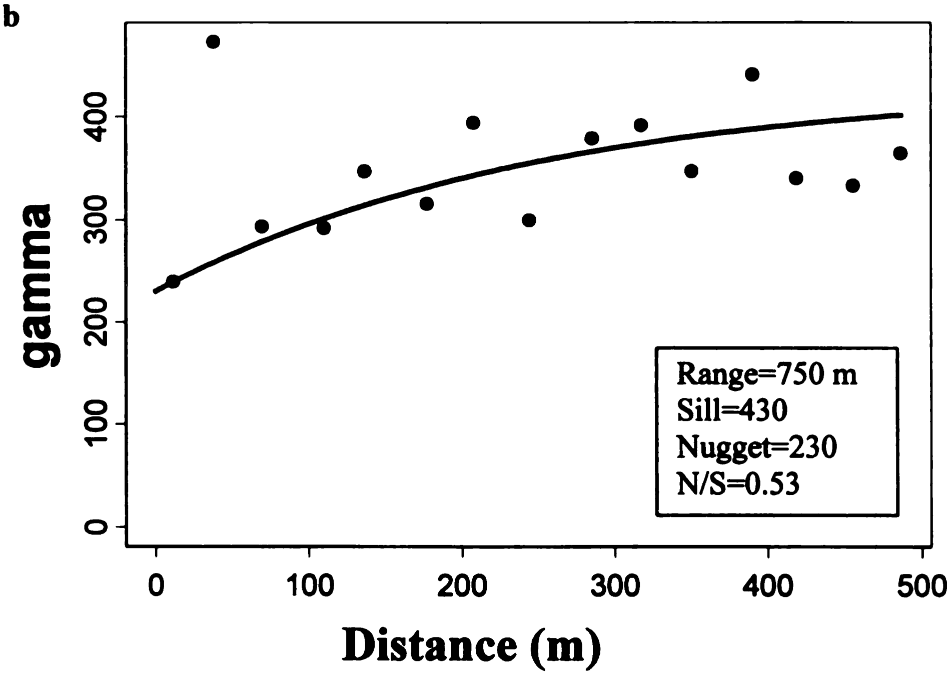
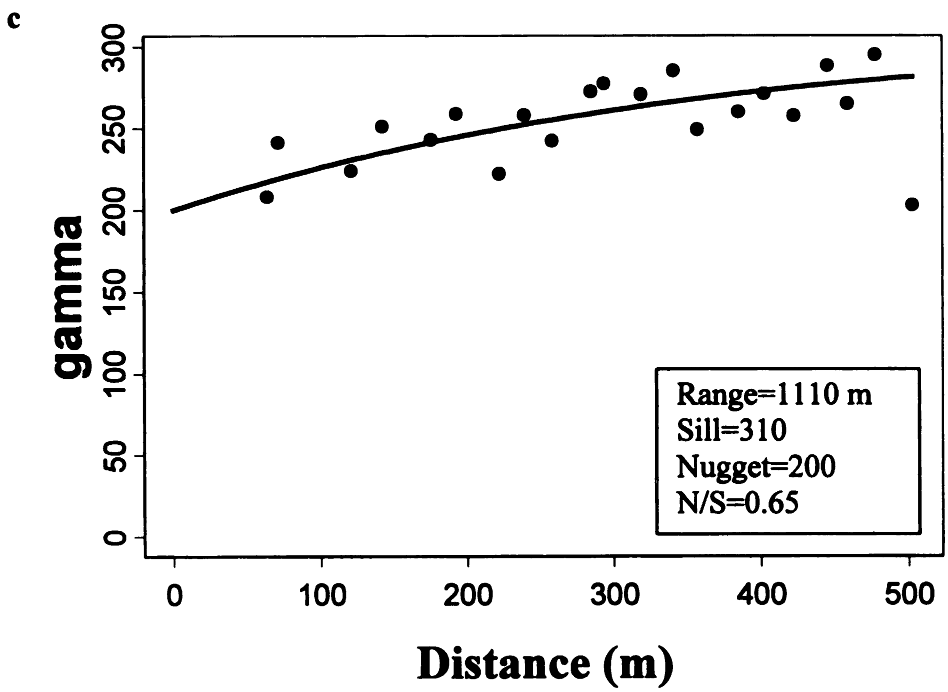


Figure 2.3. (cont'd).



