Precision Agriculture, 3, 407–417, 2002

Soil Phosphorus and Potassium Mapping Using a Spatial Correlation Model Incorporating Terrain Slope Gradient

BRIAN KOZAR bjk@montana.edu
Department of Land Resources and Environmental Sciences, Montana State University, 334 Leon Johnson Hall, PO Box 173120, Bozeman, MT 59715

RICK LAWRENCE rickl@montana.edu
Department of Land Resources and Environmental Sciences, Montana State University, 334 Leon Johnson Hall, PO Box 173120, Bozeman, MT 59715

DAN S. LONG dlong@montana.edu
Northern Agriculture Research Center, Montana State University, HC 36 Box 43, Havre, MT 59051

Abstract. Variable-rate fertilizer application requires knowledge of the spatial distribution of soil nutrients within fields. Grid soil sampling might be used for acquiring this information, but is often too expensive for resolving spatial patterns in soil nutrients at the scale of precision fertilizer application. The objective of this study was to determine whether grid sampling efficiency can be improved using cokriging estimates with slope gradient as a secondary variable, which is easily obtained from high-resolution digital elevation models. Soils in two northern Montana wheat fields were sampled at the nodes of a 100-m diagonal grid. Soil test phosphorus and potassium maps were constructed with kriging and cokriging. Co-kriging uses the spatial correlation between two variables to predict for the less intensively sampled variable of interest, often with less estimation error than a univariate method such as kriging. The average estimation variance for cokriging compared to kriging was reduced for all values of the correlation considered. The additional complexity of cokriging might be justified provided a secondary variable exists that is spatially cross correlated with the primary variable of interest.

Keywords: Co-kriging, DEM, soil mapping, soil properties, topography

Introduction

Precision agriculture often incorporates precise, spatial information about soil properties and/or nutrients (e.g., phosphorus and/or potassium) across farm fields to help meet the goal of optimizing input of fertilizers and herbicides while maintaining or increasing yields. The evaluation of soil nutrient levels across farm fields is typically performed by taking soil samples, analyzing them for nutrient content, and interpolating values between sampling points (Wollenhaupt et al., 1997). In many fields, the sampling has been conducted on a two-dimensional grid as needed to determine the spatial distribution in soil phosphorus (P) and potassium (K) levels and make site-specific recommendations

for phosphate and potash fertilizers (Wollenhaupt et al., 1994). A major limitation of grid sampling is that a large number of samples are required to resolve the spatial variability in soil nutrient levels, which can be highly complex within fields. Unfortunately, collection and analysis costs make this approach impractical on a large scale (Varvel et al., 1999).

The geostatistical procedure of kriging, which is based on regionalized variable theory, is preferred for interpolation of grid data because it allows one to exploit the spatial correlation between neighboring observations to predict attribute values at unsampled locations (Isaaks and Srivastava, 1989). In general, kriging has provided better estimates of soil properties than conventional interpolation procedures such as inverse square distance (Tabios and Salas, 1985 and Laslett et al., 1987). To this end, kriging involves the use of the variogram. A variogram has parameters that assist in describing the spatial correlation: the nugget, sill, and range. Correlations that occur most strongly at shorter distances and weaken with increasing distance indicate that spatially continuous data, including most soil properties, are strongly positively autocorrelated (Bailey and Gatrell, 1995). Comparison studies have shown that variability in soil attributes such as electrical conductivity, bulk density, pH, nutrient content, and particle distribution can be assessed better using variograms over traditional correlation models (Morkoc et al., 1987, Bresler et al., 1988, and Entz et al., 1991). In addition to providing a measure of prediction error, a major advantage of kriging over simpler methods is that sparsely sampled observations of a primary soil variable can be complemented by secondary variables that are more densely sampled.

