DIVISION S-8—NUTRIENT MANAGEMENT & SOIL & PLANT ANALYSIS

Comparison of Methods for Interpolating Soil Properties Using Limited Data

C. A. Schloeder,* N. E. Zimmerman, and M. J. Jacobs

ABSTRACT

Spatial interpolation methods are frequently used to characterize patterns in soil properties over various spatial scales provided that the data are abundant and spatially dependent. Establishing these criteria involved comparisons of abundant data from many fine-scaled (<100 ha) investigations. In this study we investigated whether it was appropriate to use spatial interpolation methods with limited (n = 46), coarse-scaled (1188 ha) soils data from a Vertisol plain. Methods investigated included ordinary kriging, inverse-distance weighting, and thin-plate smoothing splines with tensions. Comparison was based on accuracy and effectiveness measures, and analyzed using ANOVA and pairwise comparison *t*-tests. Results indicated that spatial interpolation was appropriate when the data exhibited smooth and consistent patterns of spatial dependency within the study area and the selected ranges of estimation and weighting used in this investigation. Nine of twelve soil properties we investigated exhibited characteristics other than these, however, including independent data, variable and erratic behavior, and extreme values. Our sample design may have been an important factor as well. Ordinary kriging and inverse-distance weighting were similarly accurate and effective methods; thin-plate smoothing splines with tensions was not. Results illustrate that sample size is as important for coarse-scale investigations as it is for fine-scale investigations with most soils data. However, our ability to predict successfully with some of our data raises the question as to the exact nature of the relationship between accuracy, sample size, and sample spacing, and to what extent these factors are related to the property under investigation, particularly when data are limited.

SPATIAL INTERPOLATION METHODS offer a means of characterizing a variety of factors or responses over different spatial scales. Characterization over different spatial scales has proven invaluable for pest management (Weisz et al., 1995), crop and soil management (Hosseini et al., 1994; Wollenhaupt et al., 1994), and soil properties mapping (Uehara et al., 1985; Gotway et al., 1996). Under the right circumstances it may also prove invaluable to elucidating soil–vegetation interrelationships.

Spatial interpolation methods differ from classic modeling approaches in that they incorporate information about the geographic position of the sample points (Journel and Huijbregts, 1978; Cliff and Ord, 1981; Isaaks and Srivastava, 1989; Watson, 1992; Cressie, 1993). Some methods also have the benefit of incorporating information about the degree and extent of dependence between measurements at different locations. Methods currently in use include kriging, inverse-distance weighting, and thin-plate smoothing splines, and one can use numerous approaches with each of these (see Journel and Huijbregts, 1978; Cliff and Ord, 1981; Isaaks and Srivastava, 1989; Watson, 1992; Cressie, 1993). There is some debate, however, as to which is the best or most appropriate method (Voltz and Webster, 1990; Laslett, 1994; Hosseini et al., 1994; Wollenhaupt et al., 1994; Gotway et al., 1996).

Use of any spatial interpolation method is currently based on a minimum sample size or pairwise comparison criterion and certain characteristics of the data (Journel and Huijbregts, 1978; Isaaks and Srivastava, 1989). The sample size and pairwise comparison criterion are important because of their effect on results. Specifically, it has been demonstrated that accuracy improves as sample size or the number of possible pairwise comparisons increases (Journel and Huijbregts, 1978; Uehara et al., 1985; Isaaks and Srivastava, 1989; Englund et al., 1992; Wollenhaupt et al., 1994). Map resolution may decline with fewer samples as well (Uehara et al., 1985; Gotway et al., 1996). Accuracy also depends on sample pattern and sample spacing (Voltz and Webster, 1990; Englund et al., 1992; Laslett, 1994; Wollenhaupt et al., 1994; Gotway et al., 1996). Data characteristics of importance include the coefficient of variation, skewness, and kurtosis, and whether the data contains outliers or extreme values. These are considered important because it is not certain whether data transformation improves spatial interpolation (Isaaks and Srivastava, 1989), increases accuracy (Weber and Englund, 1992), or has little effect on results (Cooke et al., 1993). It is also not certain whether highly variable data affects accuracy in general (Hosseini et al., 1994; Laslett, 1994) or the extent to which extreme values exert an influence on results (Isaaks and Srivastava, 1989; Gotway et al., 1996).

Past research has focused primarily on how soil resources vary along a transect (Trangmar, 1984; Voltz and Webster, 1990) or at fine scales (<100-ha study area) (Burgess and Webster, 1980; Diaz et al., 1992; Wollenhaupt et al., 1994; Gotway et al., 1996). Exceptions include Uehara et al. (1985) analyses of different data sets from Africa, and Hosseini et al. (1994) analyses of soils data from southwest Iran. Financial and logistical constraints are often the reasons why there are not more

C.A. Schloeder, P.O. Box 445, Fortine, MT 59918-0445; N.E. Zimmerman, Dept. Landscape Inventories, Swiss Federal Institute of Forest, Snow, and Landscape Research, Zuercherstrasse 111, CH-8903 Birmensdorf (ZH), Switzerland; M.J. Jacobs, P.O. Box 445, Fortine, MT 59918-0445. Received 29 Nov. 1999. *Corresponding author (mjjcas@libby.org).

Published in Soil Sci. Soc. Am. J. 65:470–479 (2001).

Abbreviations: BS, base saturation; CEC, cation-exchange capacity; G, goodness-of-prediction (estimate); ID, inverse-distance weighting; MAE, mean absolute error; MSE, mean-squared error; OK, ordinary kriging; OM, organic matter; P, total available phosphorus.

coarse-scaled soil investigations. These same constraints have also limited the amount of data one can collect. Lack of abundant data in turn is the reason for fine-scaled characterization by means other than spatial interpolation methods.

To maximize success with coarse-scaled soils investigations, it is important to understand the extent to which limited data from a large geographic area can affect the decision to use spatial interpolation methods. The objective of this study was to investigate the appropriateness of using spatial interpolation methods with limited surface soils data ($n = 46$) collected from a large area (1188 ha). Methods investigated included ordinary kriging and various approaches to inverse-distance weighting and thin-plate smoothing splines with tensions. The surface soils data used in this study were collected from the Omo Plain, located in Omo National Park, Ethiopia.

Review of Spatial Interpolation Methods

Interpolation using either the ordinary kriging (OK) or inverse-distance weighting (ID) methods presumes that the predictions are a linear combination of the available data, that is:

$$\hat{z}(s_o) = \sum_{i=1}^{n} \lambda_i z(s_i) \qquad [1]$$

where $\hat{z}$ is the predicted value at interpolation point s_o, $z(s_i)$ is the value of variable z at sample point s_i, λ_i is the weight given to observed value $z(s_i)$, and n is the number of observed values used in the estimation. The two methods differ, however, by how the weights are calculated. With OK, the weights are obtained by solving the kriging equation (Isaacs and Srivastava, 1989):

$$\sum_{j=1}^{n} \lambda_j \gamma[d(s_i,s_j)] + m = \gamma[d(s_o,s_i)] \quad i = 1, \ldots, n$$

$$\sum_{i=1}^{n} \lambda_i = 1 \qquad [2]$$

where m is a Lagrange multiplier and $d(s_i,s_o)$ is the distance between s_i and s_o, obtained through the semivariogram:

$$\gamma[d(s_i,s_o)] = \mathrm{var}[z(s_i) - z(s_o)] \qquad [3]$$

As specified by Isaaks and Srivastava (1989), weights are chosen to ensure that the average error for the model is 0 and the model error variance is minimized.