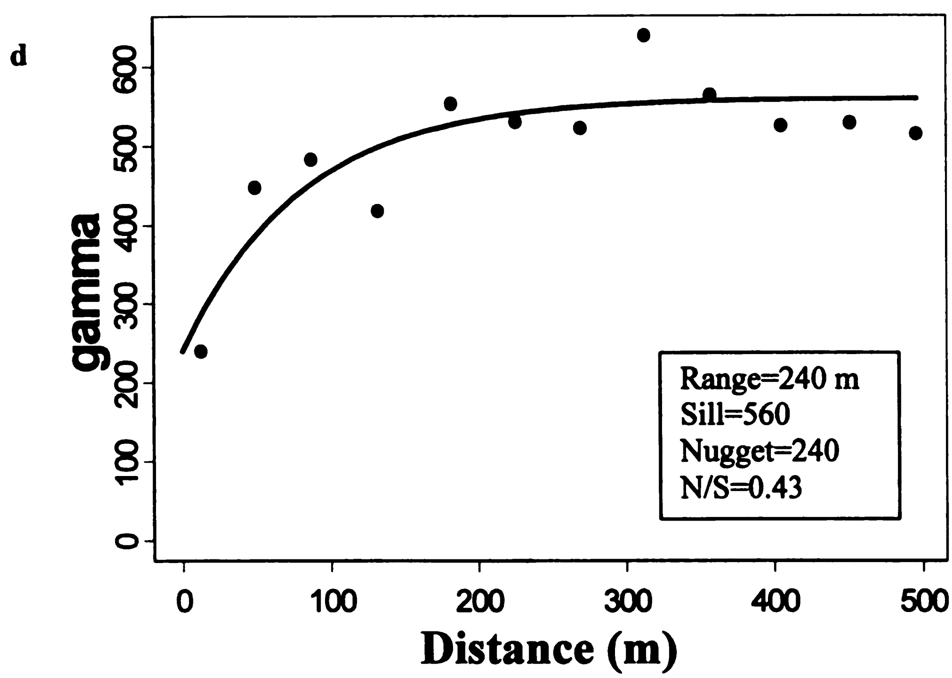
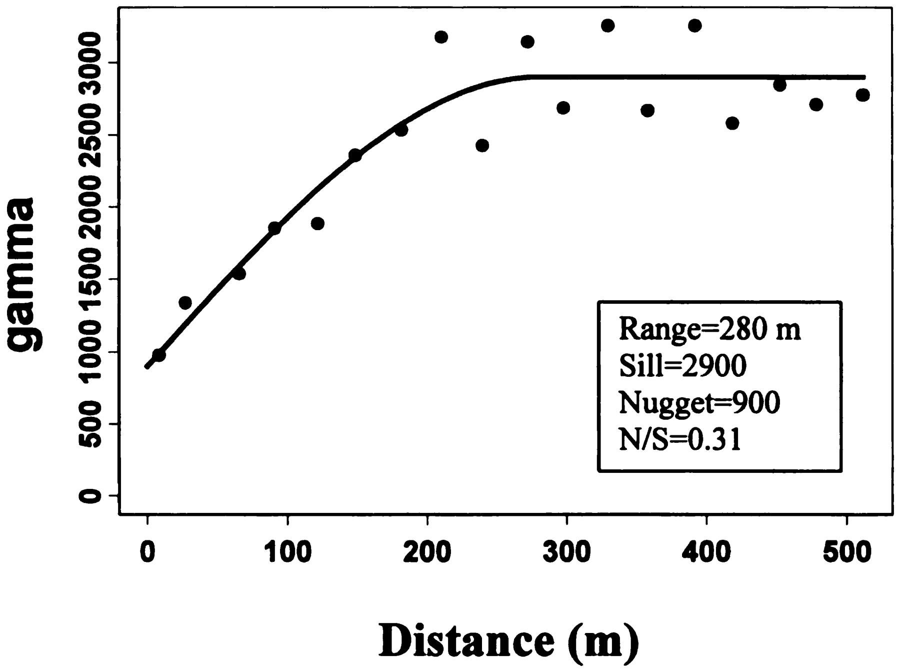


Figure 2.4. Empirical semivariograms of soil K with the semivariogram models fitted. a. Scheme I. b. Scheme II. c. Scheme III. d. Scheme IV.

