USE OF GEOSTATISTICAL TECHNIQUES IN THE DESIGN OF AN AGRICULTURAL FIELD EXPERIMENT

Wayne P. Dulaney, University of Maryland; Laura L. Lengnick and Galen F. Hart, U.S. Department of Agriculture
Wayne P. Dulaney, USDA/ARS/NRI/Remote Sensing Research Laboratory, BARC-West, Bldg. 007, Rm. 116
10300 Baltimore Ave., Beltsville, MD 20705

Key Words: Geostatistics, Kriging, Experimental design

SUMMARY

The usefulness of certain geostatistical techniques in the design of an agricultural field trial was examined. Of immediate concern was the use of these techniques to help identify an optimal sampling scheme for characterization of soil spatial variability. This was achieved by determining, *a priori*, the prediction error associated with various sampling designs. Our ultimate goal is to use geostatistical techniques to help minimize the confounding influence of location effects on treatment differences. Knowledge of soil spatial variability will be used to help locate experimental blocks on areas that are as uniform as possible, thus increasing the precision with which treatment differences can be detected. Kriging will be used to produce interpolated contour maps of various soil properties at the experimental field site.

In June 1993, a soil survey was conducted at the proposed field site. Data on various soil characteristics were collected on a 50 m x 75 m grid. Geostatistical analyses were performed on these data using public domain geostatistical software packages (GEO-EAS and GEOPACK) developed at the U.S. Environmental Protection Agency. Semivariograms were produced for each soil characteristic and were modeled using a non-linear least-squares fitting procedure (Marquardt, 1963). Soil semi-variogram model parameters were input to the Optimal Sampling Scheme For Isarithmin Mapping (OSSFIM) program which was used to determine the prediction error associated with various grid sample spacings (McBratney and Webster, 1981). Output from OSSFIM allowed us to select a soil sampling scheme--i.e., of known accuracy and cost--for a more intensive characterization of the field site.

INTRODUCTION

In April 1993, the Beltsville Agricultural Research Center (BARC) at Beltsville, MD initiated a long-term field trial in sustainable agriculture in order to better understand the environmental and economic ramifications that these crop production systems would have in the Middle Atlantic States. In contrast to conventional farming practices which use commercially available, petroleum-based inputs to maintain soil fertility and control weed and insect pests; sustainable systems make use of "on-farm" inputs such as animal and green manures, periodic crop rotations, various tillage practices, and natural biological cycles to accomplish identical ends. Because sustainable agricultural systems are inherently more self reliant, and tend to be less exploitive of the natural resource base than more conventional production systems, the long-term viability of these systems is enhanced.

A wide range of sustainable agricultural production practices will be examined at BARC using a multidisciplinary, systems approach. The inherent complexity of such an approach makes design and layout of the field plots challenging. For example, the experimental design must allow for the analysis of numerous soil and crop characteristics collected at different spatial and temporal scales. Moreover, some experiments are expected to show subtle treatment differences, while others may require a number of growing seasons before any treatment differences are expressed. In order to properly evaluate these treatment effects, it is imperative to establish a high level of experimental precision.

Achieving such a degree of experimental precision is made more difficult because large plot sizes (0.75 ha to 1.00 ha) must be used at the field site (16.2 ha) due to crop management considerations. If plots of this size were located at the field site without regard to soil spatial variability, location effects would likely confound or obscure treatment comparisons. However, if plots were located on areas as homogeneous as possible, the power to detect experimental treatment differences would be increased, and the effectiveness of blocking and replication would be improved.

In June 1993, a survey of the experimental field site was conducted. The following soil physical and chemical properties were determined from bulked samples: depth of the A horizon (A), available phosphorus (P), available potassium (K), available calcium (Ca), available magnesium (Mg), cation exchange capacity (CEC), and pH. The sample grid was approximately rectangular, consisting of 5 transects, each 75 m apart, with samples being acquired roughly every 50 m for a total of 47 soil samples. The U.S. Department of Defense's Global Positioning System (GPS), a satellite-based navigation and positioning system, was used to geographically locate the soil sample sites to sub-meter accuracies. In addition, color infrared (CIR) aerial photographs were acquired during the growing season in order to estimate relative vegetation amount at a number of significant crop development stages. Georectification of the CIR imagery was accomplished using ground control points established with the GPS.