The multivariate extension of kriging, known as co-kriging, is typically used for combining the primary and secondary attribute data. Co-kriging improves the reliability of interpolation between points by exploiting the spatial cross-correlation between the primary variable of interest that is linearly related to the secondary variable, or covariate (Isaaks and Srivastava, 1989). When spatial correlation exists, the variables are said to be co-regionalized. When two variables are co-regionalized, and the secondary variable is over-sampled with respect to the primary variable, then the sampling efficiency of the primary variable of interest is effectively increased (Yates and Warrick, 1987). Co-kriging makes use of the cross-variogram function to transfer the spatial information in the secondary variable to the primary, thus improving the reliability of the interpolation process (Yates and Warrick, 1987, Zhang et al., 1997, and Trangmar et al., 1986). Co-kriging has been used to interpolate soil chemicals, NaHCO—P, sodium adsorption ratio, and soil properties using their spatial covariances with secondary variables such as depth of the soil profile, HCL, electrical conductivity, and terrain indices, respectively (Zhang et al., 1997, Trangmar et al., 1986, Pozdnyakova and Zhang, 1999, and McKenzie and Ryan, 1999). Thus, co-kriging has the potential to result in a time and economic savings in mapping soil properties for which there are correlated variables (Trangmar et al., 1999).

The spatial distribution of soil physical and chemical properties is often related to the spatial distribution of other environmental variables with which they are correlated (Trangmar et al., 1986, Bhatti et al., 1991, and Bailey and Gatrell, 1995). A valuable and inexpensive source of secondary information is the digital elevation model (DEM). Digital elevation models have been found useful for revealing previously uncaptured spatial variability in land and soil properties, and providing explanatory variables for predictive modeling of soil properties (Moore et al., 1993, and Gessler et al., 1995; 2000). In addition, by including readily available and inexpensive terrain data derived

from a DEM, co-kriging was found to be useful for predicting primary soil variables that were more costly or difficult to sample (Odeh et al., 1994; 1995). The rationale for this approach is that topography influences soil properties due to local re-distribution of water, solar radiation, and soil materials (Gessler et al., 2000).

If the distribution of soil P or K is affected by translocation, which is a function of slope gradient, co-kriging of either soil test P or K along with a more easily obtainable DEM would potentially improve prediction results. In this study, soil test results for P and K, grid sampled from a field in northern Montana, were interpolated using ordinary co-kriging that combined these soil data with slope gradient derived from a DEM. Our objective was to determine if relationship existed between these soil test data and slope gradient derived from a DEM, and whether these relationships could be used with co-kriging to increase efficiency of grid sampling for precision nutrient management programs.

Methods

The study site is a 500 m by 2000 m portion of a farm field located in Liberty County, Montana, 13 miles north of Chester, MT (T34N, R7E, Section 17; 48°42′30″ 110°51′30″). The site has relief of 10 to 30 meters, has rolling terrain accentuated by deep swales, and has been dryland farmed with a wheat/fallow rotation. The climate is cool, semi-arid, with moderate to long winters and 100–110 frost-free days, and experiences a mean annual air temperature of 4.4°C (40°F) (Caprio et al., 1994). The area receives mean annual precipitation of about 25.4 to 30.8 cm (10 to 12 inches), with mean annual snowfall of about 63.5 to 127 cm (25 to 50 inches) (Caprio et al., 1994). The underlying geologic materials are primarily glacial till derived from the Bear Paw and Judith River formations (Veseth and Montagne, 1980). Soils are generally moderately deep and range in depth from 25 to >100 cm. We sampled soil phosphorus and potassium at 100 locations spaced 80 to 100 meters apart in a rough grid like pattern.

A survey-grade GPS receiver was used to collect a fine-resolution grid of easting, northing, and elevation points along transects over the study site. The elevation points were interpolated to a 5-m square grid with the commercial mapping software package Surfer (Golden Software) to create a fine resolution DEM. Phosphorus and potassium point sample data and the DEM were stored in a geographic information system. Slope values were derived from the DEM for each point at which soil sampling occurred. Data for each sample point were then imported into the statistical software package S-Plus data frame for geostatistical analysis, along with 400 additional random elevation points.