a



b

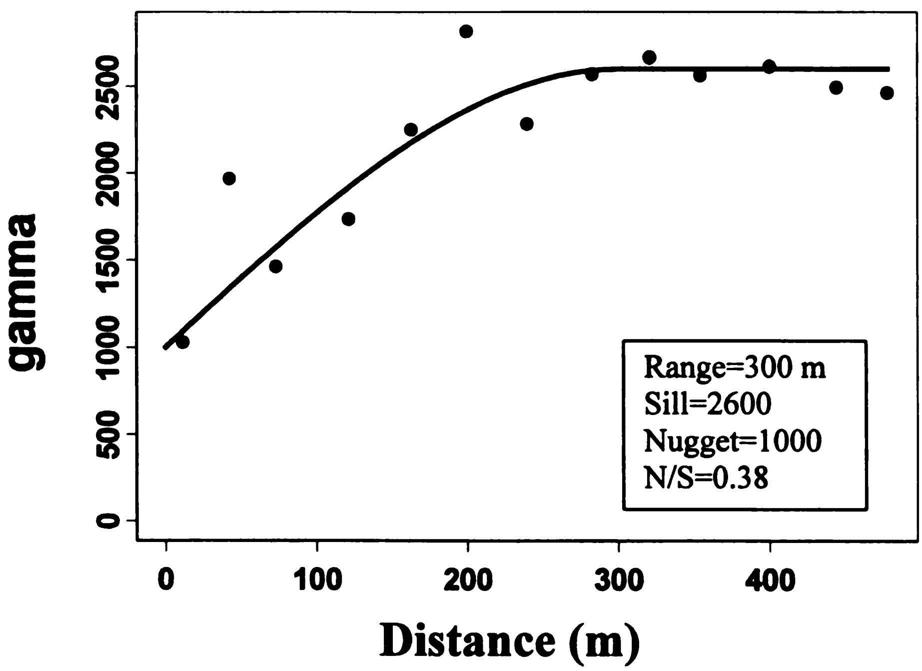
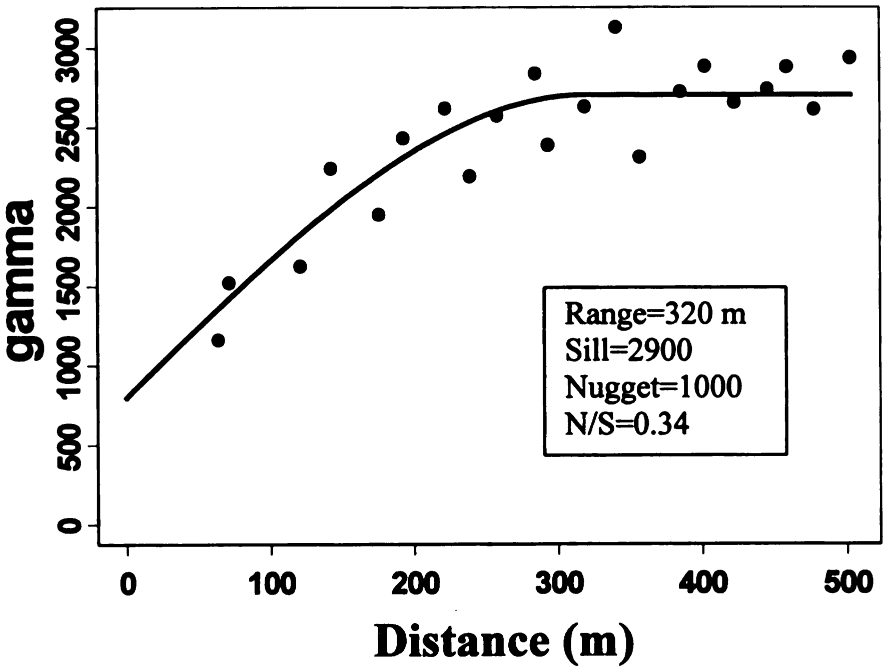
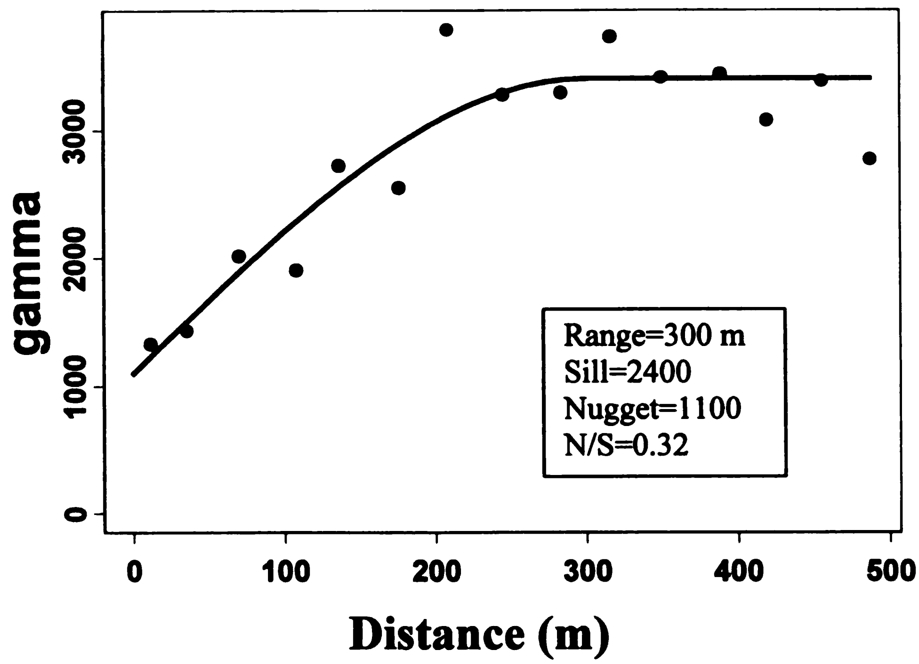


Figure 2.4. (cont'd).

c



d



Although the semivariograms appear similar at different sample sizes and configurations, they can differ greatly in confidence intervals of estimates of semi-variance for each lag. Webster and Oliver (1991) noted that the larger the sample size and the shorter the sample spacing, the smaller the confidence interval. The confidence intervals will be very important when prediction is concerned since the prediction variance can be very sensitive to the confidence intervals of sample semivariograms.

The goodness-of-prediction, G (Eq. 9), the root mean square error (RMSE, Eq.10), and the correlation coefficient (r) were summarized in Table 2.3.

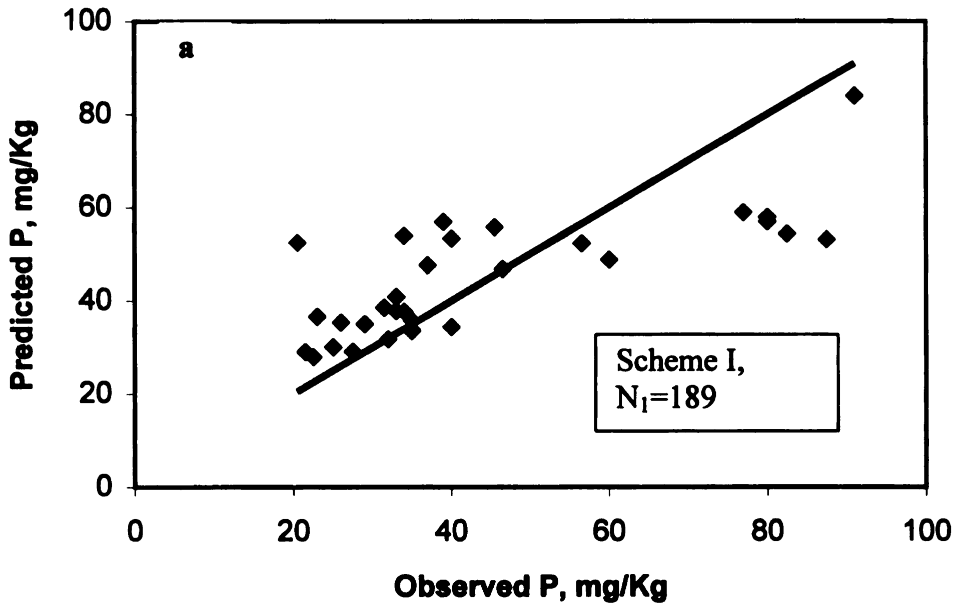
Table 2.3. The goodness of prediction (G), root mean square error (RSME), and correlation coefficient (r) between predicted and actual values.