With ID, the weights are instead inversely related to distance (Watson and Phillip, 1985):

$$\lambda_i = \frac{[d(s_i,s_0)]^{-p}}{\sum_{i=1}^{n} [d(s_i,s_o)]^{-p}} \qquad [4]$$

where p is the power parameter that controls how fast the weight of the points tends to zero with increasing distance from the interpolation site. Additionally, the inverse-distance power parameter option also enables one to control the weight assigned to sample points used for prediction to the extent that powers >1 give higher weight to the nearest points (compared with the weights of distant points) and as a result predict a more detailed surface; while powers <1 increase the importance of distant points and predict a more smoothed surface (Watson and Philip, 1985), and the predictions tend toward the sample mean (Isaaks and Srivastava, 1989; Cooke et al., 1993). With ID one is also able to define a maximum distance beyond which sample points are excluded from local predictions (or their relative contribution is close to or is zero).

The thin-plate smoothing splines with tensions (splines) method calculates a two-dimensional minimum curvature spline interpolation. The splines function used for surface interpolation is that described by Mitas and Mitasova (1988):

$$S(x,y) = t(x,y) + \sum_{j=1}^{n} \lambda j R(r_j) \qquad [5]$$

where n is the number of interpolation points, λj and $t(x,y)$ are coefficients found by a solution of a system of linear equations, r_j is the distance from the (interpolation) point to the jth point, while R represents a distance-dependent function (see Eq. [4]), which is controlled by a tension (or weight) parameter ϕ^2:

$$R(r) = -\frac{1}{2\pi\phi^2}\left[\ln\left(\frac{r\phi}{2}\right) + c + K_o(r\phi)\right] \qquad [6]$$

where r is the distance between the (prediction) point and the sample point, c is a constant equal to 0.577215, and K_o is the modified Bessel function.

The tension parameter modifies the minimization criterion so that first-derivative terms are incorporated in the minimization criteria (ESRI, 1991). A weight parameter enables one to control the amount of tension (stiffness) with higher values resulting in a coarser surface (that more closely conforms to the sample points) (Weisz et al., 1995). Lower weight values generate a smoother surface. Weights have to be ≥ 0. Again, the spatial range of influence upon predictions at given locations can be limited. However, it is done by indicating the number of points rather than by selecting a radius. Generally, the greater the number of points selected for local predictions, the smoother the resulting surface.

MATERIALS AND METHODS

Study Area

Omo National Park is located in the lower Omo Basin (Park Headquarters: 805 500 Easting, 648 200 Northing; 35°45′ E, 5°47′ N), just north of Lake Turkana and east of the Sudan–Kenya border (Fig. 1). The Omo Plain is a relatively large (≈70 by 20 km) alluvial plain that spans the eastern half of the park and forms the northern portion of the lower Omo Basin. The plain is bordered on the west by a north–south trending mountain range, on the east by the Omo River, and split north to south by the Mui River. Other features include three low-lying (<800 m) volcanic outcrops in the northern half of the plain, a seasonal spring near the middle volcanic outcrop, and a perennial thermal hotspring in the southern half of the study area (Fig. 2).

The Omo Plain and its associated features are part of a more extensive landscape that formed during the late Miocene and late Quaternary (Davidson, 1983). This landscape devel-

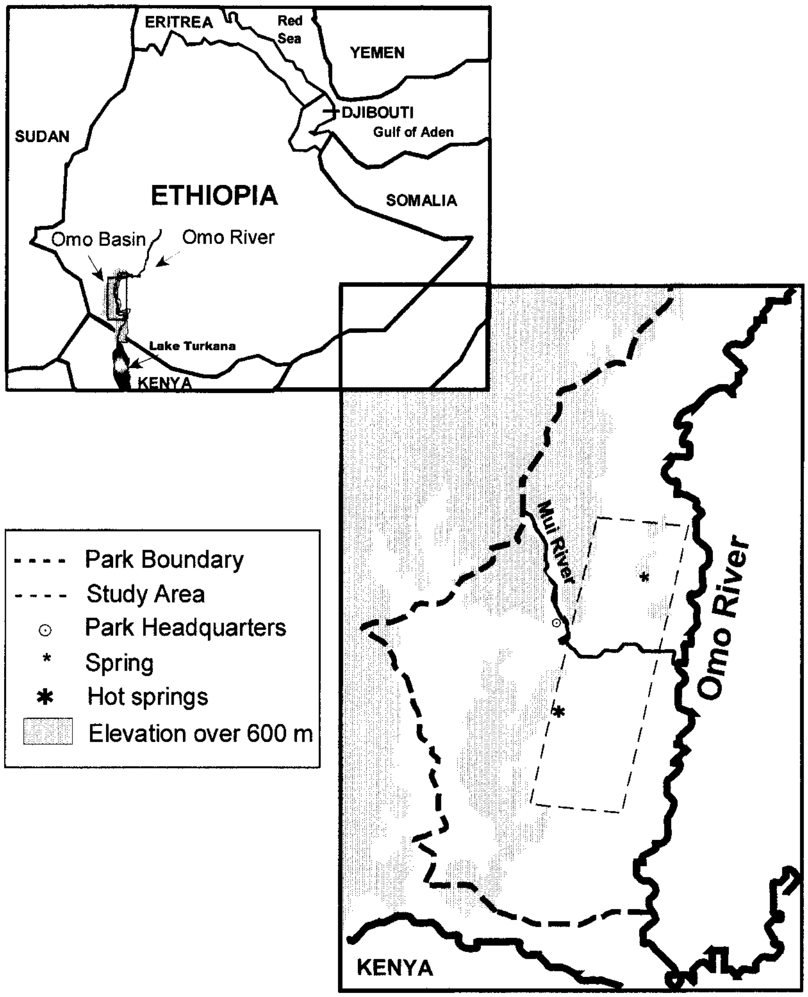

Fig. 1. Location of the lower Omo Basin, Ethiopia.

oped as a result of widespread rift activity (the Rudolf Rift), followed by repeated flooding and meandering events by the Omo and Mui Rivers and alternating changes in the level of ancestral Lake Turkana in the extreme southern part of the basin (≈3–4.3 mA) (Butzer and Thurber, 1969a; Butzer, 1971; Brown and Nash, 1976; Davidson, 1983). Currently, the Omo River rarely floods beyond 5 km and the Mui River has been described as ephemeral, beginning in the early 1970s (Sutcliffe, 1992). According to Butzer and Thurber (1969a), Lake Turkana's historic shoreline extended up to 100 km north of its current delta-fringe.

The alluvium on the Omo Plain is dated at ≈4.2 mA, at least 100 m thick, and primarily of basaltic origin (Brown and Nash, 1976). Colluvium occurs along the uplifted margins of the plain and at the base of the volcanic outcrops in the north. The source of this colluvium is salic-rich crystalline basement rocks and volcanic rock, overlaying a layer of sedimentary rock (Davidson, 1983). The soils associated with the mountain range, volcanic outcrops, and at the base of each, are classified as Stony Cambisols, and Vertisols are the predominant soil type throughout the entire basin (FAO, 1986; Sutcliffe, 1992).