Exploratory data analysis (EDA) was conducted for the soil nutrients and for slope values to look for any possible trends that might indicate lack of stationarity. Observed trends were removed using linear regression. Semivariograms were calculated individually for phosphorus, potassium, and slope gradient, using the estimator:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})]^2, \quad (1)$$

where $z(\mathbf{x}_i)$ and $z(\mathbf{x}_i + \mathbf{h})$ is the variable of interest at locations $\mathbf{x}_i$ and $\mathbf{x}_i + \mathbf{h}$ and $N(\mathbf{h})$ is the number of point pairs separated by $\mathbf{h}$. Directional variograms were computed to

identify whether the spatial autocorrelation had a directional component, or anisotropy. Calculation of the co-regionalization was then determined using the cross-variogram estimator:

$$\hat{y}_{12}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z_1(\mathbf{x}_i) - z_1(\mathbf{x}_i + \mathbf{h})][z_2(\mathbf{x}_i) - z_2(\mathbf{x}_i + \mathbf{h})], \tag{2}$$

with definitions as in Eq. (1). The semivariograms and cross-variograms were further characterized by fitting them to a spherical model, using a generalized least squares estimation approach. This process allowed for modeling the covariance structure of the data while eliminating second-order, trend effects (Bailey and Gatrell, 1995).

Parameters obtained from the semivariograms were used in a kriging procedure to predict for phosphorus and potassium. The prediction area encompassed, but did not exceed, the area where the original sample points were located. The estimated value is a weighted average of the data set:

$$\hat{z}(\mathbf{x}_0) = \sum_{i=1}^{N} \lambda_i z(\mathbf{x}_i), \tag{3}$$

where $\hat{z}(\mathbf{x}_0)$ is the estimated value, N is the number of observations $z(\mathbf{x}_i)$, and λ_i, $(i = 1, \ldots, N)$ are the weights assigned to the sampling points. Parameters obtained from the semivariograms and the cross-variogram were then used in a co-kriging procedure to predict for phosphorus and potassium based on the correlation between the soil nutrient values and the slope gradient values. The prediction area is the same as that for the kriging procedure. The estimated value is calculated through:

$$\hat{z}_1(\mathbf{x}_0) = \sum_{i=1}^{N_1} \lambda_{1i} z_1(\mathbf{x}_{1i}) + \sum_{j=1}^{N_2} \lambda_{2j} z_2(\mathbf{x}_{2j}), \tag{4}$$

where N_1 and N_2 are the number of neighbors of z_1 and z_2, respectively, and λ_{1i} and λ_{2j} are the weights associated to each sampling point. A full, comprehensive discussion of semi- and cross-variograms, the models and parameters used to characterize them, and kriging procedures is beyond the scope of this paper. For a detailed description and discussion of models, parameters, and kriging and co-kriging procedures, see, e.g., Bailey and Gatrell (1995).

Cross-validation of the predicted values obtained from both the kriging and co-kriging procedures was then conducted to determine the accuracy of both procedures for predicting phosphorus and potassium. Cross-validation systematically removed each data point from the original dataset and used the remaining observations to estimate that data value. The estimated value was then compared to the true value for evaluating a goodness of fit of both the kriging and co-kriging models. Comparison of the resulting value, the mean squared difference, between the two procedures determined which procedure was more accurate. The mean squared difference, or prediction variance, is calculated by:

$$\sigma_e^2 = \frac{1}{N} \sum_{i=1}^{N} [z(\mathbf{x}_i) - \hat{z}_{-i}(\mathbf{x}_i)]^2, \tag{5}$$

where N is the number of observations, $z(\mathbf{x}_i)$ is the actual value of variable being predicted at $\mathbf{x}_i$, and $\hat{z}_{-i}(\mathbf{x}_i)$ is the predicted value at location $\mathbf{x}_i$ without $z(\mathbf{x}_i)$. The lower the prediction variance, the more accurate is the prediction model.