A geographic information system (GIS), which is explicitly designed to handle spatially referenced data, will be used to analyze the soil and vegetation data in order to determine what areas of the field site are the most uniform. Uniformity, however, must be defined in terms of crop yield because most BARC researchers will use this variable to assess treatment differences. Although most of the GIS data layers will be derived from soil physical and chemical properties, these data can be expressed in terms of their potential effect on crop yield. Once each soil data layer is translated into its respective crop yield data layer, it will be possible to produce an integrated or summary yield map which can be used to identify suitable locations for the experimental field plots.

Each soil parameter was separated into four classes representing crop yield potentials that ranged from low to high. Most of these classes were established *a priori* according to standard agronomic recommendations; the remainder were established empirically according to the range of values found at the field site. For example, depth of the A horizon often influences crop yield because the soil in this layer has the greatest capacity to provide water and nutrients. In general, the deeper the A horizon; the more productive the soil. At the experimental field site, A ranges in depth from 16 cm to 34 cm. Four classes with means of 19 cm, 23 cm, 28 cm, and 33 cm were developed in order to describe the variation in A across the site.

It is imperative to ensure that soil sampling schemes be designed such that data are collected at a level of precision such that class means, like those determined for A horizon depth, can be separated statistically. In a manner analogous to the *t* test, class means were considered to be statistically separable if they were ±2 standard errors apart. The OSSFIM program was used to produce kriging variance vs. grid spacing plots, from which it was possible to determined what grid spacing would be necessary to achieve a desired level of error. For example, to ensure that the A horizon class means could be separated, soil data must be collected at a standard error that is less than or equal to 1.

BACKGROUND

Agricultural researchers have long understood that location effects, which are often caused by natural soil variability or previous land-use practices, can significantly reduce the ability to detect experimental treatment differences. During the 1930s, attempts were made to understand the relationship between plot size or shape and experimental error (Christidis, 1931; Smith, 1938). Various techniques such as smaller plot sizes, different blocking designs, and randomization of treatments were developed to ensure that field observations were independent--i.e., treatment comparisons could be made with equal precision, even though a spatially autocorrelated variance structure might be present at the site. Recent work has shown, however, that these traditional techniques are not as robust as might be desired (Lin et al., 1993; van Es and van Es, 1993). This has prompted the development of new statistical techniques and alternative experimental designs in order to better account for the effect of field variability on experimental results (van Es et al., 1989; Nelson and Buol, 1990; Perry et al., 1993). Little work has been done, however, to evaluate whether geostatistics can be used to facilitate the design of agricultural field experiments (Perry et al., 1993). Geostatistical techniques clearly have the potential to: (1), provide better field characterization than that obtainable from more conventional spatial interpolation methods; (2), improve plot layout, thereby increasing the level of experimental precision; and (3), increase the power of post-hoc statistical techniques which are used to correct spatial autocorrelation problems.

The term geostatistics refers to a group of spatial interpolation techniques developed from theoretical work on regionalized variables by Matheron (1963) and his colleagues. Although originally intended for use in the mining industry, these techniques are increasingly being used in a wide variety of applications. Kriging, a form of weighted local averaging, is an attractive geostatistical method because, unlike all other spatial interpolation techniques, it provides an unbiased estimate with minimum and known variance (Lam, 1983).

The kriging procedure uses weights that are determined from a semi-variogram--i.e., a function that describes how a given biophysical property varies over the landscape. The semi-variogram shows how the variance of a property changes as a function of both the distance and direction between any two points. By calculating semi-variances over a range of distances or lags, an experimental semi-variogram can be produced. The semi-variance for a particular lag, h, is given by:

$$\hat{\gamma}(h) = \frac{1}{2M(h)} \sum_{i=1}^{M(h)} \{ z(x_i) - z(x_i + h) \}^2$$

where $M(h)$ is the number of paired comparisons at that lag, and $z(x_i)$ and $z(x_i + h)$ are the observed values at points x_i and $x_i + h$, respectively (Oliver and Webster, 1990).