In the northern half of the study area, the adjacent mountains range as high as 2500 m and there is a gradual altitudinal transition between the mountain range and the plain (Fig. 2). South of the Mui River, elevations do not exceed 1100 m and there is an abrupt rather than gradual altitudinal transition between the mountain range and the plain. Topography varies little across the Omo Plain, except for the volcanic outcrops. Average elevation throughout the study area is 450 m; however, the northern half is somewhat higher than the southern half. The regional climate is semiarid, with mean annual precipitation at park headquarters averaging 793 mm and estimated mean annual temperature for the entire lower basin averaging 20°C (Gamachu, 1974). There is a gradient in rainfall across the western half of the lower Omo Basin (≈800–350 mm; north to south), with rainfall typically occurring between March and June (≈65% of the total precipitation) and between October and December (≈20% of the total precipitation) (Schloeder, 1999). Grassland and shrubland savanna vegetation characterizes the Omo landscape, and the study area is dominated by tall subhumid grasses in the northern part, short semiarid grasses in the southern part, and a range of subhumid and semiarid adapted species in the central part (Jacobs, 1999).

Soil Sampling and Analyses

Surface soil samples were collected from 46 sites in the Omo Plain (Fig. 2). Sample size and choice of sample sites were a function of our effort to contend with the minimum sample size and pairwise comparison criterion, staying within the Vertisols and the grassland (termed *soil-scape criteria*) and finances. Sampling at each site involved systematically collecting five 30-cm deep cores along a 20-m transect, using a 10-cm-diam. soil push tube. Core sample size was determined

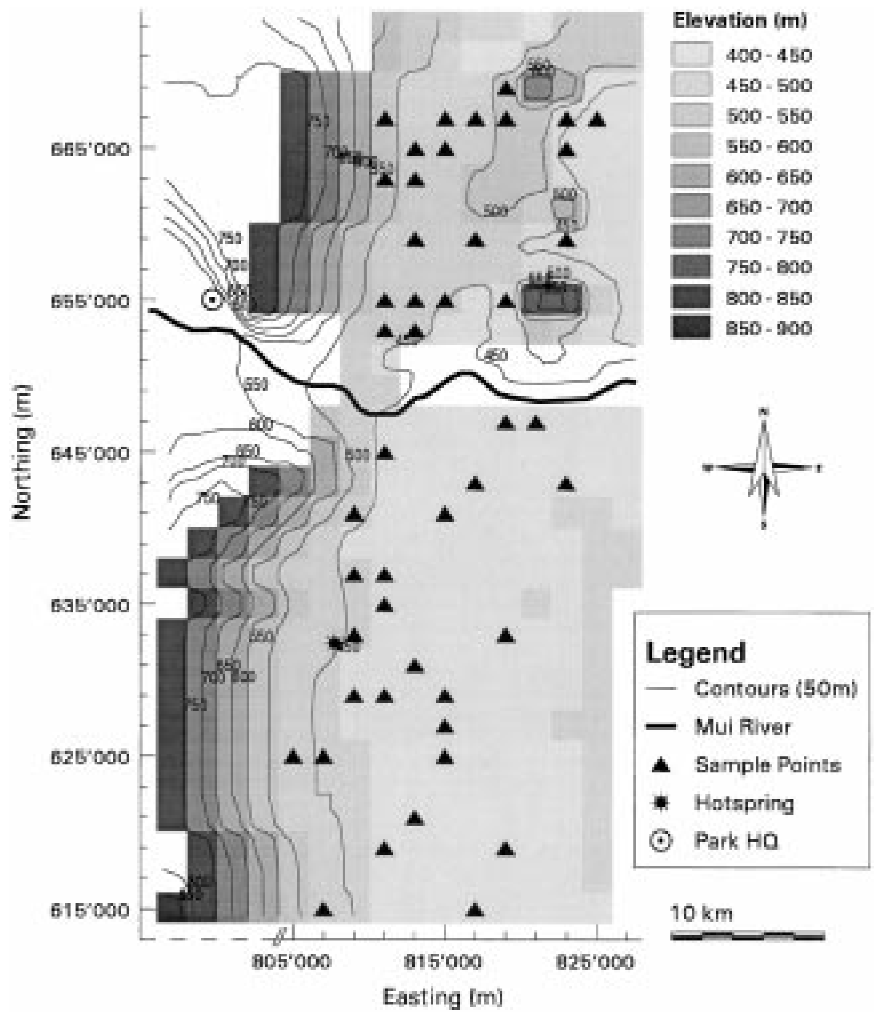

Fig. 2. Location of sample sites in the alluvial plain.

during a preliminary analysis, and core depth was based on average maximum rooting depth (Jacobs, 1999) and extent of variability reported in unpublished documents (Omo-Ghibe project, unpublished data, 1994). The five samples were bulked by site, air-dried, and passed through a 2-mm sieve. The International Livestock Research Institute soils lab in Addis Ababa, Ethiopia, performed all soil analyses following procedures outlined in Miller et al. (1998).

Surface soil properties investigated in this study included sand, silt, and clay content, pH, exchangeable Na, Ca, K, and Mg content, cation-exchange capacity (CEC), base saturation (BS), total available P, and organic matter (OM). Sand, silt, and clay content was determined by the Andreasen pipette method and by sieving, and pH was measured using the 1:1 soil/water saturation paste method. Exchangeable Na, Ca, K, and Mg content was measured using a modification of the Ammonium Acetate Method and CEC by the Ammonium Replacement Method. Base saturation was calculated as the ratio of the sum of exchangeable Na, Ca, K, and Mg content to CEC, and expressed as a percentage. Olsen's method was used to determine P. All analyses were performed in triplicate and results were averaged. Descriptive summary statistics including the mean, range, standard deviation, coefficient of variation, skewness, and kurtosis were computed in SYSTAT (SPSS, 1997), where the coefficient of variation is calculated as the standard deviation as a percentage of the mean, and skewness and kurtosis measures are centered at zero.

Approach to Spatial Interpolation

Predictions for comparison can be obtained using a variety of procedures. Because our data were limited, we used the cross validation procedure (also referred to as the jackknife procedure) (Isaaks and Srivastava, 1989). This involved consecutively removing a data value from the sample data set and interpolating to that site using the remaining data values.

Preliminary experimental analyses showed that the variograms to be used with OK were best represented by the spherical model:

$$\begin{cases} \gamma(h) = c_0 + c\left[\dfrac{3}{2}\dfrac{h}{a} - \dfrac{1}{2}\left(\dfrac{h}{a}\right)^3\right] & \text{for } 0 < h \leq a \\ \gamma(h) = c_0 + c & \text{for } h > a \end{cases}$$

where $\gamma(h)$ is the variogram for lag h, h is the distance between observations, c_o is the nugget, $c_0 + c_1$ is the sill, and a is the range. Our approach to OK involved estimating each variogram with a classical variogram estimator, fitting each variogram with a spherical variogram model, and estimating the weighting parameters by nonlinear least squares. The nonlinear least squares approach does not necessarily fair well when the points in the variogram have unequal weighting, however (Cressie, 1983). We addressed this concern by using a minimum pair criterion = 20 when estimating with the classical variogram estimator. Ordinary kriging was implemented using the S+ (SPLUS) software (Kaluzny et al., 1996) using a spherical covariance function with the estimated weighting parameters and, because our data were limited, an isotropic search radius (equal to the range of each variogram).