Table 1. Descriptive data analysis of variables for the dataset

Variable	Number	Mean	Median	Min	Max	SD
Potassium (mg kg^{-1})	100	452.2	424	132	779	142.3
Phosphorus (mg kg^{-1})	100	13.6	12	2	89	10.4
Slope (degrees)	433	5.98	5	1	20	2.3

SD = Standard deviation.

Results

Potassium demonstrated greater variability relative to the mean than did phosphorus (Table 1). The soil variables were in general normally distributed. Phosphorus contained two rather high outliers with values of 42 and 89 mg kg^{-1}. Due to the effect that these outliers may have on the structure of the variogram, these outliers were removed from the subsequent analysis. Normalization of the data did not result in significantly lower variances in the data, so data were not transformed for analysis.

EDA for phosphorus levels revealed a directional trend, as levels were generally higher in the eastern side of the plot at the higher terrain (Figure 1). The linear regression detrending model fitted phosphorus as a function of the x and y coordinates.

$$\textit{phosphorus} = 2813.87 + .0179(x) - .0022(y) \tag{6}$$

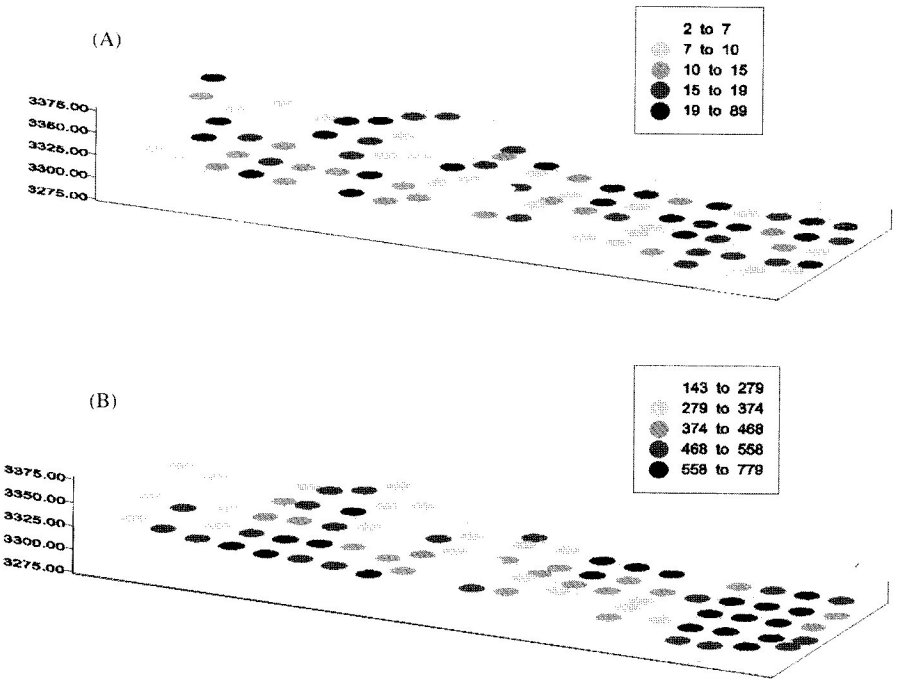

Figure 1. Sampled phosphorus (A) and potassium (B) levels overlayed onto the area landscape for purposes of detecting trend. Actual sample points measure $<1 m^2$ and size of points on figure are for visual purposes only.

The detrending model was statistically significant (p-value = 0.02), and, although the model only explained a small portion of the variability in the data ($\mathbf{R}^2 = 0.08$), it was useful in removing some trend in the data and improving the variogram fit. The residuals of the detrending model, therefore, were used to create a variogram for phosphorus. The potassium data also exhibited a moderate directional trend (Figure 1). Levels were generally higher beginning in the western portions where the terrain contained a mild northwest facing slope, then lower as the terrain undulated back down on a southeast facing slope, then higher again on another northwest facing slope in the eastern portion of the plot. The linear regression model chosen to detrend the potassium data fitted potassium as a function of elevation and the sin of the $\boldsymbol{x}$ coordinates.

$$\boldsymbol{potassium} = 12152 - 3.547(\boldsymbol{elev}) - 47.886(\sin(\boldsymbol{x})) \tag{7}$$

The sin function represented the undulation of the terrain. This model resulted in a **R**-squared of 0.28 with a significant p-value of <0.001.

