Financial assistance was provided by the
United States Department of Agriculture and the
Soil Science Society of America

SOIL SPATIAL VARIABILITY

PROCEEDINGS OF A
WORKSHOP OF THE
ISSS AND THE SSSA
LAS VEGAS, USA
30 NOVEMBER - 1 DECEMBER 1984

D. R. Nielsen and J. Bouma (Eds)

CPRO-DLO
BIBLIOTHEEK
Postbus 16
6700 AA Wageningen

Pudoc Wageningen 1985

CIP

ISBN 90-220-0891-6

© Centre for Agricultural Publishing and Documentation, Wageningen, the Netherlands, 1985

No part of this publication, apart from bibliographic data and brief quotations embodied in critical reviews, may be reproduced, re-recorded or published in any form including print, photocopy, microform, electronic or electromagnetic record without written permission from the publisher PUDOC, P.O. Box 4, 6700 AA Wageningen the Netherlands.

Printed in the Netherlands.

Contents

Foreword

D. R. Nielsen, Chairman Working Group on Spatial and Temporal Variability of Field Soils, International Society of Soil Science
J. Bouma, Secretary Working Group on Spatial and Temporal Variability of Field Soils, International Society of Soil Science

These proceedings refer to the first meeting of the Working Group on Spatial and Temporal Variability on Field Soils on behalf of Commissions I (Soil Physics) and V (Soil Genesis, Classification and Cartography), of the International Society of Soil Science. The meeting entitled "Workshop on Soil Spatial Variability", was sponsored by the International Society of Soil Science, The Soil Science Society of America and the U.S. Department of Agriculture. The objectives of the Workshop were to explore and discuss alternative statistical concepts and procedures of (1) enhancing the understanding and development of pedology, and (2) improving technology of soil survey, soil science and hydrology applied to present-day management of field soils. During the past decade, several national and international symposia have focused on the collection and analysis of soil and related environmental data as regards their spatial and temporal variations. Amongst these symposia, we consider the intellectual framework of this Workshop to be unique because it attempted to relate current efforts on statistical and mathematical interpretation of variability to contemporary soil classification programs with a viewpoint to future research.

The Workshop consisted of invited papers and extended discussions in four general areas. The resulting one-half day sessions focused on general statistical concepts of quantifying variability and upon applications to hydrology, soil survey, and miscible displacement and leaching. In each session two or three invited speakers presented 30 - 45 minute lectures designed as reviews of conceptual models, statistical approaches and experimental methods useful in studies of spatial variation. Each speaker's presentation was followed by an open discussion during which participants presented comments based upon their own experiences or questions directed towards the speaker or the audience. The contents of this book are the unabridged presentations of Workshop participants. Each gave a full measure of his knowledge and experience to contribute toward the objectives of the Workshop.

On behalf of all participants, we wish to express our gratitude for the support of the three sponsoring organizations, and especially the effort of David M. Kral and his staff of the Headoffice of the Soil Science Society of America for providing the logistics

necessary for the Workshop. We also wish to express, on behalf of the participants, our sincere appreciation to each of the invited speakers - every one of whom was articulate and erudite - for focusing our attention on the underlying principles of the Workshop.

And, finally, the two of us are especially mindful and thankful for the participation of the more than 200 persons who attended and made the Workshop a successful event. We look forward to them and future readers of this book to continue to develop our understanding of soil variability and its application to soil management.

The role of geostatistics in the design and analysis of field experiments with reference to the effect of soil properties on crop yield

Alex B. McBratney, CSIRO Division of Soils, Cunningham Laboratory, St. Lucia, Queensland 4067, Australia

"Soil isn't important for crop yield, it doesn't give a significant effect."

A statement such as this or some close version of it has been heard by the author at various agronomic centers around the world. We, as soil scientists, find these expressions frustrating. My aim here is to discuss briefly why such statements are made and to suggest how geostatistical methods can lead to improved methods of field experimental design and analysis.

The Effects of Fisherian Design

There are two main reasons for statements of the kind made above. The first is Fisher's field experimental design and analysis. Fisher's great agronomic achievement in the 'twenties was to find a technological solution to the problem of soil variation in field experiments. Through randomisation and blocking he removed, without estimating, the effect of soil and other uncontrolled environmental variables. His approach was outstandingly successful and the methodology has largely stood the test of time. From this point of view, the statement above is an artifact of the method. Probably also as a result of the success of the method, the number of uniformity trials and studies of soil and crop yield covariation diminished markedly after the 1920's (cf Vieira et al., 1983). The second reason for the statement is that soil scientists, over a period of 60 years, have evidently not explained with sufficient clarity to statisticians and agronomists the importance of soil in crop growth and how it could be taken into account.

Recent Developments in Field Experimental Design

Recently, there have been some new ideas on field experimental design and analysis which at least try to account for the spatial variation of the crop (Green et al., 1984; Wilkinson et al., 1983; Patterson and Hunter, 1984; and Williams and Patterson, in press). The first two methods assume a smooth trend plus independent errors model, which seems unrealistic. The method of Patterson and Hunter assumes an isotropic exponential semi-variogram and that of Williams and Patterson an isotropic linear semi-variogram. These latter two methods do not appear to

be sufficiently general; my studies of uniformity trials, two of which are depicted in Fig. 1, show that the form of variation of crop yields may be non-stationary or periodic or anisotropic, or some combination thereof. Clearly, a method is required which allows the form of the semi-variogram to be estimated from the experiment, such a method is outlined below.

The Role of Geostatistical Methods

Briefly, geostatistics has a part to play in field experimental design because of its ability to describe quantitatively soil and crop variation and covariation and to perform block predictions and co-predictions. I see the use of geostatistics in three, proceedingly more involved, applications.

1. The use of soil property semi-variograms to design plot and block size and shape.

Assuming that there are a few soil properties controlling the spatial yield variation of the crop and there is some proportional relation between these variances, then the anisotropy of the semi-variogram will suggest plot (and block) shape. The smaller plot dimension should be in the direction of maximum variation and the larger dimension perpendicular with the ratio of the sides equal to the geometric anisotropy ratio. The size of the plots and blocks will depend on the form of the variation and variances for them can be calculated using integrals of the semi-variogram. Qualitatively, if transitive behaviour is observed blocking will not be required if plots with dimensions similar to the range can be used. If the range is very large compared to the available experimental region then blocking will proably be required. (It appears that classical experimental designers assume a linear semi-variogram!)

2. Co-spatial soil and crop surveys

If one does not wish to assume the form of the relationship between the variances of soil and crop attributes then a co-spatial soil and crop survey should be carried out. This is a combination of a geostatistical soil survey and a uniformity trial. For this to be fruitful, the size of yield plots and soil grids should be less than the range of spatial dependence. This should allow the calculation of cross semi-variograms of the soil and crop attributes as well as the regression of yield on soil attributes (allowance should be made for spatial dependence and a method such as that described by Cook and Pocock (1983) should be used).

3. Embedded field experimental designs

In an attempt to account for spatial soil and crop variation in field experiments, it is possible to perform a co-spatial soil

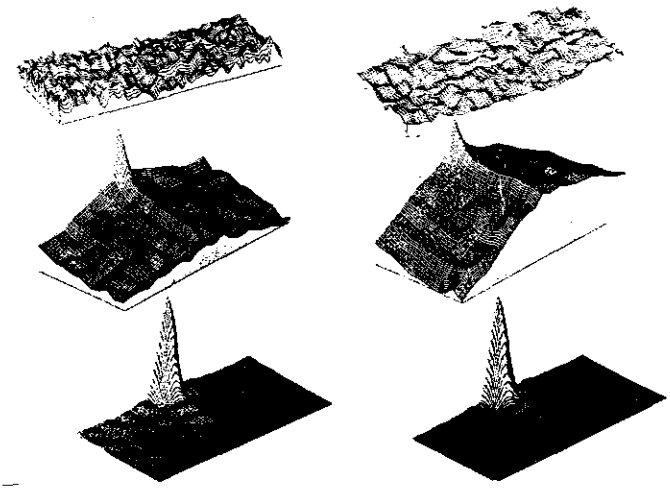


Fig. 1 Spatial analysis of two uniformity trials.