With ID and splines we opted to use six different approaches per method because we wanted to evaluate the level of subjectivity involved in the choice of parameter values. For ID the approaches involved using combinations of three different distance power parameter (p) values: 0.5, 1, and 2, and two different search radii (r): 22 and 12 km, when calculating the individual predictions. These power parameter values represented commonly used weighting values and the two search radii represented distances slightly less than half the study area's length and width. The larger search radius also was similar to the range used when first estimating each variogram with a classical variogram estimator. For splines, the approaches involved using combinations of three variations of the weight (w) parameter: 0.01, 0.10, and 0.50; and two variations of the number of points parameter (p): 8 and 16, when calculating the individual predictions. For comparison, the two values selected for the number of points parameter represented the average number of points that would fall within either of the search radii used with ID, and within the range of the variogram. We used the GRID-module in Arc/Info (ESRI, 1991) to calculate the ID and splines predictions.

Prediction Comparison and Statistical Analyses

Comparison of predictions was based on two measures of accuracy: the mean absolute error (MAE) and the mean-squared error (MSE) measure; and one measure of effectiveness: the goodness-of-prediction (G) estimate. The MAE is a measure of the sum of the residuals (e.g., predicted minus observed) (Voltz and Webster, 1990):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left[\left|z(x_i) - \hat{z}(x_i)\right|\right] \qquad [7]$$

where $z(x_i)$ is the observed value at location i, $\hat{z}(x_i)$ is the predicted value at location i, and n is the sample size. Small MAE values indicate a method with few errors, overall.

The MAE measure, however, does not reveal the magnitude of error that might occur at any point (Voltz and Webster, 1990; Laslett, 1994; Gotway et al., 1996). For this reason we chose to calculate the MSE measure, which is a measure of the sum of the squared residuals:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}[z(x_i) - \hat{z}(x_i)]^2 \qquad [8]$$

Squaring the difference at any point gives an indication of the magnitude of differences such that small MSE values indicate more accurate predictions, point-by-point.

The G measure gives an indication of how effective a prediction might be, relative to that which could have been derived from using the sample mean alone (Agterberg, 1984):

$$G = \left(1 - \left\{\sum_{i=1}^{n}\left[z(x_i) - \hat{z}(x_i)\right]^2 / \sum_{i=1}^{n}\left[z(x_i) - \bar{z}\right]^2\right\}\right)100 \qquad [9]$$

where $\bar{z}$ is the sample mean. A G value equal to 100% indicates perfect prediction while negative values indicate that the predictions are less reliable than if one had used the sample mean instead.

Interpolation methods were statistically analyzed using the Analysis of Variance (ANOVA) procedure, using the lowest MAE and MSE values and the highest G values. Pairwise comparison t-tests were calculated to determine whether the methods predicted the same at the same locations (Hollander and Wolfe, 1973). All analyses were performed in SYSTAT (SPSS, 1997). Significance was based on a probability of 0.05 using a Bonferroni-correction (Cooper, 1968) to take into account multiple testing, when applicable.

RESULTS AND DISCUSSION

These soils data were consistent with what was expected from a Vertisol-dominated landscape (Table 1). However, the calculated coefficient of variation was high for P content, sand content, silt content, and clay content because of a few extreme values (n = 2, 2, 1, and 3, respectively). Furthermore, the exchangeable Na data exhibited bimodality. In general, the soils in this specific soil-scape were highly alkaline, high in clay content, and low in sand content, and exhibited high base saturation (>75%).

Interpolation Success

Spatial interpolation of the sand content, silt content, exchangeable K content, CEC, and BS data proved unreliable, using the OK method. Furthermore, we always obtained negative G values for these same properties, when using the ID and splines methods (data not presented because OK was not possible). One reason for this may have been that the scale of our sampling scheme (e.g., sample distance ≥2 km) was either too small or too large. Limited data may be another reason because on occasion, we did observe some degree of spatial dependency in the variograms at distances <10 km.

Table 1. Summary statistics for all soil properties: mean, range, standard deviation (SD), percentage coefficient of variation (CV%), skewness, and kurtosis.

Property†	Mean	Range	SD	CV %	Skewness	Kurtosis
Sand, g kg^{-1}	171	18–518	98	57	1.17	2.64
Silt, g kg^{-1}	213	100–560	95	44	1.68	3.28
Clay, g kg^{-1}	631	281–841	105	71	−0.15	0.71
pH	7.82	6.87–8.49	0.38	5	−0.44	−0.41
Exchangeable bases						
Na, cmol$_c$ kg^{-1}	3.61	0.24–8.77	2.40	66	−0.05	0.50
Ca, cmol$_c$ kg^{-1}	27.56	14.36–39.11	5.49	20	−0.27	0.28
K, cmol$_c$ kg^{-1}	1.22	0.65–2.34	0.35	28	0.93	1.41
Mg, cmol$_c$ kg^{-1}	7.64	0.92–14.43	3.24	42	0.30	−0.37
CEC, cmol$_c$ kg^{-1}	43.22	26.06–62.79	8.14	19	0.13	0.41
BS, %	91.98	76.68–100.00	7.15	7	−0.54	−1.00
P, mg kg^{-1}	4.68	0.01–21.58	5.65	121	1.74	2.80
OM, g kg^{-1}	14	8–26	4	27	0.63	0.21

† CEC, cation-exchange capacity; BS, base saturation; P, total available phosphorus; OM, organic matter.

The pattern was erratic, however. We would need more samples to determine whether the pattern was an artifact of the data or evidence of spatial dependency, and whether our sample design was inappropriate. What is also interesting to note as a result of this exercise is that we would not have been able to determine success if we had based our decision solely on either the MAE or MSE measures. It was only because we had calculated the G measure that we were able to determine that interpolation with these properties, using any of these methods, was inappropriate.

Example variograms and the model fit obtained by nonlinear least squares for those soil properties that exhibited spatial dependency are illustrated in Fig. 3. The variogram for clay content and interpolation of this data were based on cross validation using 44 rather than 46 samples, however. This was because the inclusion of two of the extreme clay content values (the lowest ones) resulted in weak to no spatial dependency, whereas their exclusion resulted in spatial dependency.

There were only three soil properties for which we obtained reasonably accurate and effective predictions (Tables 2–9). These were pH, exchangeable Na content, and OM content. Analyses of the OK variance statistics and variograms indicated the possible reasons for this. For all soil properties except these and P content, interpolation was based on models with either very high nuggets or very high sill-to-nugget ratios (>30%). The effect of high nugget values was a more equal distribution of weights that resulted in higher predictions, while high sill-to-nugget ratios resulted in an increase in the prediction variances. Both were a consequence of limited, variable, or weakly autocorrelated data, particularly with respect to clay content. These same characteristics also contributed to similarly high MAE and MSE values, and similarly low G values, using ID and splines.