Semi-variograms for each of the two soil variables (Figure 2 and Table 2) and for slope (Figure 3 and Table 2) were generated, along with a cross-variogram for potassium and slope (Figure 3 and Table 2), after the removal of trend, as discussed earlier. The variogram for phosphorus illustrated a pure nugget effect, i.e., the nugget (small scale spatial variability) equaled the variance of the data. This corresponded to a complete lack of spatial dependence between phosphorus samples. Due to this lack of spatial dependence, no model could be fitted to the semi-variogram that would yield useful predictive value (Bailey and Gatrell, 1995). No predictions of phosphorus based on slope gradient could therefore be constructed using semi- and cross-variograms. Potassium and the slope gradient did exhibit spatial structure and stationarity after detrending. The semi-variogram for slope and the cross-variogram were assumed isotropic, while an anisotropic calculation was used for the potassium semi-variogram due to directional dependency. The crossvariogram for potassium and slope showed a strong cross correlation. The crossvariogram demonstrated that a positive increment in slope value corresponded to a negative increment in potassium. As slope gradient increases, therefore, the amount of potassium in the soil decreases.

A total of 334 equally spaced point estimates of potassium were generated in a rectangular grid, using both the kriging and co-kriging methods. The prediction are encompassed, but did not exceed, the area from which the original 100 samples were taken.

The predicted values obtained from the kriging and co-kriging procedures were then plotted in a map for purposes of visual comparison (Figure 4). Cross-validation resulted in a range of residuals from the original values of -289 to 170, although a majority of the residuals fell in the range of -50 to 50. Kriging resulted in a prediction variance of 9,769.2, while co-kriging resulted in a prediction variance of 5,430.7. This translated into a 44% improvement in prediction variance when employing co-kriging rather than kriging for estimating potassium across our filed.

Discussion

Both phosphorus and potassium displayed first order variation across the field, indicating large-scale trends in distribution of these soil nutrients. Sample size and spatial intensity

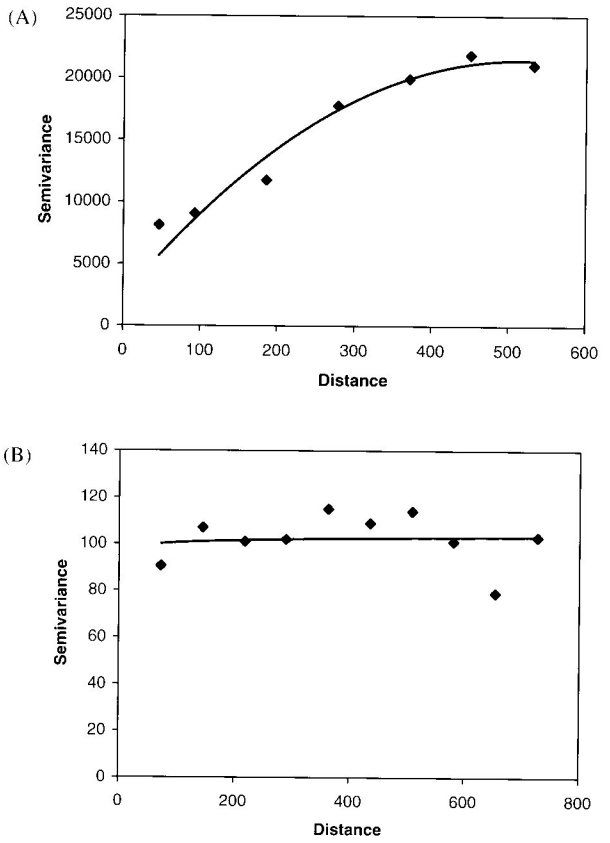

Figure 2. Semi-variograms of potassium (A) and phosphorus (B).