Left upper — The navel orange uniformity trial of Batchelor and Reed (1918). Surface showing the yield of oranges on 20 by 50 trees on a 22 ft. grid.

middle — Two dimensional autocorrelation surface to lag 10 in both directions. The ridge from center back to center front of this surface is in the same direction as the yield surface from back to front. Note the slight anisotropy.

lower — Two dimensional spectral density surface smoothed with a two dimensional Barlett window to lag 10. Directions are the same as the autocorrelation surface. There is no significant periodic behavior; all the power is at low frequencies.

Right upper — The IR8 rice uniformity trial of Gomez and Gomez (1976). Surface showing the yield of rice on 18 by 36, 1 meter square plots. Note the non stationarity from left to right.

middle — Two dimensional autocorrelation surface to lag 10 in both directions. The ridge from center back to center front of this surface is in the same direction as the yield surface from back to front. Note the strong anisotropy probably caused by the non stationarity.

lower — Two dimensional spectral density surface smoothed with a two dimensional Barlett window to lag 10. Directions are the same as the autocorrelation surface. There is no significant periodic behavior; almost all the power is at low frequencies.

For further discussion of the methods used to compute these diagrams, and for an example of periodic variation in a uniformity trial, see McBratney and Webster (1981).

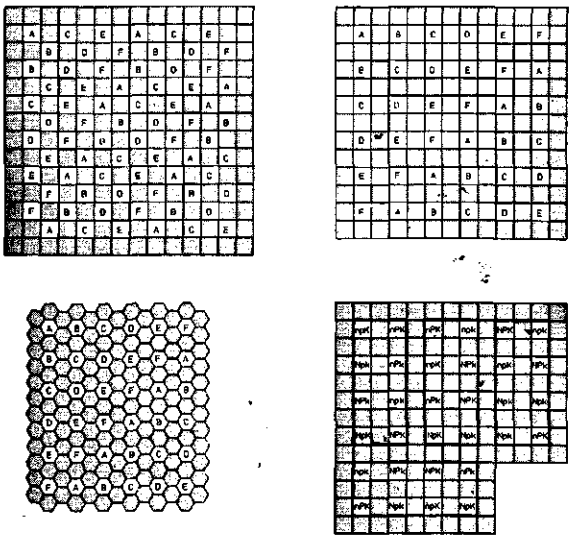


Fig. 2 Embedded field experimental designs

Left	upper	First order nearest neighbor (4 neighbors) embedded Latin square design with 2 replicates on a square grid.
	lower	First. order nearest neighbor embedded Latin square design on an equilateral triangular grid.
Right	upper	Second order nearest neighbor (8 neighbors) embedded Latin square design on a square grid.
	lower	Second order nearest neighbor embedded 2 cubed factorial design with 4 replicates on a square grid.
	N.B.	The grey plots need not necessarily be as large as the main treatment plots.

and crop survey as described above and follow it up in the next growing season with a classical field experiment on the same area. If spatial patterns vary from season to season then perhaps a better approach is to combine the co-spatial soil and crop survey and classical experimental designs - such a combination may be called an embedded design. Such schemes based on square and triangular grids are shown in Fig. 2. These diagrams are diagrammatic and are primarily intended to show the spatial nature of the designs. For example, in Fig. 2, it is questionable whether blocking of the replicates is required because of the local control afforded by these designs. The grey plots in each diagram are control plots and the named white plot is a treatment plot. The control plots can be measured for soil attributes and crop yield and a regression model obtained. These values can then be used to krige (or co-krige) the environmental yield on the treatment plots. Further soil analyses on the treatment plots will allow detection of any deviation between predicted soil effects and those observed, suggesting some interaction between treatments and soil properties.

Statisticians will eschew these embedded designs because relative to classical designs they will be regarded as inefficient. Presumably this inefficiency is measured by equating information gained with increased effort and area of the experiment. It may not be a simple matter to estimate the information gained from an experiment and the formulation of an economic or scientific loss function seems difficult. Statisticians have been extremely parsimonious in their experimental designs, however this has not led to complete satisfaction with the users of such schemes. Inefficiency is refuted on the grounds that these designs allow realistic spatial models to be fitted explicitly. To the experimenter there is no real substitute for this.

Conclusions

(1) At present there is a lack of quantitative co-spatial soil and crop data.

(2) The design, execution and analysis of spatially integrated soil and crop surveys and field experiments should become a joint research topic for agronomists, soil scientists and statisticians.

REFERENCES

1. Batchelor, L.D. & Reed, H.S., 1918. Relation of the variability of fruit trees to the accuracy of field trials. J. Agric. Res. 12:245-283.

2. Cook, D.G. & Pocock, S.J., 1983. Multiple regression in geographical mortality studies, with allowance for spatially correlated errors. Biometrics 39:361-371.

3. Gomez, K.A. & Gomez, A.A., 1976. Statistical methods for agricultural research with emphasis on rice. IRRI, Los Banos. pp. 208-209.

4. Green, P.J., Jennison, C., & Seheult, A.H., 1984. Analysis of field experiments by least-squares smoothing. J.R. Statist. Soc., B46 (in press).

5. McBratney, A.B. & Webster, R., 1981. Detection of ridge and furrow pattern by spectral analysis of crop yield. Int. Statist. Rev. 49:45-52.

6. Patterson, H.D. & Hunter, E.A., 1983. The efficiency of incomplete block designs in National List and Recommended List cereal variety trials. J. Agric. Sci. 101:427-433.

7. Vieira, S.R. et al., 1983. Geostatistical theory and application to variability of some agronomical properties. Hilgardia 51(3):1-75.

8. Wilkinson, G.N. et al., 1983. Nearest neighbour (NN) analysis of field experiments. J. R. Statist. Soc., B 45: 151-211.

9. Williams, E.R. & Patterson, H.D. A neighbor model for field experiments. Submitted to Biometrika.

Spatial variability: geostatistical methods

Allan Gutjahr, Department of Math and Statistical Research Center
New Mexico Institute of Mining and Technology, Socorro, NM 87801,USA

Introduction

Recent research in the soil sciences (Byers and Stephens, 1983; Bresler and Dagan, 1983) and hydrology (Delhomme, 1978, 1979; Bakr et.al. 1978; Gutjahr et. al., 1978; Gutjahr and Gelhar, 1981) emphasize the concepts of spatial variability in the study of soil properties and flow problems. The papers cited emphasize the statistical and stochastic nature of the phenomena.

In this paper I will present some of the basic concepts of spatial variability and illustrate their meaning. I will also discuss the assumptions involved, the type of data needed, and the advantages and limitations of these procedures.

Problems of variation in soil properties have certainly been recognized for many years. In fact much of the statistical work on design of experiments and analysis of variance originated in agricultural research. The main concern, however, was with differences between mean values and variation was often viewed as a nuisance to be controlled for in some manner.

By contrast, the geostatistical study of spatial variability deals with data and problems that involve uncontrollable variation that still has some kind of structure. Thus the data in space (or time) is presumed to have some connectedness or continuity embedded within the randomness. The objectives of such a study vary and include attempts to explain variation, to build predictive models, to interpolate or extrapolate values, to design sampling plans and to interrelate variations of different properties (e.g. to relate the variability of conductivity and head in a hydrologic context). Thus this approach views variation as part of an overall problem which can convey vital information about the phenomena studied.