Scheme (N)	**G**	**RSME**	**r**	**G**	**RMSE**	**r**
		P			K	
I (189)	55.1	14.2	0.76	50.0	37.4	0.72
II (148)	45.9	15.6	0.76	42.9	40.0	0.68
III (106)	38.9	16.6	0.69	34.1	43.0	0.63
IV (109)	46.0	15.6	0.71	45.3	39.1	0.69

The G value is a measure that provides information on the improvement made by the used prediction method as compared to using the sample mean as a predictor. The greater the G value, the better the prediction method performed (Gotway, 1996). The highest G value of P and K both occurred at Schemes I (55.1 and 50.0, respectively). Also, the RMSE of both Schemes I was the lowest (14.2 and 37.4 for P and K, respectively). Scheme II and IV of both P and K performed almost equally well in general. Their G values and RMSE were similar (Table 2.3), although Scheme II (N_2=148) had approximately twice as many as acre grids than Scheme IV (N_4=109) did. Schemes III for both P and K had the lowest G value (38.9 and 34.1, respectively) and the highest RMSE (16.6 and 43.0, respectively). The correlation coefficients (r) between predicted and actual values again were the highest in Schemes I (0.76 and 0.72 for P and K, respectively), the lowest in Schemes III (0.69 and 0.63 for P and K). These results indicate that a lack of short spaced samples considerably influences the accuracy of prediction. Webster and Oliver (1992) noted that since ordinary kriging put more weight on nearby samples and less weight on remote ones, the close-by samples are of more importance than distant ones, and short spaced samples can provide more information than regular grid samples at longer distances. With a similar number of short spaced samples, the extra one-acre-grid samples in Scheme II (N_2=148) did not improve the prediction accuracy compared to Scheme IV (N_4=109). On the other hand, at similar sample sizes, Scheme IV (N_4=109) performed considerably better than Scheme III (N_3=106) due to the short spaced samples.

Figure 2.5 and 2.6 plotted the predicted values versus actual values of soil P and K, with a 1:1 line presented. Compared with the 1:1 line in the plots, it can be observed that the kriging predictions at all sampling schemes tended to overestimate the low data values and underestimate the high data values. This trend was particularly evident in P predictions (Fig. 2.5) and less so in K predictions (Fig. 2.6). This implied that even the largest sample size was not able to accurately predict local high and low values. If a higher level of accuracy of prediction is desired, additional samples would be needed, even in Scheme I, to capture local high variation. However, if an accurate soil property map is of interest, a sampling scheme like Scheme IV in this study can serve as a start with relative small number of samples and reasonable prediction accuracy. The residual map could be helpful to identify the areas where the prediction error is high. Extra samples can be taken selectively at those areas to improve the predictions for high and low observations.

Figure 2.5a-d. Predicted values vs. observed values of soil P of four sampling schemes. a. Scheme I. b. Scheme II. c. Scheme III. d. Scheme IV.



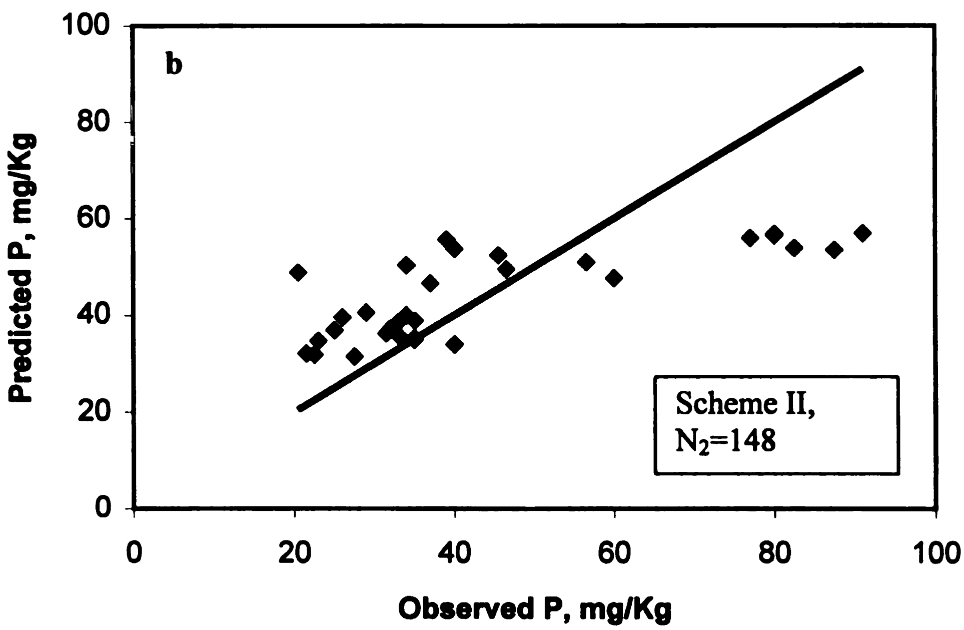
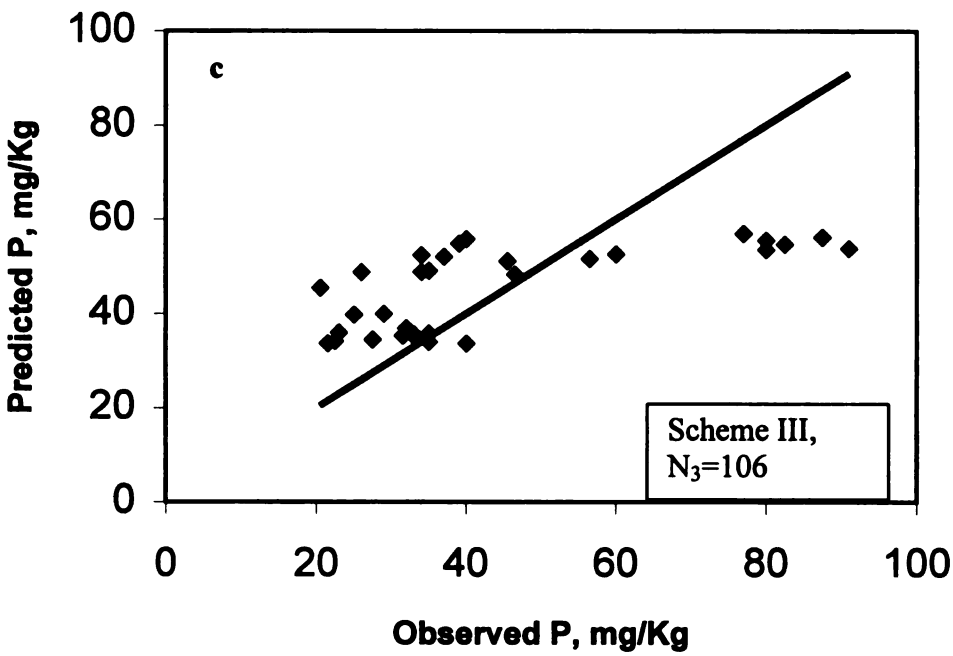