of the sampling scheme influence the quality of semi- and cross-variograms. Phosphorus failed to demonstrate autocorrelation at the scale of our sampling necessary for the construction of the semi-variogram, both before and after potential outliers were removed. This result indicated that phosphorus varied at scales below 80 to 100 m. Thus, any attempt to accurately estimate soil phosphorus at our study site would require intense sampling at distances considerably shorter than 100 m. Potassium showed a field-wide trend related both to elevation and the rolling terrain. In addition, autocorrelation was

Table 2. Parameters of the semivariograms and cross-variogram

	Lag	Azimuth	MaxD	Tol	Sill	Nugget	Range
Potassium	100	0	800	11.5	20140	2349	416
Slope	100	none	800	none	16.7	4.5	580
K/Slope	100	none	800	none	−357	4.11	499

MaxD = maximum distance to compute variograms. Tol = tolerance angle of azimuth. Nugget, and Range measured in meters.

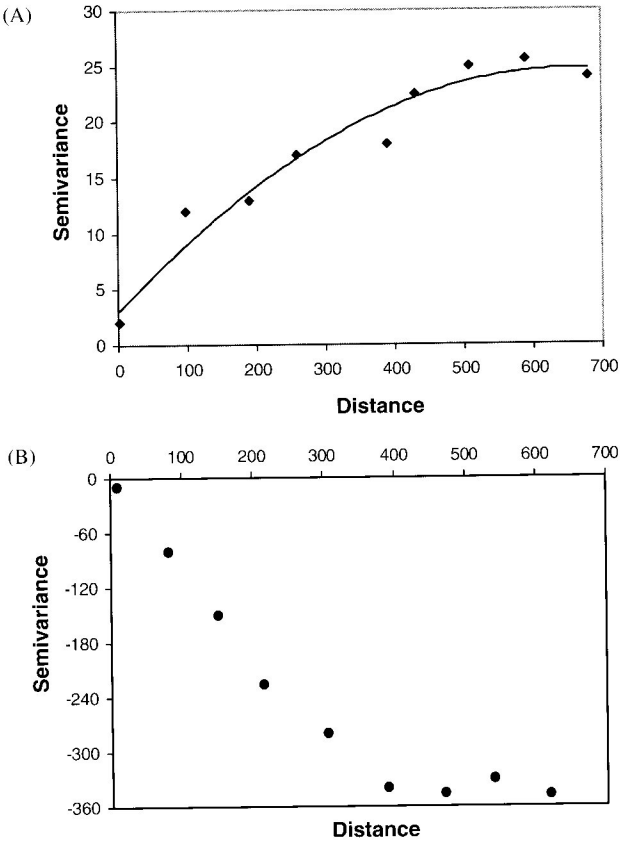

Figure 3. Semi-variogram of slope (A) and cross-variogram of potassium and slope (B).

evident at our sample separation, indicating that predictions could be made through estimation procedures using values of potassium sampled at this intensity (80 to 100 m). The difference in spatial dependence, or autocorrelation, for the two soil nutrients suggested that, for the purpose of spatial modeling, optimal distances between samples varies among soil properties. The negative autocorrelation between potassium and slope meant that for this field steeper slopes had lower potassium levels. This relationship also meant that knowledge of slope gradient, which was easily derived, could improve our estimates of soil potassium for this field.

A comparison between the original sampled potassium values and the kriged and co-kriged estimates (Figure 4) indicated that much of the prediction error occurred where there were steeper slopes and on the knoll area near the middle of the field. In these areas potassium was primarily under-predicted. Co-kriging, however, did result in markedly increased accuracy among these areas. Based on our cross-validation, co-kriging with the terrain index of slope improved our mapping of soil potassium by 44% when compared to kriging. The results indicated that allowing slope angle to be incorporated into the prediction of potassium through co-kriging accounted for some of the variability necessary for accurate predictions.