Random Fields

The starting point in the geostatistical study of spatial variability is the notion of a random field or

random function in space ($\underset{\sim}{x}$ will designate a vector in 1, 2, or 3, space. The Glossory contains a more complete definition of terms). Simply, put a random field, $V(\underset{\sim}{x})$ or spatial stochastic process is a random variable for any fixed value of $\underset{\sim}{x}$. $V(\underset{\sim}{x})$ might represent quantities like hyraulic conductivity, grain size or head at location $\underset{\sim}{x}$.

A complete description of $V(\underset{\sim}{x})$ would require the joint probability distribution at any set of locations, $\underset{\sim}{x}_j$, $j = 1 \ldots n$. In practice this amount of information is virtually impossible to obtain and instead only descriptions of moments are used. These include the mean, $V(\underset{\sim}{x}) \equiv E(V(\underset{\sim}{x}))$, the variance, $Var(V(\underset{\sim}{x}))$ and, most importantly, a measure of the statistical relationship between $V(\underset{\sim}{x}_1)$ and $V(\underset{\sim}{x}_2)$ known as the covariance $cov(V(\underset{\sim}{x}_1), V(\underset{\sim}{x}_2))$.

We start by examining the concept of statistical homogeneity or second-order stationary. $V(\underset{\sim}{x})$ is statistically homogeneous (second-order stationary) if $V(\underset{\sim}{x})$ has constant mean and if $Cov(V(\underset{\sim}{x}_1), V(\underset{\sim}{x}_2))$ only depends upon the vector difference $\underset{\sim}{x}_1 - \underset{\sim}{x}_2$. Namely if

(i) $E(V(\underset{\sim}{x})) = m$, a constant

and (ii) $Cov(V(\underset{\sim}{x}_1), V(\underset{\sim}{x}_2)) = R(\underset{\sim}{x}_1 - \underset{\sim}{x}_2)$, or

$$Cov(V(\underset{\sim}{x}+\underset{\sim}{y}), V(\underset{\sim}{x})) = R(\underset{\sim}{y}).$$

$R(\underset{\sim}{y})$ is the covariance function evaluated at lag $\underset{\sim}{y}$.

The covariance condition states the statistical or probabilistic dependence is only related to the separation between the points. The constant mean condition is not a controlling factor if $m(\underset{\sim}{x}) = E(V(\underset{\sim}{x}))$ is known. Note also that $Var(V(\underset{\sim}{x})) = R(\underset{\sim}{0})$ for the statistically homogeneous case. In some applications the homogeneity assumption is further specialized by assuming $R(\underset{\sim}{y})$ only depends upon $y = |\underset{\sim}{y}|$, the length of the separation vector in which case we say the process is statistically isotropic.

How can $R(\underset{\sim}{y})$ be estimated? In the general statistically isotropic case we could first estimate the mean, $\bar{V}$, based upon observations $V(\underset{\sim}{x}_1) \ldots V(\underset{\sim}{x}_n)$ and then estimate R(y) by grouping points. Thus if Δ is some fixed value, we could group those points $(\underset{\sim}{x}_j, \underset{\sim}{x}_k)$ into the set A(y) where $y-\Delta \leq |\underset{\sim}{x}_j - \underset{\sim}{x}_k| \leq y+\Delta$ and then estimate R(y) as follows:

$$\hat{R}(y) = \sum_{\text{pts in } A(y)} [V(\underset{\sim}{x}_j)-\bar{V}][V(\underset{\sim}{x}_k)-\bar{V}]/\#\text{ pts in } A(y)$$

Some examples of possible one-dimensional or isotropic covariance functions are shown in Figure 1.

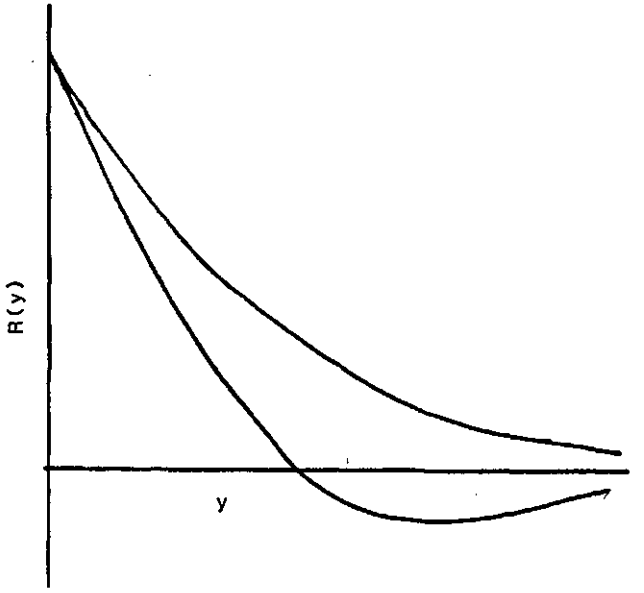


Fig. 1 Isotropic covariances

Generally, we expect covariances to decrease as y increases - i.e. the "influence" of $V(\underset{\sim}{x})$ dies out as we move further away from that point. An average distance over which significant correlation ($R(\underset{\sim}{y})/R(\underset{\sim}{0})$ = the correlation function) is called a scale. A common procedure is to take the scale λ to be that value where $R(\lambda)/R(0) = e^{-1}$ (e-fold drop).

There are cases where, at least within the area studied, $V(\underset{\sim}{x})$ does not appear to be statistically homogeneous. To study these situations the French geostatistics school (Matheron, 1971; 1973) has introduced the notion of intrinsic random functions where the increments or changes are assumed to be statistically homogeneous.

An intrinisic random function of order zero is a random field $V(\underset{\sim}{x})$ with

(i) $E(V(\underset{\sim}{x})) = m$

(ii) $\gamma(\underset{\sim}{y}) = E\{[V(\underset{\sim}{x}+\underset{\sim}{y})-V(\underset{\sim}{x})]^2\}/2$

$$= \frac{1}{2} \mathrm{Var}[V(\underset{\sim}{x}+\underset{\sim}{y})-V(\underset{\sim}{x})]$$

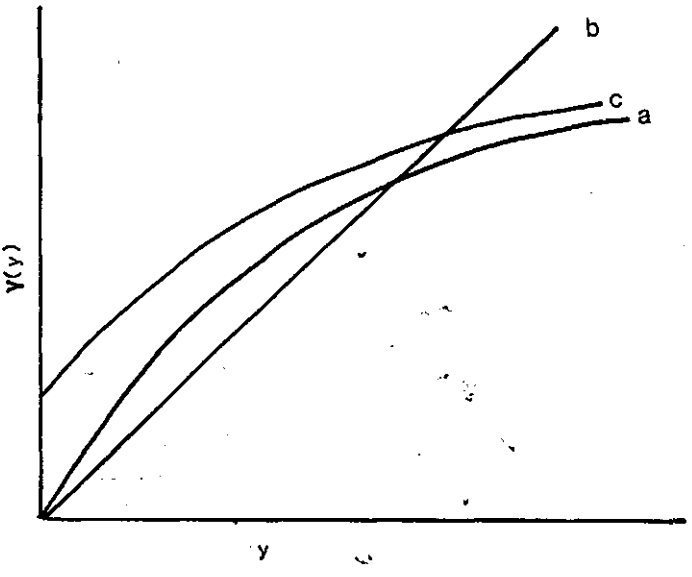


Fig. 2 Isotropic variograms

The function $\gamma(\underset{\sim}{y})$ only depends on the lag $\underset{\sim}{y}$ and is called the semi-variogram or just variogram. Typical variograms are shown in Figure 2 for 1-dimension. If $V(\underset{\sim}{x})$ is also statistically homogeneous, $\gamma(\underset{\sim}{y}) = R(\underset{\sim}{0}) - R(\underset{\sim}{y})$ and as $\underset{\sim}{y}$ increases, $\gamma(\underset{\sim}{y})$ approaches $\text{Var}(V(\underset{\sim}{y}))$ which is also called the sill. Correspondingly, the scale λ is sometimes called the range of the variogram.