The experimental semi-variogram is usually modeled by fitting one of a group of authorized models to it (e.g., linear, spherical, or exponential). Three terms are generally used to describe specific features of the semi-variogram model: the *nugget*, or y-intercept, provides an estimate of the amount of "white noise" present--i.e., error resulting from measurement errors and from spatial variation that occurs over distances much shorter than the shortest lag interval; the *range* marks the limit of spatial dependency; and the *sill* is the upper bound where the semi-variance reaches its maximum value (over the sill, observations are spatially independent).

The kriging variance or prediction error of a block estimate is given by:

$$\sigma^2(B) = \sum_{i=1}^{n} \lambda_i \, \bar{\gamma}(x_i, B) + \psi - \bar{\gamma}(B, B)$$

where λ_i is the coefficient or kriging weight, (x_i, B) is the average semi-variance between all points within the block and the ith sampling point, ψ is a Lagrange multiplier to minimize the kriging variance, and (B, B) is the average within-block variance. Kriging variance depends on the form of the semi-variogram, the size of the block to be estimated, and the configuration of the sampling points with respect to the block (Burrough, 1991). More importantly, prediction error does not depend directly on the actual values of the observation points. This is highly significant because if the spatial variability of a parameter can be described in the form of a semi-variogram, and what block size to estimate can be determined, then the kriging variance associated with various sampling strategies can be calculated. This makes it possible for a researcher to choose an optimal sampling strategy in advance--i.e., a plot of sampling frequency versus kriging variance or prediction error can be produced (Burgess and Webster, 1980; Burgess et al., 1981; Oliver and Webster, 1991). This technique was used to select an optimal sampling strategy for characterization of soil spatial variability at the experimental field site.

METHOD

Before semi-variograms of the soil parameters were produced, various descriptive statistics (e.g., histogram and probability plots; skewness and kurtosis) were calculated to check for outliers and to determine the distribution of each data set. The data were transformed to a normal distribution if their statistics indicated that the distribution was highly skewed.

The GEO-EAS (version 1.2.1) geostatistical software package (Englund and Sparks, 1988) was used to estimate the soil semi-variograms. An important prerequisite to production of the nondirectional (360 degree) semi-variograms was determination of an appropriate lag interval. A number of "rules-of-thumb" were followed in order to arrive at this decision: (1), the lag interval should result in 6-10 points on the semi-variogram; (2), the lag interval should result in at least 20 paired comparisons for the first lag, and approximately 100 comparisons for the remaining lags; and (3), the length of the lag interval should be approximately equal to 1/2 the maximum field site dimension divided by the number of points or lags. Keeping these guidelines in mind, test semi-variograms were produced at every 2.5 m interval from 35 m to 55 m. A lag interval of 52.5 m was found to produce semi-variograms with characteristics that closely matched the "rules-of-thumb" criteria, and was thus used to produce the nondirectional soil semi-variograms.

The GEOPACK (version 1.0) geostatistical software package (Yates and Yates, 1990) was used to model the experimental soil semi-variograms. Linear, spherical, and exponential models were fit to the semi-variograms using a non-linear least-squares procedure (Marquardt, 1963). The Akaike Information Criterion (AIC), a goodness-of-fit statistic, was determined for each model (Webster and McBratney, 1989). The model having the lowest AIC value--i.e., the "best" fit--was then cross-validated in order to reduce the mean and variance to 0 and 1, respectively. Results from the cross-validation procedure were considered to be the best description of the nondirectional experimental semi-variogram.

A final procedural step was performed in order to determine if direction, as well as distance, had a significant influence on spatial correlation. Four directional variograms (0, 45, 90, and 135 degrees) were produced for each soil parameter. By superimposing the nondirectional model over each directional semi-variogram, it was possible to tell if there was a strong directional component to the spatial correlation--i.e., was it anisotropic. Most of the semivariograms from the initial soil survey showed a marked degree of spatial dependence. The form of these relationships was generally isotropic and was best fit by spherical models. Experimental semi-variograms of four of the more significant soil parameters are shown below.