Figure 2.5. (cont'd)



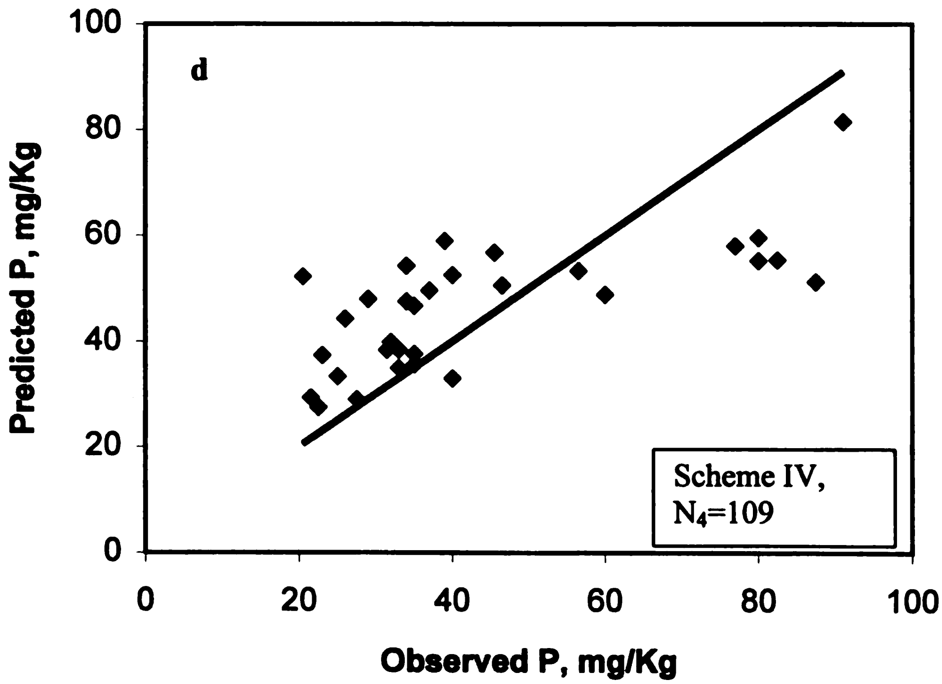
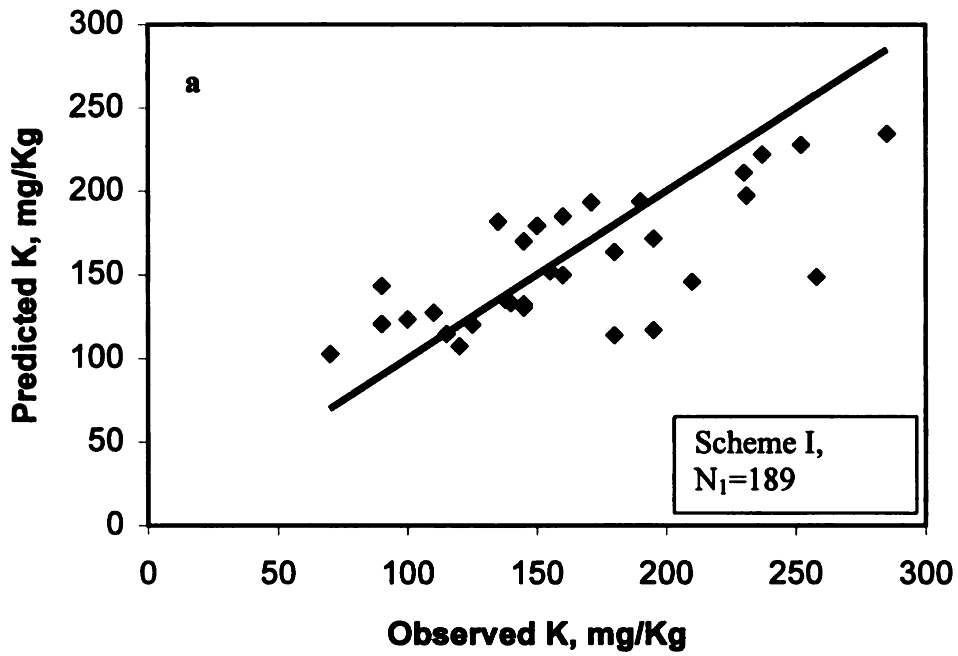


Figure 2.5a-d. Predicted values vs. observed values of soil K of four sampling schemes. a. Scheme I. b. Scheme II. c. Scheme III. d. Scheme IV.