If $V(\underset{\sim}{x})$ is not statistically homogeneous $\gamma(\underset{\sim}{y})$ will not approach a sill value (e.g. Figure 2-b). Figure 2-c shows a variogram with a jump at 0 - called a nugget effect. Such an effect isn't really possible if $V(\underset{\sim}{x})$ is continuous but it may be observed in estimates because fine-scale estimates aren't available; it can also be used to model measurement error.

Using the same set A(y) introduced for estimating R(y) an isotropic variogram can be estimated by

$$\hat{\gamma}(y) = \sum_{\text{pts in A}} [V(\underset{\sim}{x}_j) - V(\underset{\sim}{x}_k)]^2 / [2 \cdot \#\text{pts in A}(y)]$$

In some cases non-isotropy can be detected by calculating $\gamma(y)$ in different directions and possibly correcting for trends in the mean (see Figure 3)

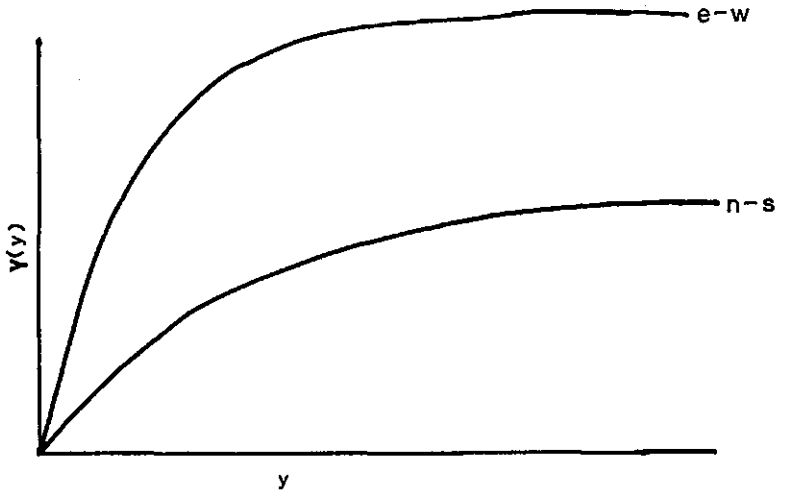


Fig. 3 Non-isotropic variograms

$R(\underset{\sim}{y})$ and $\gamma(\underset{\sim}{y})$ are thus used to characterize the variability that can exist for a random field. What else can be done with these functions?

Kriging

If $V(\underset{\sim}{x}_1)$, $V(\underset{\sim}{x}_2)$ ... $V(\underset{\sim}{x}_n)$ are measurements at n locations we may want to use them to interpolate or predict V at some other location $\underset{\sim}{x}$. How can this be done using either $R(\underset{\sim}{y})$ or $\gamma(\underset{\sim}{y})$? Kriging is the procedure commonly used to do this.

In kriging, linear estimators,

$$\hat{V}(\underset{\sim}{x}) = \sum_{j=1}^{n} \lambda_j V(\underset{\sim}{x}_j),$$

are used for the unknown value $V(\underset{\sim}{x})$. The λ_j's are chosen so that

(1) the estimator is unbiased:

$$E(\hat{V}(\underset{\sim}{x})) = E(V(\underset{\sim}{x}))$$

and (2) The mean squared error $E\{[\hat{V}(\underset{\sim}{x})-V(\underset{\sim}{x})]^2\}$ is minimum.

In both the statistically homogeneous and the intrinsic random function case minimizing the mean square error while using linear unbiased estimators will lead to n+1 linear equations in n+1 unknowns: $\lambda_1 \ldots \lambda_n$ and μ a Langrange multiplier. Note that the λ's and μ depend upon the location $\underset{\sim}{x}$ so that at each location we get different weights. Usually this dependence on $\underset{\sim}{x}$ is not explicitly shown.

The kriging variance, or minimum mean squared error is also usually obtained

$$\sigma_k^2 = \text{var}[\hat{V}(\underset{\sim}{x})-V(\underset{\sim}{x})] = E[\hat{V}(\underset{\sim}{x})-V(\underset{\sim}{x})]^2.$$

The kriging equations and kriging variances for the two cases discussed are given below.

a. Statistically homogeneous case

$$\sum_{j=1}^{n} \lambda_j R(\underset{\sim}{x}_i-\underset{\sim}{x}_j)-\mu = R(\underset{\sim}{x}-\underset{\sim}{x}_i);\ i = 1 \ldots n$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\sigma_k^2 = R(0)-\sum_{j=1}^{n} \lambda_j R(\underset{\sim}{x}_j-\underset{\sim}{x})+\mu.$$

b. Intrinsic Random Function of order 0

$$\sum_{j=1}^{n} \lambda_j \gamma(\underset{\sim}{x}_i-\underset{\sim}{x}_j)+\mu = \gamma(\underset{\sim}{x}-\underset{\sim}{x}_i);\ i = 1 \ldots n$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\sigma_k^2 = \sum_{j=1}^{n} \lambda_j \gamma(\underset{\sim}{x}_j-\underset{\sim}{x})+\mu$$

The equations in (b) can be obtained formally from those in (a) by taking $R(\underset{\sim}{0}) = 0$, $\gamma(\underset{\sim}{y}) = -R(\underset{\sim}{y})$.

The kriging estimators or interpolators have several other interesting properties in addition to unbiasedness and minimum mean square error.

(1) $\hat{V}(\underset{\sim}{x})$ is an exact interpolator: $\hat{V}(\underset{\sim}{x}_j) = V(\underset{\sim}{x}_j)$ at the observation points

(2) The weights λ_j and μ depend on the covariance function or variogram and the locations, but not on the actual values observed

(3) m, the mean, is not needed to calculate the estimate $\hat{V}(\underset{\sim}{x})$

(4) σ_k^2 yields a measure of the precision and again only depends on the λ's, μ, and the covariance function or variogram.

In a kriging study the above points are used in a variety of ways. For example, to validate a kriging model, an estimate is made at $\underset{\sim}{x}_{j_0}$, an observation point, by leaving out that point and using the remaining n-1 points to develop the kriging equations. This procedure is repeated with sucessively different points are excluded. The values $[\hat{V}_{j_0}(\underset{\sim}{x}_{j_0}) - V(\underset{\sim}{x}_{j_0})]^2/\sigma_k^2$ are calculated and averaged where the subscript j_0 indicates the omitted point. This average value should be close to 1 if the model assumptions are true and the correct covariance function (variogram) are used.

The fact that kriging variances only depend upon location and not actual V values can be used to see what effects added sampling can have on the estimates. Thus a fictitious point can be placed and the kriging weights and variances calculated for the changed situation. By moving the point one can decide on an "optimal" location for an added sample.

Figure 4 shows schematically what kriging estimates would look like. The kriged path is smoother than the actual path, as one might suspect. In Figure 5 the effect on σ_k of an added point is illustrated: note $\sigma_k = 0$ at observations since values are exact at those points.