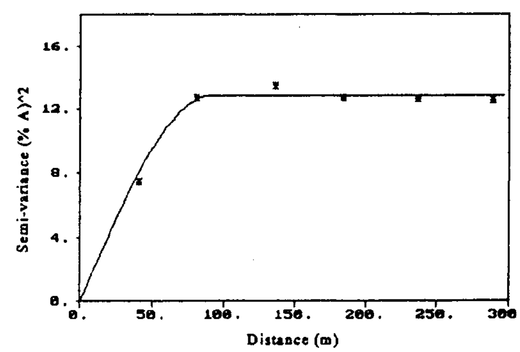

Figure 1. Depth of A horizon semi-variogram.

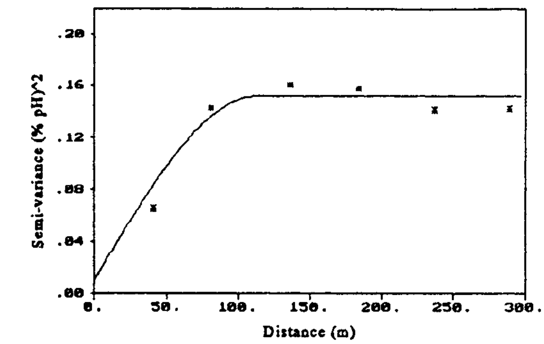

Figure 2. pH semi-variogram.

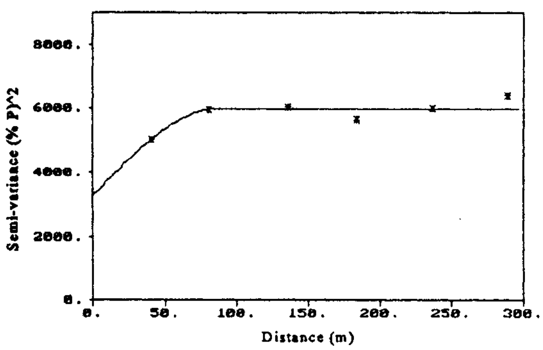

Figure 3. Phosphorus semi-variogram.

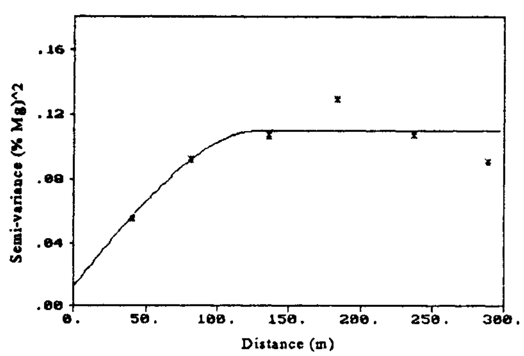

Figure 4. Magnesium semi-variogram.

Once the experimental soil semi-variogram models were estimated, it was possible to determine an optimal sampling scheme for future soil characterization. This was accomplished using OSSFIM version CG+2 (McBratney and Webster, 1981). OSSFIM requires input of the semi-variogram model type, nugget, range, sill, anisotropy ratio, and block size in order to produce a plot of grid sample spacing vs. kriging variance. This plot was then used to determine what sampling intensity was required to achieve a particular degree of kriging variance or prediction error. Table 1 shows what sampling frequencies are required to obtain soil data at a level of precision high enough to allow for separation of the class means.

Table 1. Grid spacing required to achieve desired level of precision.

Soil Parameter	Std. Error	Grid Spacing (m)
A Horizon Depth (cm)	1.0	23
pH (cmolH)	0.1	12
Phosphorus (kg/ha)	17.0	12
Magnesium (meq/100g)	0.1	26

CONCLUDING REMARKS

The design and layout of the experimental field plots at BARC has doubly benefited from the use of geostatistical techniques. By determining the spatial variability of potential crop yield, kriging will be used to help minimize the impact of location effects on treatment comparisons, and thus increase the power to detect true treatment differences. Moreover, geostatistical techniques have been used to design optimal sampling schemes for future data collection. This application was very relevant to our research because the expense associated with conducting long-term agricultural experiments makes it incumbent to obtain at least some level of assurance that the data used to establish field trials is precise enough for its intended purpose.