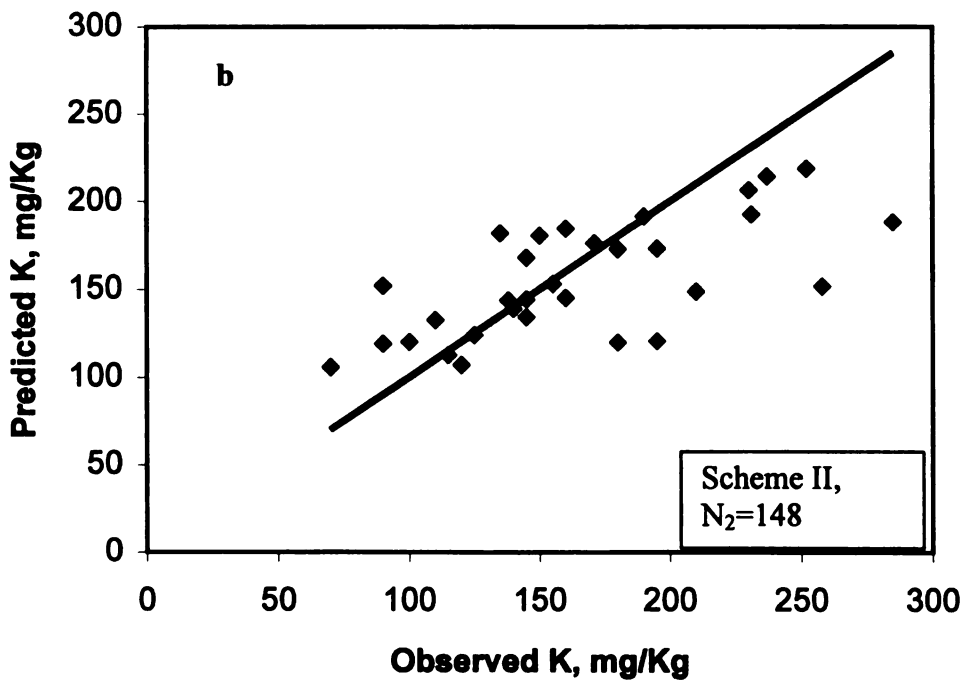
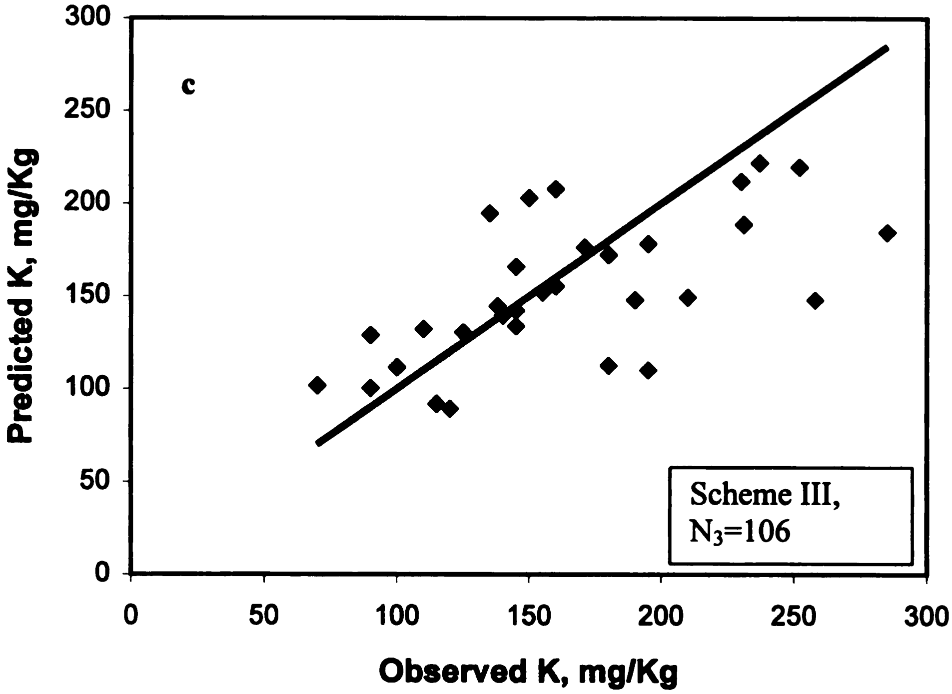
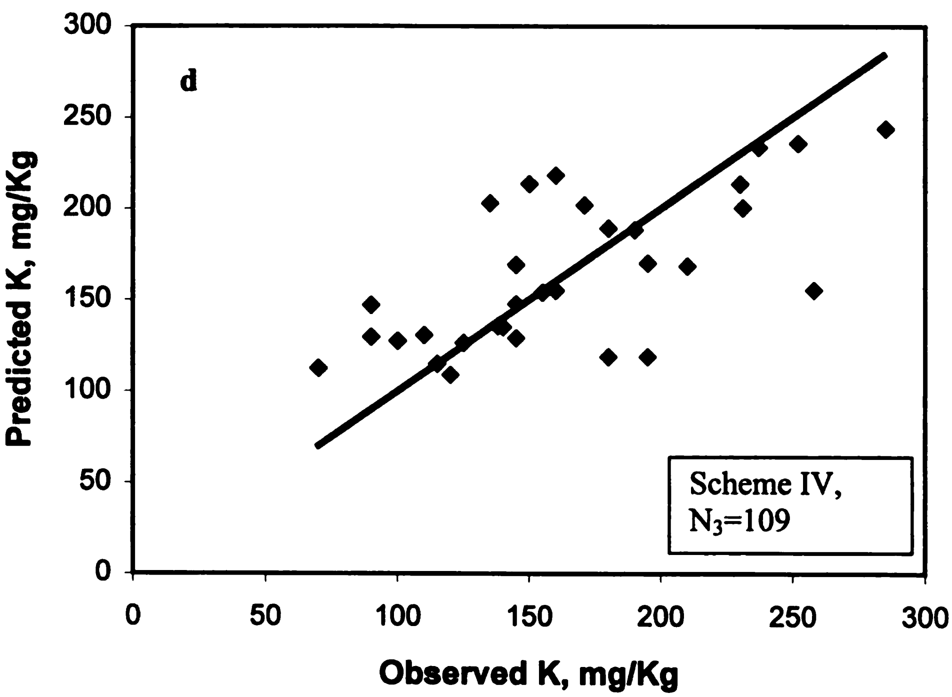