Extensions and Modifications

The kriging procedure can be extended to cases

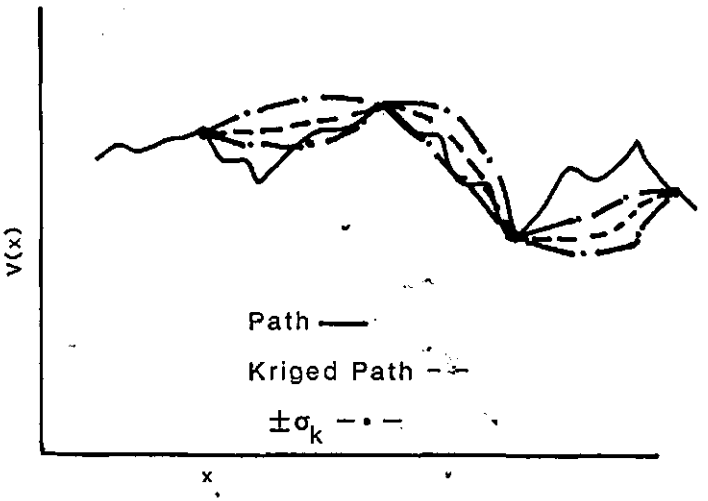


Fig. 4 Kriging illustration

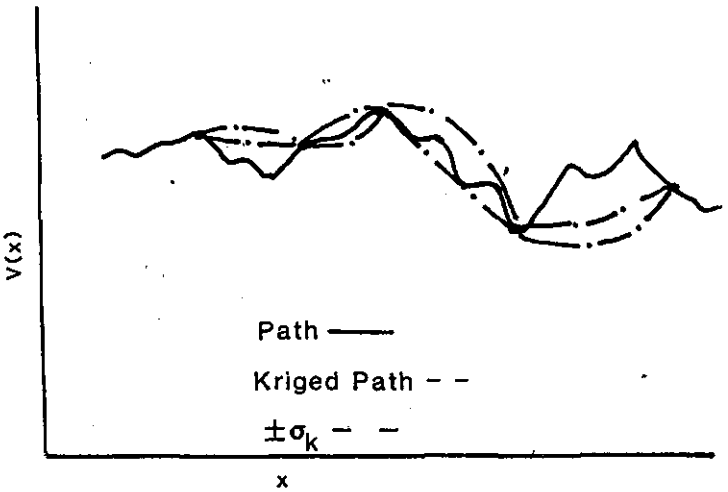


Fig. 5 Added Point effect

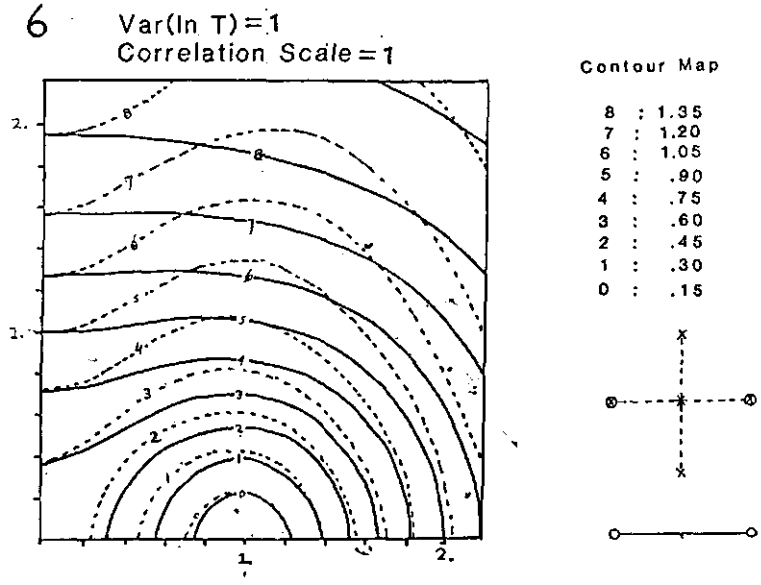
where $m(\underset{\sim}{x}) = E(V(\underset{\sim}{x}))$ depends on $\underset{\sim}{x}$ but the covariance or variogram assumptions still apply.

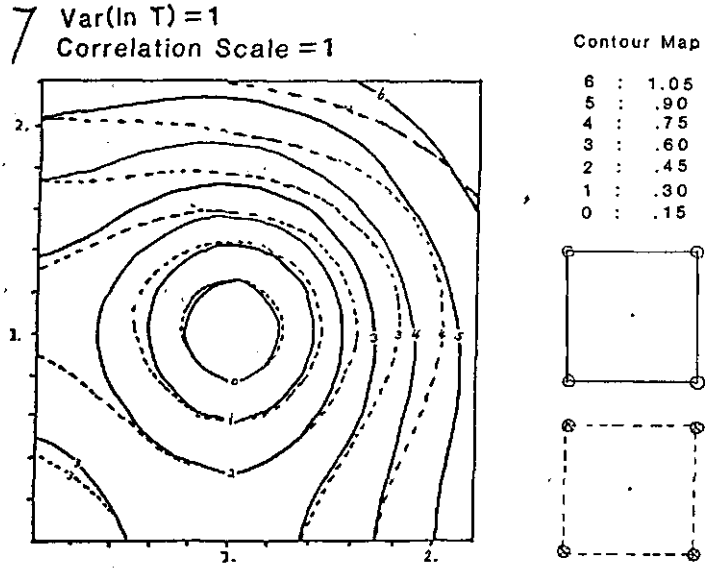
One procedure is to approximate $m(\underset{\sim}{x})$ by $\sum_{k=1}^{K} a_k P_k(\underset{\sim}{x})$ where the P's are a set of "basis" functions e.g polynomials). The unbiasedness conditions lead to K constraints and K Lagrange multipliers. The resulting kriging equations are called the universal kriging equations (Matheron, 1971).

An alternative is to look at higher-order differences and assume some kind of stationarity or homogeneity for these differences. This leads to higher-order intrinsic random functions and generalized covariances (Matheron, 1973). Again kriging equations can be developed for this case.

Another extension covers cases where we have more than one random field. This is an especially interesting case because one can speak of the "transfer" of information from one field to the other. In this case, for statistically homogeneous random fields $U(\underset{\sim}{x})$, $V(\underset{\sim}{x})$, the cross-covariance $R_{uv}(\underset{\sim}{y}) = \mathrm{cov}(U(\underset{\sim}{x}+\underset{\sim}{y}), V(\underset{\sim}{x}))$ is used as well as the auto-covariance functions. The kriging question now becomes the following: Given $U(\underset{\sim}{x}_{11}) \ldots, U(\underset{\sim}{x}_{1n})$ and $V(\underset{\sim}{x}_{21}) \ldots, V(\underset{\sim}{x}_{2m})$ find the prediction of $V(\underset{\sim}{x})$. The unbiasedness conditions can become more complicated. The use of this co-kriging procedure along with physical models that interlink the two fields is especially intriging because it can tell us something about the worth of different pieces of data.

In Figures 6 and 7 I show some plots of kriging variance for a head/transmissivity model where the two quantities are related by a flow equation. The cross-covariances needed are found by using the flow equation and a spectral approach I will touch upon briefly and that I'm sure Lynn Gelhar will refer to in his talk. In the Figures kriging variance contours are shown for 2 networks of observations in each case. The observations are taken at points separated by one or two correlation lengths of the lnT random function where O's denote transmissivity (T) measurements and crosses indicate head measurements. In Figure 6 the dashed contours correspond to the network with five head measurements (each one correlation scale apart) and 2 transmissivity measurements (2 correlation scales apart), while in Figure 7 the contours are for measurements on observations 2 correlation scales apart. The contours are only shown for the first quadrant and the networks are presumed to have their centers at the origin. This model contains within it the seeds for a kind of inverse procedure and a method





Figs 6 and 7 Kriging variance contours.
Co-kriging Transmissivity/Head

for evaluating worth of different types of data.

Kriging has also been used to study methods for creating what might be called possible realities or realizations of random fields. Recall the kriged paths (e.g. Figure 4) are smoother than the real paths. In some studies we want to re-create paths through the sampled data points (as a kriging estimator does) that exhibit the kinds of variation and correlation seen in reality.

Thus we want a conditional random field $V(\underset{\sim}{x})$, given observations $V(\underset{\sim}{x}_1)$ ... $V(\underset{\sim}{x}_n)$. Delhomme (1978, 1979) gives examples of such conditional simulations which use the kriging procedure. One need not use kriging to do this but it does lead to a rapid method for generating these realizations.