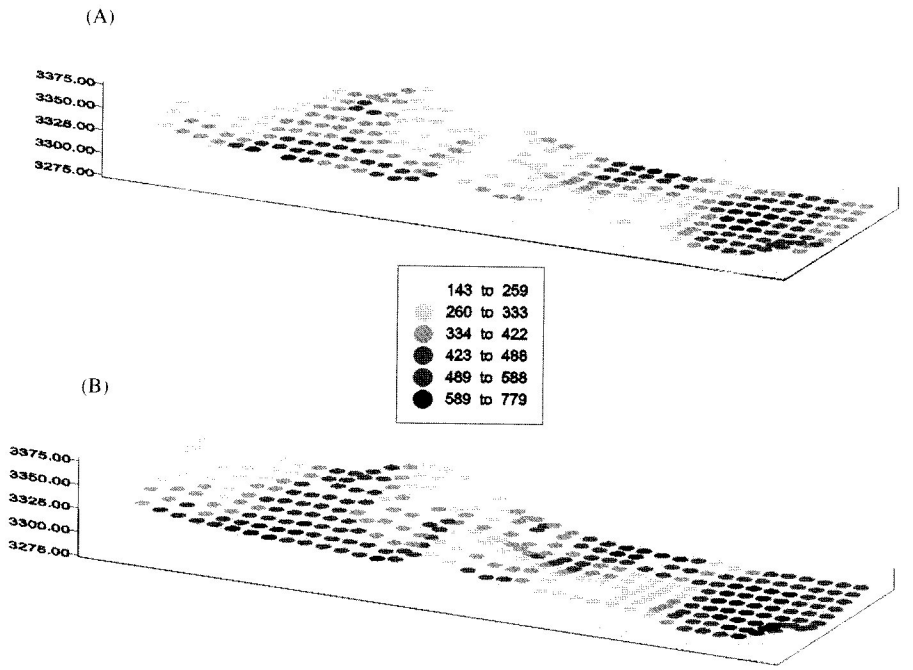

Figure 4. Potassium estimated by the kriging (A) and co-kriging (B) procedures.

The influence topography has on soil formation, and the strong role topography can play in the spatial distribution of soil properties, can result in strong correlation between soil properties and terrain indices. Our results suggested that co-kriging with terrain indices can take advantage of this correlation by making a prediction for a soil property, based not only on surrounding values of that soil property, but also on surrounding values of the terrain index as well. This method can be particularly useful when the soil property (primary variable) is undersampled, and the value of the terrain index (secondary variable) at the locations being predicted is known. Here co-kriging should show the most improvement over kriging.

The correlation discerned here between slope and potassium suggests translocation of potassium within the soil profile. However, the relationship between terrain and potassium is influenced by other environmental factors present in the system, such as geology, climate, and other soil properties, and the processes that occur between them. Furthermore, as the processes that define these relationships vary between fields, it is likely that so to will the relationship between potassium and terrain vary between fields. Therefore, any one particular geostatistical model will have varied levels of success from field to field. Geostatistics presents an exciting option for prediction of soil properties and nutrients, but these techniques are only beneficial if they can accurately model the system in which it is being used. Other terrain indices (e.g., plan and profile curvature; compound topographic index) might capture the processes that underlie the relationship between topography and the spatial distribution of soil properties more accurately, dependent on the system in which the analysis is being conducted.

Improved estimates of soil nutrients at sub-field scales are critical for the success of precision farming. A detailed survey of individual fields is generally not practical, and methods that incorporate field sampling are necessary. This study demonstrated that geostatistical methods can provide reasonable estimates of soil nutrients, if sampling is sufficiently intense to capture the scale at which the nutrients vary. Incorporation of correlates, such as terrain variables that influence soil formation and nutrient distribution, can improve these estimates dramatically and can allow for further reduction in sampling. Furthermore, these techniques can aid us in understanding the processes that occur between topography and soils that drive the spatial distribution of soil properties and nutrients across landscapes. Given increased understanding of these processes and the system in which they occur, this can in turn allow for further implementation of geostatistical and terrain modeling prediction techniques.