Figure 2.6. (cont'd).





Conclusion

The effects of sampling scheme on the spatial variation of soil P and K and the accuracy of prediction were investigated by four sampling schemes designed using 189 grid and off-grid samples. Scheme I had all data points (N_1=189), Scheme II had all grid samples and 42 randomly selected off-grid samples (N_2=148), Scheme III had grid samples only (N_3=106), and Scheme IV had 59 grid samples and 50 randomly selected off-grid samples (N_4=109). The sample semivariograms of soil P calculated based on the data from the four different sampling schemes varied considerably, whereas those of K were fairly consistent. Schemes I with the largest number of samples (N_4=189) produced the most accurate predictions for both soil P and K. Schemes IV (N_4=109) performed as accurately as Scheme II (N_2=148) for both P and K, however with approximately 40 samples less. Scheme III (N_3=106) performed noticeably less accurate than Scheme IV (N_4=109), with approximately the same number of samples. The results indicated that the short spaced samples provided more valuable information than regularly spaced samples separated by long distances. All four sampling schemes tended to overestimate low values, and underestimate high values. This indicated that even the most dense sampling scheme (Scheme I) in the study couldn't predict well the local high and low values, and they were somewhat "averaged" by the neighbor values used to predict them. If appreciably higher prediction accuracy is desired, this problem can be improved by checking the residual map for high error areas and taking additional samples in those areas.

Reference

Bailey, T.C., and A.C. Gatrell. 1995. Interactive Spatial Data Analysis. Longman Group Limited.

Cambardella, C.A., T.B. Moorman, J.M. Novak, T.B. Parkin, D.L. Karlen, R.F. Turco, and A.E. Konopka. 1994. Field-Scale Variability of Soil Properties in Central Iowa Soils. Soil Sci. Soc. Am. J. 58:1501-1511.

Campbell, J.B.1978. Spatial Variation of Sand Content and pH within Single Contiguous Delineations of Two Soil Mapping Units. Soil Sci. Soc. Am. J. 42:460-464.

Ferguson, R.B., C.A. Gotway, G.W. Hergert, and T.A. Peterson. 1996. Soil Sampling for Site-Specific Nitrogen Management. Proceedings of 3rd Intl. Conference, Precision Agriculture.

Gotway, C.A., R.B. Ferguson, G.W. Hergert, and T.A. Peterson. 1996. Comparison of Kriging and Inverse-Distance Methods for Mapping Soil Parameters. Soil Sci. Soc. Am. J. 60:1237-1247.

Hemlett, J.M., R. Horton, and N.A.C. Cressie. 1986. Resistant and Exploratory Techniques for Use in Semivariogram Analyses. Soil Sci. Soc. Am. J. 50: 868-875.

Isaaks, E.H., and R.M. Srivastava. 1989. An Introduction to Applied Geostatistics. Oxford University Press, Inc.

Kravchenko, A, and D.G. Bullock. 1999. A Comparative Study of Interpolation Methods for Mapping Soil Properties. Agron. J. 91:393-400.

Laslett, G.M., A.B. McBratney, P.J. Pahl, and M.F. Hutchinson. 1987. Comparison of Several Spatial Prediction Methods for Soil pH. J. Soil Sci. 38:325-341.

Laslett, G.M., and A.B. McBratney. 1990. Further Comparison of Spatial Methods for Predicting Soil pH. Soil Sci. Soc. Am. J. 54:1553-1558.

McBratney, A.B., and R. Webster. 1983. Optimal Interpolation and Isarithmic Mapping of Soil Properties. J. Soil Sci. 34, 137-162.

McBratney, A.B., and R. Webster. 1986. Choosing Functions For Semi-variograms of Soil Properties and Fitting Them to Sampling Estimates. J. of Soil Sci. 37, 617-639.

Mohamed, S.B., E.J. Evans, and R.S. Shiel. 1996. Mapping Techniques and Intensity of Soil Sampling for Precision Farming. Proceedings of 3rd Intl. Conference, Precision Agriculture.

Mulla, D.J. 1991. Mapping and Managing Spatial Patterns in Soil Fertility and Crop Yield. p15-26. *In* P.C. Robert et al. (ed.) Soil Specific Crop Management: A workshop in research and development issues. Proc. Workshop, Minneapolis, MN. 1992. ASA, CSSA, and SSSA, Madison, WI.

Rao, P.S.C., and R.J. Wagenet. 1985. Spatial variability of Pesticides in Field Soils: Methods for Data Analysis and Consequences. Weed Sci. 33 (Suppl. 2):18-24

Rochette, P., R.L. Desjardins, and E. Pattey. 1991. Spatial and Temporal Variability of Soil Respiration in Agricultural Fields. Can. J. Soil Sci. 71:189-196.

Sutherland, R.A., C. van Kessel, and D.J. Pennock. 1991. Spatial Variability of Nitrgen-15 Natural Abundance. Soil Sci. Soc. Am. J. 55:1339-1347.