The conditional simulation approach first finds the kriging weights for the given covariance function and data locations. Then an unconditional random field $s_u(\underset{\sim}{x})$ is generated with the desired covariance behaviour. Finally this field is kriged, using the previous weights to get $s_u(x)$ and then the conditioned field,

$$s_c(\underset{\sim}{x}) = \hat{V}(\underset{\sim}{x}) + [s_u(\underset{\sim}{x}) - \hat{s}_u(\underset{\sim}{x})]$$

will have the desired properties. Conditional simulations are useful for studying input-output relationships in cases where two fields are related to each other.

Kriging - uses, abuses, advantages and disadvantages

The ready availability of computer codes makes kriging and its relatives good candidates for use and abuse in a variety of applications. In carrying out kriging studies there are things to be aware of and in this section some uses, warnings about misuses and problems that may be encountered will be discussed.

First let me discuss advantages of kriging some of which I have already mentioned.

(1) The λ's depend on $\underset{\sim}{x}$, $R(y)$ (or $\gamma(y)$) and the $\underset{\sim}{x}_i$ locations but not on the $V(\underset{\sim}{x}_i)$'s. This makes sample design studies possible and studies of where points should be placed to minimize variances.

(2) $\hat{V}(\underset{\sim}{x})$ is an exact interpolator. If the $V(\underset{\sim}{x}_i)$'s contain measurement error that can also be incorporated into the kriging equation.

(3) The mean m is not needed to get $V(\underset{\sim}{x})$.

(4) The kriging variance yields a useful measure of accuracy.

(5) The procedure is general and flexible if the assumption are valid.

What are some disadvantages and things to watch out for?

(1) The covariance (variogram) must be known. Very often the estimates of the covariance (variogram) behave poorly and are difficult to interpret. In addition the same points used to estimate the covariance (variogram) are often used in the kriging procedure. To minimize the effects of correlation and bias, standard functions are fit by eye to the experimentally observed functions. The effects on the kriging variance of using the same data for both the covariance (variogram) estimates and for obtaining the weights still needs exploration.

(2) The mean must be constant or have a known form. One can get around some of the problems by using intrinsic random function theory and generalized differences or by using universal kriging. However both of these options are considerably harder to apply and interpret. In addition these procedures are not as readily implemented. An alternative is to fit a mean equation or trend and remove that mean from the data. This may, however, have an undesirable effect on the covariance (variogram) estimates.

(3) The statistics of covariance (variogram) estimates are difficult to study and not well known.

(4) In theory, kriging can handle anisotropic covariance functions or variograms but again in practice that is hard to do. Virtually all standard kriging packages don't include options for anisotropy.

Kriging is not an automatic procedure. In carrying it out one should validate the model as much as possible

and be aware of the limitations. Used in that way, with thought, it can be a useful technique for studying spatial variability in space and time and its consequences.

Spectral Methods and Representations

Kriging is a kind of orthogonal projection procedure involving projection onto the data. It is similar in that respect to regression and other multivariate techniques. In carrying out analytical calculations that involve equations connecting statistically homogeneous random fields we often use yet another "orthogonalization" procedure. This is generally referred to as the spectral representation theorem (Lumley and Panofsky, 1964; Koopmans, 1974; Rosenblatt, 1973). I will state it and then discuss what is says and why it is so important.

The Spectral Representation Theorem

If $V(\underset{\sim}{x})$ is a statistically homogeneous random field with mean 0 and $R(\underset{\sim}{y}) = \mathrm{cov}(V(\underset{\sim}{x}+\underset{\sim}{y}), V(\underset{\sim}{x}))$ continuous at $\underset{\sim}{0}$, then there is a unique (with probability 1) complex process $Z(\underset{\sim}{k})$ and a positive non-decreasing function $F(\underset{\sim}{k})$ with the following properties:

$$\text{(i)} \quad R(\underset{\sim}{y}) = \int_{-\infty}^{\infty} \exp\{i\underset{\sim}{k}\cdot\underset{\sim}{y}\}dF(\underset{\sim}{y})$$

$$\text{(ii)} \quad V(\underset{\sim}{x}) = \int_{-\infty}^{\infty} \exp\{i\underset{\sim}{k}\cdot\underset{\sim}{x}\}dZ(\underset{\sim}{k})$$

$$\text{(iii)} \quad E(dZ(\underset{\sim}{k})) = 0, \quad E(dZ(\underset{\sim}{k})dZ^*(\underset{\sim}{k}')) = 0, \ \underset{\sim}{k} \neq \underset{\sim}{k}'$$

$$E(|dZ(\underset{\sim}{k})|^2) = dF(\underset{\sim}{k})$$

In the above the integrals are 1-2 or 3 dimensional depending on the region, $dF(\underset{\sim}{k}) = f(\underset{\sim}{k})d\underset{\sim}{k}$ if a spectral density $f(\underset{\sim}{k})$ exists; $F(\underset{\sim}{k})$ is the spectral distribution and * designates complex conjugate.

Now what does this formidable expression mean?

We can take the original process which involves correlation and inter-relationship and write it as an integral of a (complex) process that has uncorrelated components. It involves a disentangling of the $V(\underset{\sim}{x})$ process into "independent" pieces in the $Z(\underset{\sim}{k})$ process. For many problems the spectral representation will allow us to study terms in isolation because they don't mix or interact.

This procedure is analoguous to decompositions of sums of squares encountered in analysis of variance. There one tries to split the sources of variability into different parts and assign them to different causes. Namely, (ss = sum of squares)

ss total = ss between groups + ss within groups

is a basic identity in one-way analysis of variance. The same kind of decomposition occurs in the spectral representation above. In fact,

$$\mathrm{var}(V(\underset{\sim}{x})) = \int f(\underset{\sim}{k})d\underset{\sim}{k}$$

$$\cong f(\underset{\sim}{k}_1)\Delta\underset{\sim}{k} + f(\underset{\sim}{k}_2)\Delta\underset{\sim}{k} + \ldots$$

With $f(\underset{\sim}{k}_j)\Delta\underset{\sim}{k}$ = variance associated with frequency $\underset{\sim}{k}_j$, the decomposition of variance is like that in the analysis variance.

The study of spatial variability by using stochastic methods is still being developed. It can yield insights into the variation that exists in the field, the continuity of the field and the uncertainty one can expect. With proper care the methods should help in interpretation of data and in developing fundamental understanding of physical models.

BIBLIOGRAPHY

Bakr, A.A., L.W. Gelhar, A.L. Gutjahr and J.R. MacMillan, Stochastic analysis of spatial variability in subsurface flows, 1, comparison of one- and three-dimensional flows, Water Resour. Res., 14(2), 263-271, 1978.

Bresler, E. and G. Dagan, Unsaturated flow in spatially variable fields 2. Application of water flow models to various fields Water Resour. Res., 19, 421-428, 1983.

Byers, E. and D. Stephens, Statistical and stochastic analysis of hydraulic conductivity and particle size in a fluvial sand, Soil Science So. American Journal, 47, No. 6, 1072-1081.

Delhomme, J.P., Kriging in the hydrosciences, Adv. Water Resour., 1(5), 251-266, 1978.

Delhomme, J.P., Spatial variability and uncertainty in groundwater flow parameter: A geostatistical approach, Water Resour. Res., 15(2), 269-280, 1979.

Gutjahr, A.L., L.W. Gelhar, A.A. Bakr and J.R. MacMillan, Stochastic analysis of spatial variability in subsurface flows, 2, evaluation and application, Water Resour. Res., 14(5), 953-959, 1978.

Gutjahr, A.L. and L.W. Gelhar, Stochastic models of subsurface flow: Infinite versus finite domains and stationarity, Water Resour. Res., 17(2) 337-350, 1981.

Koopmans, L.H., The spectral analysis of time series, 366 pp., Academic Press, New York, 1974.

Lumley, J.L., and Panofsy, H.S., The structure of atmospheric turbulence, 239 pp., John Wiley, New York, 1964.

Matheron, G., The theory of regionalized variables and its applications, Ecole Des Mines De Paris, Fontainebleau, France, 1971.