Acknowledgments

We thank Carl and Janice Mattson, Chester, Montana, for giving us permission to use their fields for this and other research. We also thank Dr. Robin Reich, Associate Professor of Forest Biometrics, Department of Forest Sciences, Colorado State University, for his assistance with the statistical analysis.

References

Bailey, T. C. and Gatrell, A. C. 1995. *Interactive Spatial Data Analysis* (Addison Wesley Longman Limited, Essex, England), p. 162–166, 176–181, 213–215.

Bhatti, A. U., Mulla, D. J. and Frazier, B. E. 1991. Estimation of soil properties and wheat yields on complex eroded hills using geostatistics and Thermatic Mapper images. Remote Sensing of Environment **37**, 181–191.

Bresler, E. and Laufer, A. 1988. Statistical inferences of soil properties and crop yields as spatial random functions. Soil Sci. Soc. Am. J. **52**, 1234–1244.

Caprio, J. M., Cooksey, D. I., Jacobsen, J. S., Nielsen, G. A. and Roche, R. R. 1994. MAPS Atlas Version 5.0, Montana State University Extension Bulletin 125, Bozeman, MT.

Entz, T. and Chang, C. 1991. Evaluation of soil sampling schemes for geostatistical analyses: A case study for bulk density. Can J. Soil Sci. **71**, 165–176.

Gessler, P. E., Moore, I. D., McKenzie, N. J. and Ryan, P. J. 1995. Soil-landscape modelling and spatial prediction of soil attributes. Int J. Geog. Inf. Sys. **9**(4), 421–432.

Gessler, P. E., Chadwick, O. A., Chamran, R., Althouse, L. and Holmes, K. 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Sci. Soc. Am. J. **64**, 2046–2056.

Laslett, G. M., McBratney, A. B., Pahl, P. J. and Hutchinson, M. F. 1987. Comparison of several spatial prediction methods for soil pH. J. Soil Sci. **38**, 325–341.

McKenzie, N. J. and Ryan, P. J. 1999. Spatial prediction of soil properties using environmental correlation. Geoderma **89**(1–2), 67–94.

Moore, I. D., Gessler, P. E., Nielsen, G. A. and Peterson, G. A. 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. **57**, 443–452.

Morkoc, F., Biggar, J. W., Nielsen, D. R. and Myers, D. E. 1987. Kriging with generalized covariances. Soil Sci. Soc. Am. J. **51**(5), 1126–1131.

Odeh, I. O. A., McBratney, A. B. and Chittleborough, D. J. 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma **63**(3–4), 197–214.

Pozdnyakova, L. and Zhang, R. 1999. Geostatistical analyses of soil salinity in a large field. Precision Agriculture **1**(2), 153–165.

Tabios, G. Q. and Salas, J. D. 1985. A comparative analysis of techniques for spatial interpolation of precipitation. Water Resources Bulletin **21**, 365–380.

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

Varvel, G. E., Schlemmer, M. R. and Schepers, J. S. 1990. Precision Agriculture **1**, 291–300.

Veseth, R. and Montagne, C. 1980. Geologic parent material of Montana soils. Montana Agric. Exp. Sta. Bulletin 721 (Montana St. Univ., Bozeman, MT and USDA-SCS, Bozeman, MT).

Wollenhaupt, N. C., Wolkowski, R. P. and Clayton, M. K. 1994. Mapping soil test phosphorus and potassium for variable-rate fertilizer application. J. Prod. Agric. **7**, 441–448.

Wollenhaupt, N. C., Mulla, D. J. and Gotway-Crawford, C. A. 1997. *American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America*, edited by F. J. Pierce and E. J. Sadler (Madison, WI), p. 19–54.

Yates, S. R. and Warrick, A. W. 1987. Estimating soil water content using cokringing. Soil Sci. Soc. Am. J. **51**, 23–30.

Zhang, R., Shouse, P. and Yates, S. 1997. Use of pseudo-crossvariograms and co-kriging to improve estimates of soil solute concentrations. Soil Sci. Soc. Am. J. **61**(5), 1342–1347.