Trangmar, B.B., R.S. Yost, and G. Uehara. 1985. Application of Geostatistics to Spatial Studies of Soil Properties. Advances in Agron. Vol.38.

Vieira, S.R., D.R. Nielsen, and J.W. Biggar. 1981. Spatial Variability of Field-Measured Infiltration Rate. Soil Sci. Soc. Am. J. 45:1040-1048.

Weber, D.D., and E.J. Englund. 1994. Evaluation and Comparison of Spatial Interpolators II. Math. Geol. 26:589-603.

Webster, R., and H.E. Cuanalo DE LA C. 1975. Soil Transect Correlograms of North Oxfordshire and Their Interpretation. J. Soil Sci. 26, 176-194.

Webster, R, and M.A. Oliver.1992. Sample Adequately to Estimate Variograms of Soil Properties. J. of Soil Sci. 43, 177-192.

Wibawa, W.D, D.L. Dludlu, L.J. Swenson, D.G. Hopkins, and W.C. Dahnke. 1993. Variable Fertilizer Application Based on Yield Goal, Soil Fertility , and Soil Map Unit. J. Prod. Agric. 6:255-261

Wollenhaupt, N.C., R.P. Wolkowski, and M.K. Clayton. 1994. Mapping Soil Test Phosphorus and Potassium for Variable-Rate Fertilizer Application. J. Prod. Agric. 7:441-448.

APPENDICES

Appendix 1. Summary statistics of monitor yield data from 1996 - 2001.

Year	Crop	Number of data points	Mean	Max.	Min.	S.D.
			---------------------Bu/acre------------------			
			Field 1			
1996	Corn	35,567	105.6	176.2	32.6	16.7
1997	Soybean	27,302	39.0	65.1	15.3	6.4
1998	Corn	31,156	161.1	218.5	83.8	15.1
1999	Soybean	32,061	33.9	54.2	15.1	4.9
2000	Corn	31,379	152.5	215.1	86.2	15.1
2001	Corn	23,789	123.5	189.9	49.3	15.1
			Field 2			
1997	Corn	17,636	147.3	224.6	21.7	27.2
1999	Corn	17,419	145.0	248.9	10.1	27.5
2001	Corn	12,968	120.9	193.0	49.3	20.3

Appendix 2. Summary statistics of the selected soil properties from the A and B horizons.

Soil Properties	Max.	Min.	Mean	S.D.
		A Horizon		
Horizon Thickness (cm)	32.5	17.5	24.0	3.8
Bulk Density (g/cm^3)	1.9	1.2	1.4	0.15
Total Clay (%)	19.1	3.8	13.6	3.93
Total Silt (%)	66.1	12.5	47.7	12.56
Total Sand (%)	83.6	17.8	38.6	15.86
Fine Silt (%)	46.6	7.7	29.2	9.56
Coarse Silt (%)	24.4	4.8	18.6	4.70
Very Fine Sand (%)	14.3	4.9	10.0	2.40
Fine Sand (%)	56.3	5.9	17.6	11.41
Medium Sand (%)	15.4	2.3	7.0	3.26
Very Coarse Sand (%)	5.8	1.2	2.9	1.11
Coarse Sand (%)	5.3	0.1	1.2	1.15
Ca (meq/100g)	12.6	1.9	5.9	2.33
Mg (meq/100g)	2.4	0.4	1.2	0.45
K (meq/100g)	1.2	0.1	0.4	0.22
Sum of Basis (meq/100g)	14.9	2.5	7.5	2.69
Acidity (meq/100g)	14.0	2.4	8.0	3.05
CEC (meq/100g)	18.0	3.5	10.8	3.24
Base Saturation (%)	100	37.0	68.8	14.84
Organic Carbon (%)	2.2	0.5	1.3	0.39
Total N (%)	0.3	0.0	0.1	0.05
pH ($CaCl_2$)	6.7	4.4	5.4	0.54
EC (mmho/cm)	1.1	0.2	0.4	0.15

Appendix 2. (cont'd)

Soil Properties	Max.	Min.	Mean	S.D.
		B Horizon		
Horizon Thickness (cm)	60	12.5	27.3	10.2
Bulk Density (g/cm^3)	2.1	1.2	1.5	0.16
Total Clay (%)	28.1	3.4	18.2	5.87
Total Silt (%)	67.1	9.3	40.1	15.66
Total Sand (%)	87.3	15.3	41.7	18.23
Fine Silt (%)	54.4	5.5	24.0	12.67
Coarse Silt (%)	23.9	3.8	16.1	4.94
Very Fine Sand (%)	16.7	2.7	10.3	3.00
Fine Sand (%)	63.2	5.7	20.2	13.15
Medium Sand (%)	17.8	2.2	7.4	4.01
Very Coarse Sand (%)	8.1	0.9	2.9	1.76
Coarse Sand (%)	5.9	0.1	1.0	1.09
Ca (meq/100g)	8.7	0.7	5.7	1.99
Mg (meq/100g)	2.8	0.4	1.2	0.52
K (meq/100g)	1.0	0.1	0.3	0.18
Sum of Basis (meq/100g)	11.6	1.2	7.2	2.41
Acidity (meq/100g)	15.9	1.8	8.1	3.34
CEC (meq/100g)	18.4	2.4	12.0	3.71
Base Saturation (%)	95.0	33.0	61.7	14.72
Organic Carbon (%)	2.0	0.1	0.6	0.47
Total N (%)	0.2	0.0	0.1	0.04
pH ($CaCl_2$)	6.2	4.0	4.9	0.54
EC (mmho/cm)	1.7	0.2	0.5	0.52