Matheron, G., The intrinsic random functions and their applications, Adv. Appl. Prob., 5, 439-468, 1973.

Rosenblatt, M., Random processes, Oxford University press, 208 pp., 1962.

Random field or spatial stochastic process $\underset{\sim}{V}(\underset{\sim}{x})$: This means for each fixed $\underset{\sim}{x}$ (e.g. a location in space) $\underset{\sim}{V}(\underset{\sim}{x})$ is a random variable.

Probability density of $\underset{\sim}{V}(\underset{\sim}{x})$: The function $g(v; \underset{\sim}{x})$ such that the Probability $\underset{\sim}{V}(\underset{\sim}{x})$ is between a and b, denoted by $P(a < \underset{\sim}{V}(\underset{\sim}{x}) \leq b)$, is given by

$$P(a < \underset{\sim}{V}(\underset{\sim}{x}) \leq b) = \int_a^b g(v; \underset{\sim}{x})dv.$$

Expected value or mean of $\underset{\sim}{V}(\underset{\sim}{x})$: The expected value or mean (denoted by $E(\underset{\sim}{V}(\underset{\sim}{x}))$ or $\bar{\underset{\sim}{V}}(\underset{\sim}{x})$) is the probability-weighted average,

$$E(\underset{\sim}{V}(\underset{\sim}{x})) = \int_{-\infty}^{\infty} v\, g(v; \underset{\sim}{x})dv.$$

Covariance function for a random field: This is designated by $\mathrm{cov}(\underset{\sim}{V}(\underset{\sim}{x}_1), \underset{\sim}{V}(\underset{\sim}{x}_2))$ and is defined as

$$\mathrm{cov}(\underset{\sim}{V}(\underset{\sim}{x}_1), \underset{\sim}{V}(\underset{\sim}{x}_2)) = E[(\underset{\sim}{V}(\underset{\sim}{x}_1)-\bar{\underset{\sim}{V}}(\underset{\sim}{x}_1))(\underset{\sim}{V}(\underset{\sim}{x}_2)-\bar{\underset{\sim}{V}}(\underset{\sim}{x}_2))]$$

It measures statistical relationship between field values at two different locations.

Statistical homogeneity or second-order stationarity: The random field $\underset{\sim}{V}(\underset{\sim}{x})$ is statistically homogeneous or second-order stationary if

(i) $E(V(\underset{\sim}{x})) = m$, a constant.

and (ii) $\text{cov}(V(\underset{\sim}{x}_1), V(\underset{\sim}{x}_2)) \equiv R(\underset{\sim}{x}_1 - \underset{\sim}{x}_2)$ only depends upon the separation vector $\underset{\sim}{x}_1 - \underset{\sim}{x}_2$. $\underset{\sim}{y} = \underset{\sim}{x}_1 - \underset{\sim}{x}_2$ is called the lag vector.

Covariance function for statistically homogeneous processes: $R(\underset{\sim}{y}) = \text{cov}(V(\underset{\sim}{x}+\underset{\sim}{y}), V(\underset{\sim}{x}))$, the covariance function as a function of the lag $\underset{\sim}{y}$.

Statistically isotropic process: A statistically homogenous process where $R(\underset{\sim}{y}) \equiv R(|\underset{\sim}{y}|) = R(y)$ only depends on the separation distance, $y = |\underset{\sim}{y}|$.

Correlation function: $\rho(\underset{\sim}{y}) = R(\underset{\sim}{y})/R(\underset{\sim}{0})$.

Scale: An average distance over which points are significantly correlated. For an isotropic covariance function this is sometimes taken as that value λ where $e^{-1} = \rho(\lambda)$, an e-fold drop.

Intrinsic random function of order 0: A random field $V(\underset{\sim}{x})$ with constant mean where $E([V(\underset{\sim}{x}+\underset{\sim}{y})-V(\underset{\sim}{x})]^2)$ only depends on $\underset{\sim}{y}$.

Variogram or semi-variogram: The function

$$\gamma(\underset{\sim}{y}) = E([V(\underset{\sim}{x}+\underset{\sim}{y})-V(\underset{\sim}{x})]^2)/2$$

for an intrinsic random function of order 0.

Sill: If $\gamma(\underset{\sim}{y})$ has a limiting value as y increases, the limit is called the sill and equals $\text{var}(V(\underset{\sim}{x}))$.

Range: The scale in a variogram for a statistically homogeneous process.

Unbiasedness: If $G(V(\underset{\sim}{x}_1) \ldots V(\underset{\sim}{x}_n))$ is some function of $V(\underset{\sim}{x}_1) \ldots V(\underset{\sim}{x}_n))$ it is an unbiased estimator of $V(\underset{\sim}{x}_0)$ if

$$E(G(V(\underset{\sim}{x}_1) \ldots V(\underset{\sim}{x}_n)) = E(V(\underset{\sim}{x}_0)).$$

Mean squared error: If $G(V(\underset{\sim}{x}_1) \ldots V(\underset{\sim}{x}_n))$ is an estimator of $V(\underset{\sim}{x}_0)$ the mean squared error of G is

$$E([G(V(\underset{\sim}{x}_1) \ldots V(\underset{\sim}{x}_n)) - V(\underset{\sim}{x}_0)]^2).$$

It is a measure of how "close" the estimator G is to the quantity being estimated, $V(\underset{\sim}{x}_0)$.

Estimator, $\hat{V}$: For estimators the notation

$$\hat{V}(\underset{\sim}{x}_0) \equiv G(V(\underset{\sim}{x}_1) \ldots V(\underset{\sim}{x}_n)) \text{ is often used.}$$

Linear Estimator: An estimator of the form

$$\hat{V} = \lambda_1 V(\underset{\sim}{x}_1) + \lambda_2 V(\underset{\sim}{x}_2) + \ldots \lambda_n V(\underset{\sim}{x}_n)$$

$$= \sum_{j=1}^{n} \lambda_j V(\underset{\sim}{x}_j)$$

where the λ's are constants.

Kriging: The procedure that finds the best (minimum mean square error) linear unbiased estimator of $V(\underset{\sim}{x})$ based upon observations $V(\underset{\sim}{x}_1) \ldots V(\underset{\sim}{x}_n)$. For a statistically homogeneous process with covariance function $R(\underset{\sim}{y})$, this yields a set of linear equations for the "weights" λ_j;

$$\sum_{j=1}^{n} \lambda_j R(\underset{\sim}{x}_j - \underset{\sim}{x}_i) - \mu = R(\underset{\sim}{x}_i - \underset{\sim}{x}), \quad i = 1 \ldots n$$

$$\sum_{j=1}^{n} \lambda_j = 1.$$

μ is a Lagrange multiplier. For an intrinsic random function of order zero the covariance function $R(y)$ can be replaced by $-\gamma(\underset{\sim}{y})$, the negative of the variogram, to get the kriging equations.

Kriging variance: The variance associated with the kriging estimator, designated by σ_k^2. This is also the minimum mean squared error. For a second-order stationary process it is

$$E([\hat{V}(\underset{\sim}{x})-V(\underset{\sim}{x})]^2 = \sigma_k^2 = R(\underset{\sim}{0}) - \sum_{j=1}^{n} R(\underset{\sim}{x}-\underset{\sim}{x}_j)+\mu$$

For an intrinsic random function, again R is replaced by $-\gamma$.

Co-kriging: The extension of kriging to the case where $\underset{\sim}{V}(x)$ is estimated using observations from two random fields where now the cross-covariance, $\text{cov}(\underset{\sim}{V}(x+y), \underset{\sim}{U}(x))$, also enters in.