A high nugget or high sill/nugget ratio could also have resulted from using a spherical model as the pattern of spatial continuity. The effect of an inappropriate choice of pattern of spatial continuity would have been specific to only a few instances when a point was dropped during cross validation, however, rather than to all 46 instances.

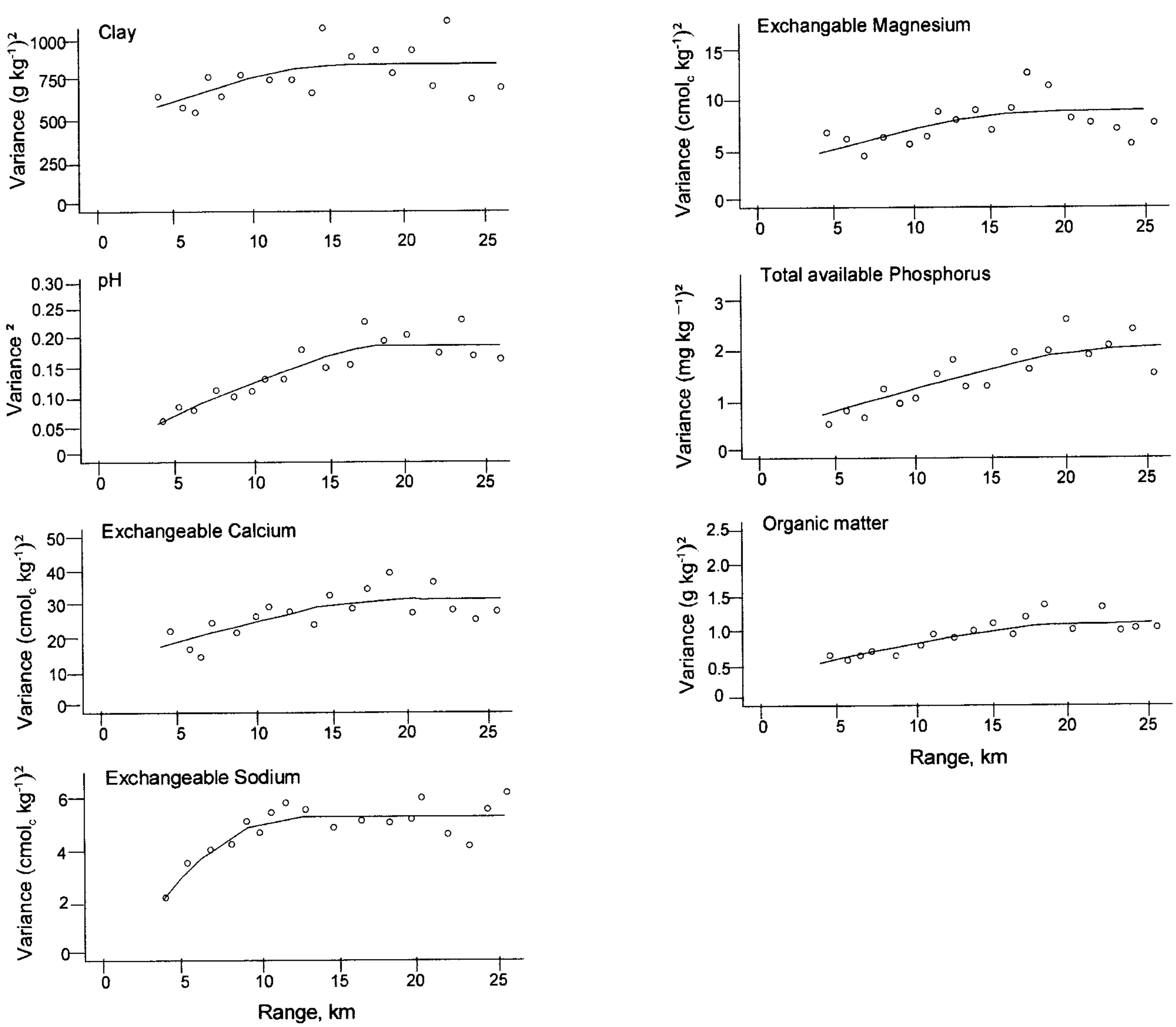

Fig. 3. Example omnidirectional variograms from cross-validation exercises.

Table 2. Range in variance statistics (range, sill, and nugget) and variances (vp) of predictions.

Property	Range (km)	Sill†	Nugget†	vp†
Clay, g kg^{-1}	11–17	618–996	357–580	40–112
pH	17–21	0.16–0.21	0.01–0.02	0.18–0.38
Exchangeable bases				
Na, cmol$_c$ kg^{-1}	13–24	4.99–9.12	0.21–0.14	0.26–2.43
Ca, cmol$_c$ kg^{-1}	18–21	17.48–37.71	7.88–14.33	1.56–5.71
Mg, cmol$_c$ kg^{-1}	16–22	3.36–7.13	3.66–6.14	0.89–3.25
P, mg kg^{-1}‡	24–38	1.75–2.61	0.19–0.51	0.48–1.22
OM, g kg^{-1}§	20–47	0.59–1.10	0.35–0.55	0.20–0.30

† Values indicate squared units.
‡ P, total available phosphorus.
§ OM, organic matter.

Alternatively, either could have resulted because estimation by nonlinear least squares was based on unequally represented variogram values (range in data pairs: 20 to 34; mean = 24), despite our effort to minimize differences. This would not explain why we obtained accurate predictions (with low variances) for pH, exchangeable Na content, and OM, however.

Data clustering is another possible explanation, at least for some of the soil properties. Clustering causes the variogram to be more representative of a particular region than of the entire study area, particularly when the data may actually represent two separate soil populations (Isaaks and Srivastava, 1989). However, in another investigation (Schloeder, 1999) it was determined that the only differences that existed between the northern and southern halves of the study area were that there was a decreasing trend in OM (north to south), and the ratio of exchangeable Na content to clay content was twice as high in the south as it was in the north.

Data clustering can also lead to higher variances for predictions in areas where the sample points are more widely spaced. In this investigation the sample points were in general farther apart in the southern half of the study area than they were in the northern half, partly because we limited our sampling to the Vertisol-dominated grassland areas only (the northern half of the grassland plain was smaller than the southern half). An analysis of the variances of the predictions from both areas revealed that they tended to be lower in the north than in the south for all properties except pH, exchangeable Na content, and OM. These results suggest then that our data do not come from separate soil populations and that data clustering may in fact be the problem, at least for most of the data sets.

Extreme values were certainly important, particularly with respect to clay content, because the inclusion of two low values prevented interpolation with OK. We considered this same possibility for the P content data and performed a separate cross validation exercise, whereby we removed the three highest values from the P data set ($n = 43$) and recalculated the MAE, MSE, and G measures. Following, we determined that accuracy increased and effectiveness declined (MAE = 2.4 mg kg^{-1}; MSE = 10.4 [mg kg^{-1}]^2; G = 15.3%; using ID: p = 0.50, r = 16), and there was little change in the average nugget and sill-to-nugget ratios (nugget = 0.14 and sill = 1.61 [mg kg^{-1}]^2). Unlike the results we obtained by removing the two low clay values, we found that inclusion of the three highest P content values resulted in an increase, rather than a decrease, in the amount of spatial dependency in the P data. These results also explained why the G result was less than satisfactory for P content when we used all 46 values, in that the inclusion of all 46 values meant that a higher mean value was used with the G measure. They also indicate that either response is possible if an investigator were to remove an extreme value from their data. Caution should be exercised when considering their removal, as a result, particularly when data are limited.