This investigation has shown that the kriging variance of the data collected using the 1993 sample grid (50 m x 75 m) was too high to allow for separation of the class means. To remedy this problem, a second soil characterization was performed in April 1994 using a 25 m x 25 m sample grid. Although Table 1 shows that both pH and phosphorus require a finer resolution sample grid, sampling at this frequency (12 m) was deemed to be prohibitively expensive; especially since these soil properties were felt to be less important predictors of crop yield than was depth of the A horizon.

Unfortunately, using geostatistical techniques to design optimal sampling schemes is hindered by one significant limitation: before an appropriate sampling scheme can be determined, the parameter's semi-variogram must be known; yet, estimation of this semi-variogram itself requires some sampling. This Catch-22 may not be a problem if the spatial behavior of the parameter is well known, or the costs associated with conducting a limited, exploratory survey in order to determine the semi-variogram are small compared to the costs of a full survey where the quality of the data collected is not explicitly known.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to P.A. Burrough and C. Wesseling for their expeditious reply to our request for a copy of the most recent version of OSSFIM. We would also like to thank W. Potts for the expert guidance and assistance he has provided us on this project.

REFERENCES

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

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

Burrough, P.A. 1991. Sampling designs for quantifying map unit composition. p. 89-125. *In* M.J. Mausbach and L.P. Wilding (ed.) Spatial variabilities of soils and landforms. SSSA Special Publication 28. SSSA, Madison, WI.

Christidis, B.G. 1931. The importance of the shape of plots in field experimentation. Journal of Agricultural Science, Cambridge 21:14-37.

Englund, E. and A. Sparks. 1991. GEO-EAS 1.2.1 Geostatistical environmental assessment software. EPA/600/8-91/008. U.S. Environmental Protection Agency.

Lam, N.S. 1983. Spatial interpolation methods: a review. The American Cartographer 10:129-149.

Lin, C.S., M.R. Binns, H.D. Voldeng, and R. Guillemette. 1993. Performance of randomized block designs in field experiments. Agronomy Journal 85:168-171.

Marquardt, D.W. 1963. An algorithm for least-squares estimation of non-linear parameters. J. Soc. Ind. Appl. Math. 11:431-441.

Matheron, G. 1963. Principles of geostatistics. Economic Geology 58:1246-1266.

McBratney, A.B. and R. Webster. 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables. II. Program and examples. Computers and Geosciences 7:335-365.

McBratney, A.B., R. Webster, and T.M. Burgess. 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I. Theory and method. Computers and Geosciences 7:331-334.

Nelson, L.A. and S.W. Buol. 1990. Experimental designs to evaluate crop response on adjacent soil mapping units. Soil Science Society of America Journal 54:841-849.

Oliver, M.A. and R. Webster. 1990. Kriging: a method of interpolation for geographical information systems. International Journal of Geographical Information Systems 4:313-332.

Oliver, M.A. and R. Webster. 1991. How geostatistics can help you. Soil Use and Management 7:206-217.

Perry, E.M., D.D. Kaufman, G.F. Hart, and B. Payne. 1993. Spatial analysis of remotely-sensed and groundbased data to explain location effects in crop yields. p. 265-271. *In* 1993 ACSM/ASPRS Annual Convention and Exposition: Technical Papers, Vol. 2, New Orleans, LA. 15-18 February. ASPRS and ACSM, Bethesda, MD.

Smith, F.H. 1938. An empirical law describing heterogeneity in the yield of agricultural crops. Journal of Agricultural Science, Cambridge 28:1-23.

van Es, H.M., C.L. van Es, and D.K. Cassel. 1989. Application of regionalized variable theory to large-plot field experiments. Soil Science Society of America Journal 53:1178-1183.

van Es, H.M. and C.L. van Es. 1993. Spatial nature of randomization and its effect on the outcome of field experiments. Agronomy Journal 85:420-428.

Webster, R. and A.B. McBratney. 1989. On the Akaike information criterion for choosing models for variograms of soil properties. Journal of Soil Science 40:493-496.

Yates, S.R. and M.V. Yates. 1990. GEOPACK 1.0 Geostatistical software system. EPA/600/8-90/004. U.S. Environmental Protection Agency.