A pattern of smooth and nonerratic behavior, either globally or locally (e.g., within the range of the search neighborhoods) explains successful prediction of the pH, exchangeable Na, and OM data. In another investi-

Table 3. Mean absolute error (MAE) and mean square error (MSE), standard errors (SE), and goodness-of-prediction values (G %), using the ordinary kriging method.

Property	MAE	SE	MSE†	SE†	G %
Clay, g kg^{-1}	83	9	1042	208	2.79
pH	0.21	0.03	0.08	0.02	44.27
Exchangeable bases					
Na, cmol$_c$ kg^{-1}	0.98	0.11	1.50	0.31	73.33
Ca, cmol$_c$ kg^{-1}	4.44	0.54	32.68	7.63	12.22
Mg, cmol$_c$ kg^{-1}	2.37	0.28	9.20	1.92	10.25
P, mg kg^{-1}‡	3.12	1.07	22.86	7.18	26.88
OM, g kg^{-1}§	2.20	0.30	0.80	0.20	46.61

† Values indicate squared units.
‡ P, total available phosphorus.
§ OM, organic matter.

Table 4. Mean absolute error values and standard errors (SE) by approach, using the inverse-distance weighting method.

	Approach†					
Property	1	2	3	4	5	6
Clay, g kg^{-1}	81.80	80.10	79.50	79.60	78.90	80.40
SE	8.30	8.20	8.30	8.30	8.30	8.30
pH	0.26	0.25	0.22	0.25	0.24	0.21
SE	0.03	0.03	0.03	0.03	0.03	0.03
Exchangeable bases						
Na, cmol$_c$ kg^{-1}	1.53	1.43	1.29	1.41	1.33	1.21
SE	0.14	0.15	0.16	0.14	0.14	0.15
Ca, cmol$_c$ kg^{-1}	4.26	4.25	4.39	4.38	4.36	4.48
SE	0.36	0.41	0.44	0.37	0.40	0.42
Mg, cmol$_c$ kg^{-1}	2.16	2.14	2.10	2.19	2.16	2.11
SE	0.26	0.26	0.27	0.26	0.26	0.27
P, mg kg^{-1}‡	3.49	3.43	3.61	3.44	3.45	3.63
SE	0.48	0.52	0.54	0.48	0.50	0.51
OM, g kg^{-1}§	2.10	2.20	2.20	2.20	2.20	2.20
SE	0.30	0.30	0.30	0.30	0.30	0.30

† Model parameters: Approach 1–3, r (search radii) = 22; Approach 4–6, r = 12; Approach 1 and 4, p (distance power) = 0.5; Approach 2 and 5, p = 1; Approach 3 and 6, p = 2.
‡ P, total available phosphorus.
§ OM, organic matter.

Table 5. Mean square error and standard errors (SE) by approach, using the inverse-distance weighting method. Values indicate squared units.

	Approach†					
Property	1	2	3	4	5	6
Clay, g kg^{-1}	984	943	937	935	924	949
SE	177	160	158	188	174	170
pH	0.11	0.10	0.09	0.10	0.09	0.08
SE	0.02	0.02	0.02	0.02	0.02	0.02
Exchangeable bases						
Na, cmol$_c$ kg^{-1}	3.54	3.16	2.62	2.92	2.67	2.34
SE	0.53	0.61	0.66	0.48	0.54	0.59
Ca, cmol$_c$ kg^{-1}	30.06	29.55	30.80	31.51	31.14	32.25
SE	3.60	4.16	4.79	3.65	4.00	4.39
Mg, cmol$_c$ kg^{-1}	7.82	7.55	7.41	7.73	7.50	7.39
SE	1.80	1.70	1.74	1.82	1.70	1.71
P, mg kg^{-1}‡	25.67	24.08	23.44	24.11	23.47	23.62
SE	5.72	7.09	7.95	5.61	6.59	7.20
OM, g kg^{-1}§	0.80	0.80	0.80	0.80	0.80	0.80
SE	0.20	0.20	0.20	0.20	0.20	0.20

† Model parameters: Approach 1–3, r (search radii) = 22; Approach 4–6, r = 12; Approach 1 and 4, p (distance power) = 0.5; Approach 2 and 5, p = 1; Approach 3 and 6, p = 2.
‡ P, total available phosphorus.
§ OM, organic matter.

Table 7. Mean absolute error values and standard errors (SE) by approach, using the thin-plate smoothing splines with tension method.

	Approach†					
Property	1	2	3	4	5	6
Clay, g kg^{-1}	84.70	85.40	86.40	82.70	83.30	85.00
SE	8.30	8.29	8.30	8.30	8.30	8.30
pH	0.25	0.25	0.25	0.24	0.24	0.25
SE	0.03	0.03	0.03	0.03	0.03	0.03
Exchangeable bases						
Na, cmol$_c$ kg^{-1}	1.07	1.07	1.07	1.06	1.06	1.06
SE	0.10	0.10	0.10	0.10	0.10	0.10
Ca, cmol$_c$ kg^{-1}	5.90	5.97	6.06	5.76	5.88	6.06
SE	0.36	0.41	0.44	0.37	0.40	0.42
Mg, cmol$_c$ kg^{-1}	2.67	2.68	2.69	2.70	2.73	2.80
SE	0.35	0.34	0.34	0.34	0.33	0.33
P, mg kg^{-1}‡	4.14	4.21	4.30	3.90	4.01	4.17
SE	0.56	0.55	0.55	0.55	0.55	0.54
OM, g kg^{-1}§	2.50	2.50	2.50	2.50	2.50	2.50
SE	0.30	0.30	0.30	0.30	0.30	0.30

† Model parameters: Approach 1–3, p (number of sample points) = 16; Approach 4–6, p = 8; Approach 1 and 4, w (weight) = 0.01; Approach 2 and 5, w = 0.10; Approach 3 and 6, w = 0.50.
‡ P, total available phosphorus.
§ OM, organic matter.

gation, pH and exchangeable Na content were determined to be positively correlated, and their spatial patterns related to the seasonal spring in the north, the hotspring in the south, and the gradient in rainfall (Schloeder, 1999). The nature of the spatial pattern displayed by OM was a south-trending decrease in content, in response to a change in species composition, vegetation height, and underground biomass, and the gradient in rainfall (Jacobs, 1999; Schloeder, 1999). This decreasing trend was reflected in the higher nugget estimates. Unfortunately, using an anisotropic search radius would have led to fewer pairwise comparisons and worse, rather than better, predictions.

Accuracy and Effectiveness

Results from the ANOVA tests indicated that the three methods did not differ in accuracy, and all approaches were similar, regardless of which type of residual measure or soil property we examined. The pairwise comparison t-test results, however, indicated that splines differed from ID by frequently producing residuals that were of greater magnitude at specific locations, for exchangeable Mg content (MAE: P = 0.03; MSE: P = 0.02) and exchangeable Ca content (MSE: P = 0.04). Splines tended to yield higher residuals because this method, which is based on a polynomial function, uses a smoothing approach when interpolating, which results in a loss of local detail at specific points. Ordinary kriging and ID, on the other hand, calculate an estimated value that is based on a direct summation of the data values within a specified distance or radii of the predicted point. This direct summation approach tends to preserve both local detail and trend. More localized detail in the exchangeable Ca and exchangeable Mg data were the cause of less accurate site-specific predictions using splines, and no difference in predictions when comparing ID with OK.

An analysis of the G values indicated that splines was also for the most part not as effective an interpolator as OK and ID because of its smoothing approach to interpolation. Furthermore, splines was less effective

Table 6. Percentage goodness-of-prediction values by approach, using the inverse-distance weighting method.

	Approach†					
Property	1	2	3	4	5	6
Clay, g kg^{-1}	8.14	11.96	12.59	12.79	13.78	11.42
pH	23.94	31.15	40.36	28.21	35.24	43.02
Exchangeable bases						
Na, cmol$_c$ kg^{-1}	36.92	43.65	53.32	47.92	52.47	58.31
Ca, cmol$_c$ kg^{-1}	7.51	9.09	5.22	3.04	4.18	0.78
Mg, cmol$_c$ kg^{-1}	23.76	26.36	27.76	24.67	26.90	27.94
P, mg kg^{-1}‡	17.90	22.97	25.04	22.90	24.93	24.45
OM, g kg^{-1}§	49.17	49.37	48.99	44.04	45.63	47.11

† Model parameters: Approach 1–3, r (search radii) = 22; Approach 4–6, r = 12; Approach 1 and 4, p (distance power) = 0.5; Approach 2 and 5, p = 1; Approach 3 and 6, p = 2.
‡ P, total available phosphorus.
§ OM, organic matter.

Table 8. Mean square error and standard errors (SE) by approach, using the thin-plate smoothing splines with tension method. Values indicate squared units.

	Approach†					
Property	1	2	3	4	5	6
Clay, g kg^{-1}	1050	1067	1091	986	1016	1066
SE	177	160	158	188	174	170
pH	0.10	0.11	0.11	0.10	0.11	0.11
SE	0.03	0.03	0.03	0.03	0.03	0.03
Exchangeable bases						
Na, cmol$_c$ kg^{-1}	1.65	1.66	1.64	1.64	1.61	1.60
SE	0.24	0.24	0.24	0.23	0.23	0.23
Ca, cmol$_c$ kg^{-1}	58.26	60.10	62.72	55.59	58.43	62.83
SE	3.60	4.16	4.79	3.65	4.00	4.39
Mg, cmol$_c$ kg^{-1}	12.69	12.82	12.95	12.43	12.78	13.25
SE	3.06	3.03	3.03	2.97	2.95	2.94
P, mg kg^{-1}‡	30.11	30.98	32.20	27.73	28.94	30.94
SE	6.97	6.80	6.71	6.51	6.37	6.22
OM, g kg^{-1}§	1.00	1.00	1.00	1.00	1.00	1.00
SE	0.20	0.20	0.20	0.20	0.20	0.20

† Model parameters: Approach 1–3, p (number of sample points) = 16; Approach 4–6, p = 8; Approach 1 and 4, w (weight) = 0.01; Approach 2 and 5, w = 0.10; Approach 3 and 6, w = 0.50.
‡ P, total available phosphorus.
§ OM, organic matter.

Table 9. Percentage goodness-of-prediction values by approach, using the thin-plate smoothing splines with tension method.

	Approach†					
Property	1	2	3	4	5	6
Clay, g kg^{-1}	1.99	0.41	−1.80	8.02	5.21	0.51
pH	23.94	31.15	40.36	18.12	35.24	43.02
Exchangeable bases						
Na, cmol$_c$ kg^{-1}	70.54	70.68	70.81	70.87	71.27	71.56
Ca, cmol$_c$ kg^{-1}	−79.24	−84.91	−92.99	−71.04	−79.79	−93.31
Mg, cmol$_c$ kg^{-1}	−23.73	−24.98	−26.25	−21.23	−24.59	−29.21
P, mg kg^{-1}‡	3.71	0.90	−3.00	11.30	7.45	1.03
Om, g kg^{-1}§	39.30	39.02	37.95	37.07	36.10	35.00

† Model parameters: Approach 1–3, p (number of sample points) = 16; Approach 4–6, p = 8; Approach 1 and 4, w (weight) = 0.01; Approach 2 and 5, w = 0.10; Approach 3 and 6, w = 0.50.
‡ P, total available phosphorus.
§ OM, organic matter.

because this method calculated predictions of greater magnitude of error for sites located along the edge of the study area. Ordinary kriging and ID, however, calculated edge-site predictions that were more similar to the actual data. Splines, OK, and ID differed in their calculation of edge-site predictions because in the absence of external data points for reference, predictions along the edge followed the specific pattern of the model being used. For OK, this meant that each edge prediction was a product of our effort to predict what the pattern in spatial continuity was from the variogram, and solved for using a spherical covariance function, with the criterion of minimizing the error variance. Consequently, it was the model of the pattern of spatial continuity that determined the influence of the nearby sample points when deriving a prediction for each of the edge sites. Alternatively, ID conserved the nearest values such that when the nearby values were high, ID calculated a prediction that was similarly high. Following, when the nearest values were low, ID calculated a prediction that was similarly low. Splines, however, followed the model trend such that when the nearby values were on an increasing trajectory, splines calculated a prediction that was higher than the nearby values. Alternatively, when the nearby values were on a decreasing trajectory, splines calculated a prediction that was lower than the nearby values. It should not be surprising then that splines proved less effective than OK or ID, since 37% of our sample sites occurred along the edge of the study area and there was detail or trend in these data.

Accurate and effective prediction of the pH, exchangeable Na, and OM data was achieved because of the nature of the spatial dependency exhibited by these data within the study area, and within the selected range(s) of estimation and weighting. Additionally, similar results using OK and ID indicated these methods to be equally accurate and effective predictors of the pattern of dependency, and that splines was a poor predictor. Ordinary kriging was preferable to ID, however, because it was the only method where we could obtain an estimate of the variance for every prediction and where some of the subjectivity was removed when selecting the weighting values (Burgess and Webster, 1980; Isaaks and Srivastava, 1989). This method is also useful in that it does not tend to produce a bullseye pattern around the sample locations (Gotway et al., 1996).

CONCLUSIONS

The results of this investigation demonstrated that spatial interpolation of coarse-scaled limited surface soils data was mostly inappropriate. For most of the data sets our inability to predict, or the ability to predict without much accuracy, could be attributed to either spatially independent data, limited data, sample spacing, extreme values, and variable or erratic behavior. There were three data sets, however, where spatial interpolation was not inappropriate. These were pH, exchangeable Na content, and OM. Our ability to predict successfully with these data raises the question as to the exact nature of the relationship between accuracy, sample size, and sample spacing, and to what extent these factors are related to the soil property under investigation, particularly when data are limited. It also leads us to speculate as to which data characteristics are important and when they might be important. To illustrate, the exchangeable Na data used in this investigation were highly variable and exhibited bimodality. However, these characteristics did not appear to negatively affect the predictions because the characteristics represented statistical evidence of spatial dependency in the data rather than a problem with the data. Alternatively, the extreme characteristics exhibited by the clay data represented a random pattern of data distribution as well as a problem with the data. To date, there have been few coarse-scaled, data-intensive soil investigations. This leaves us with little understanding of the exact pattern and nature of spatial dependency that is inherent in many soil resources, and few statistical means with which to spatially characterize a resource. The questions raised here, as well as the need for spatial characterization, suggest the need for more coarse-scaled data-intensive soil investigations.

ACKNOWLEDGMENTS

This research was funded by the Wildlife Conservation Society and International Livestock Research Institute under a cooperative agreement with the Ethiopian Wildlife Conservation Organization. Technical support was provided by the Utah State Univ.–USGS– Biol. Res. Div. Coop. Fisheries and Wildlife Research Unit and the Utah State Univ. Ecology Center Program. N.E. Zimmerman was supported by a grant from the Swiss Federal Research Institute–WSL (Project no. 4.92.771) and of the Novartis Foundation. We thank J. Kern for his technical advice and C.A. Gotway for reviewing the manuscript.

REFERENCES

Agterberg, F.P. 1984. Trend surface analysis. p. 147–171. *In* G.L. Gaile and C.J. Willmott (ed.) Spatial statistics and models. Reidel, Dordrecht, the Netherlands.

Brown, F.H., and W.P. Nash. 1976. Radiometric dating and tuff mineralogy of Omo Group deposits. p. 50–623. *In* W.W. Bishop (ed.) Geological background to fossil man. Scottish Acad. Press, Edinburgh.

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

Butzer, K. 1971. Recent history of an Ethiopian Delta: The Omo River and the level of Lake Rudolf. Dep. Geog. Res. Pap. 136. Univ. Chicago, Chicago, IL.

Butzer, K., and D. Thurber. 1969a. Some late Cenozoic sedimentary formations of the lower Omo Basin. Nature 222:1138–1143.

Butzer, K., and D. Thurber. 1969b. Horizontal sediments of the lower Omo valley: The Kibbish Formation. Quaternaria 11:15–29.

Cliff, A.D., and J.K. Ord. 1981. Spatial processes: Models and applications. Pion Limited, London, UK.

Cooke, R.A., S. Mostaghimi, and J.B. Campbell. 1993. Assessment of methods for interpolating steady-state infiltrability. ASAE Publ. 36:1333–1341.

Cooper, D.W. 1968. The significance level in multiple tests made simultaneously. Heredity 23:614–617.

Cressie, N.A. 1993. Statistics for spatial data. John Wiley & Sons, Ontario, Canada.

Davidson, A. 1983. The Omo River Project: Reconnaissance geology and geochemistry of parts of Illubabor, Kefa, Gemu Gofa and Sidamo, Ethiopia. Min. Mines & Ener., Ethiopia Inst. Geol. Surv. Addis Ababa, Ethiopia.

Diaz, O.A., D.I. Anderson, and E.A. Hanion. 1992. Soil nutrient variability and soil sampling in the everglades agricultural area. Comm. Soil Sci. Plant Anal. 23:2313–2337.

Englund, E., D. Weber, and N. Leviant. 1992. The effects of sampling design parameters on block selection. Math. Geol. 24:329–343.

Environmental Systems Research Institute. 1991. Grid command reference. ESRI, Badlands, CA.

Food and Agricultural Organization. 1965. Survey of the Awash River Basin. Vol. 11. FAO, Rome, Italy.

Gamachu, D. 1974. Aspects of climate and water budget in Ethiopia. Tech. Mono. Addis Ababa Univ., Ethiopia.

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. Am. J. 60:1237–1247.

Hollander, M., and D.A. Wolfe. 1973. Nonparametric Statistical Methods. Wiley & Sons, New York.

Hosseini, E., J. Gallichand, and D. Marcotte. 1994. Theoretical and experimental performance of spatial interpolation methods for soil salinity analysis. Trans. ASAE 37:1799–1807.

Isaaks, E.H., and R.M. Srivastava. 1989. An introduction to applied geostatistics. Oxford Univ. Press, New York.

Jacobs, M.J. 1999. Influence of grazing, fire, and rainfall regime on plant species dynamics in an Ethiopian perennial grassland. Ph.D. diss. Utah State Univ., Logan, UT.

Journel, A.G., and C.H. Huijbregts. 1978. Mining geostatistics. Academic Press, New York.

Kaluzny, S.P., S.C. Vega, T.P. Cardoso, and A.A. Shelly. 1996. S+ Spatialstats users manual. Mathsoft, Seattle, WA.

Laslett, G.M. 1994. Kriging and splines: An empirical comparison of their predictive performance in some applications. J. Am. Stat. Assoc. 89:391–409.

Miller, R.O., J. Kotuby-Amacher, and J.B. Rodriguez. 1998. Western states laboratory proficiency testing program: Soil and plant analytical methods. Ver. 4.10. Utah State Univ., Logan, UT.

Mitas, L., and H. Mitasova. 1988. General variational approach to the interpolation problem. Comp. Math. Appl. 16:983–992.

Schloeder, C.A. 1999. Determinants of African savanna vegetation distribution in an African savanna landscape ecotone. p. 94–151. *In* C.A. Schloeder. Investigation of the determinants of African savanna vegetation distribution: A case study from the lower Omo Basin, Ethiopia, Chapter 4. Ph.D. diss. Utah State Univ., Logan.

Sutcliffe, J.P. 1992. Peoples and natural resources in the north and south Omo and Kefa administrative regions of southwestern Ethiopia: A case study in strategic natural resource planning. Min. Plan. Econ. Dev., Addis Ababa, Ethiopia.

SPSS. 1997. SYSTAT 6.0 for Windows: Statistics and Graphics. SPSS, Inc. Chicago, IL.

Trangmar, B.B. 1984. Spatial variability of soil properties in Sitiung, West Sumatra, Indonesia. Ph.d. Diss. Univ. of Hawaii, Honolulu.

Uehara, G., B.B. Trangmar, and R.S. Yost. 1985. Spatial variability of soil properties. p. 61–95. *In* D.R. Nielsen and J. Bouma (ed.) Soil Spatial Variability. Proc. ISSS and SSSA, Las Vegas, NV. 30 Nov.–1 Dec. 1984. PUDOC, Wageningen, the Netherlands.

Voltz, M., and R. Webster. 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. J. Soil Sci. 41:473–490.

Watson, D.F. 1992. Contouring: A guide to the analysis and display of spatial data. Pergammon Press, New York.

Watson, D.F., and G.M. Philip. 1985. A refinement of inverse distance weighted interpolation. Geo-process. 2:315–327.

Weber, D., and E. Englund. 1992. Evaluation and comparison of spatial interpolators. Math. Geol. 24:381–391.

Weisz, R., S. Fleischer, and Z. Smilowitz. 1995. Map generation in high-value horticultural integrated pest management: Appropriate interpolation methods for site-specific pest management of Colorado Potato Beetle (Coleoptera: Chrysomelidae). J. Econ. Entomol. 88:1650–1657.

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.