Spectral density for a statistically homogeneous random field: A Fourier transform of the covariance function ($\underset{\sim}{k}$ = wave number or frequency),

$$f(\underset{\sim}{k}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{R(\underset{\sim}{y})}{(2\pi)^p} \exp\{-i\underset{\sim}{k}\cdot\underset{\sim}{y}\} dy_1 \cdots dy_p$$

$$\equiv \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} R(\underset{\sim}{y}) \exp\{-i\underset{\sim}{k}\cdot\underset{\sim}{y}\}\, d\underset{\sim}{y}$$

where p = the dimension of the space. The

covariance function can be recovered from the spectral density by an inverse Fourier transform.

Spectral representation theorem for a statistically homogeneous, mean zero, process: This states that if $\underset{\sim}{V}(\underset{\sim}{x})$ is statistically homogeneous random field with $E(\underset{\sim}{V}(\underset{\sim}{x})) = 0$, covariance function $\underset{\sim}{R}(\underset{\sim}{y})$, spectral density $\underset{\sim}{f}(\underset{\sim}{k})$ then there is a unique (probability 1) complex process $\underset{\sim}{Z}(\underset{\sim}{k})$ with the following properties

(i) $$\underset{\sim}{V}(\underset{\sim}{x}) = \int_{-\infty}^{\infty} e^{-i\underset{\sim}{k}\cdot\underset{\sim}{x}}\, d\underset{\sim}{Z}(\underset{\sim}{k})$$

(ii) $$E(d\underset{\sim}{Z}(\underset{\sim}{k})) = 0,\ E(d\underset{\sim}{Z}(\underset{\sim}{k}_1)\, d\underset{\sim}{Z}^*(\underset{\sim}{k}_2)) = 0,$$

$$\underset{\sim}{k}_1 = \underset{\sim}{k}_2,$$

and $$E(|d\underset{\sim}{Z}(\underset{\sim}{k})|^2) = \underset{\sim}{f}(\underset{\sim}{k})\, d\underset{\sim}{k}$$

(* = complex conjugate.)

The spectral representation theorem resolves the original process into a "uncorrelated" complex process which greatly simplifies some calculations.

Discussion

D. Goss and L. Wilding: What is meant by a constant mean and what are the requirements for a constant mean?

A. Gutjahr: The constant mean assumption means there is no trend within the range of study. It is not as crucial to the analysis if (i) the trends can be estimated and removed or (ii) one uses generalizations of kriging like universal kriging, where the trend can be estimated, or intrinsic random functions of higher order.

H. ten Berge: Is the condition that the variance be constant across the field implied in second order stationarity and is it a severe condition?

A. Gutjahr: Yes. It is implied by the definition of second order stationarity and it may be a serious restriction in some applications. In such cases one might look at some transformed values (like logarithmic transformations) of the data, try multiplicative types of corrections or use variograms and the intrinsic random function approach.

H. ten Berge: Is biasedness only a characteristic of the sampling procedure?

A. Gutjahr: No. It can also be a characteristic of the statistical procedures used in estimation and prediction.

B. Luxmoore and B. Schuhway: In the case of a linear variogram is there some of deciding between a trend in the data and a stationary process without a trend in the data?

A. Gutjahr: One can sometimes detect trends in the data by examining the variogram in different directions. For example, if there is a linear trend in a particular direction, the variogram in that direction will have a parabolic behavior and one can then remove the trend in that direction (see Delhomme, 1978). The question of a trend is often a difficult one to decide on the basis of the data and may again require higher order intrinsic random functions, or using subsidiary knowledge of the process to remove that trend.

C. Kirda: If a set of data has trends should the trend be removed before finding the correlation length from an auto-correlation analysis?

A. Gutjahr: Yes. Often trends can be spotted from the covariance functions if they have shoulders or tend to remain high as the separation distance increases. Correlation lengths calculated from the data without the trend removed are then too large, and also tend to include a deterministic component in the random model.

M. Nash: Is there any test for the range or correlation distance in correlation functions?

A. Gutjahr: There are tests for correlation distances (using, for example spectral analysis) and in the stationary case the tests would be the same for the range in the variogram. However these procedures often yield estimates with high variances and hence are not very precise.

W. Jury: There have been several instances of an apparent correlation between sampling grid density and correlation scale on a given field or even on the same transect of a field. It has been postulated that this would always result from non-linear drifts which are difficult to remove. How does this affect the measurability of the correlation scale?

A. Gutjahr: Yes, this can affect the measurement of the correlation scale. It can also occur that within the region of interest there may be several scales. As Lynn Gelhar indicated in his talk, the scale of the problem and the scale of the correlation function are interrelated. The region of study should be large enough to include several correlation scale lengths in which case the larger scale variation may be treated as a trend. In addition if there are significant non-linear trends one might again go to an intrinsic higher order function approach (using generalized differences) and estimate the generalized covariance.

L. Wilding: How many couples of observation are needed at the greatest distance for statistical reliability or validity?

A. Gutjahr: Unfortunately the statistical behavior of the variogram is not a simple one and still deserves

more study. As a rule of thumb one should have about 50 data points to get reasonably reliable estimates and then perhaps a good second rule of thumb is not to place much reliance on variogram estimates beyond 20% of the maximum distance (this second rule of thumb is often used in time series analyses). Namely if the biggest pairwise separation is 100 meters only look at the variogram out to 20 meters as having reasonable reliability. This should be tempered, of course, by looking at the number of data pairs in the estimates.

R. Horton: You have described two tools: correlogram and the semi-variogram. The correlogram requires second order stationarity in order to be valid while the semi-variogram requires less confining intrinsic assumption. Should one use the semi-variogram instead of the correlogram?

A. Gutjahr: If non-stationarity is a possibility then I would examine the semi-variogram. If a sill is shown then I would use a correlation function because it is often easier to interpret and also treat statistically (especially if the data is equally spaced).

S. Rao and K. Cassel: Is there any implicit assumption made about the distribution (e.g. normality) for data used in kriging. If not why do some researchers first transform the data?

A. Gutjahr: In both variogram estimation and kriging the assumption of normality is not needed - only expected values or averages enter in. Transformation may be used however for a variety of reasons like stabilizing the variance, making probability statements about kriging estimates and for developing procedures for estimating variograms (e.g. via maximum likelihood procedures). However the kriging procedure itself doesn't require distributional assumptions.

P. Nkedi-Kizza and L. Stroosnyder: What are the basic differences between classical agricultural statistics and geostatistics? At what stage should either be used? Can one use both? What about independent assumptions?

A. Gutjahr: Generally in classical agricultural statistics one is interested in estimating means and also, generally, the data are presumed to be independent or contain independent random error. Of course situations do exist where variables are assumed to be related (like analysis of covariance) but the main focus is on controlling for variability with appropriate designs and procedures.

In geostatistics the focus is on the inhomogeneities that exist and on the relationship between values - correlation occupies center stage. Thus if one wants to estimate the variability and continuity that exists, geostatistics is one way to do that.

The two (classical and geostatistical methods) can be used together. For example one might use geostatistics to estimate underlying variability within a field and then classical statistics to compare treatments in different areas after correcting for inherent variability.

In addition if measurements are far enough apart (as judged by examining the variogram or covariance function) observations become uncorrelated and again one might use classical procedures for widely separated points.

R. Bruce: Kriging is called an "exact" interpolator. What happens if measurement errors exists? Is kriging as a predictor to be applied only within the domain of the original study?

A. Gutjahr: One can extend the kriging procedure to include measurement error by modifying the covariance or variogram function in which case we no longer have an exact interpolator. Kriging is an exact interpolator within the field of interest and I should not really have referred to it as a predictor.

C. Gantzon and M. Nash: Can you discuss the difference in efficiency between taking samples randomly within the field versus ordered regularly spaced sampled? If the samples are placed randomly how is the semivariogram estimated?

A. Gutjahr: I don't know what the efficiency would be. I would prefer a regular grid or smaller grids superimposed on larger grids for estimating

variograms if there are many sample points because this would allow the use of spectral methods if needed. However in general this is not an easy question to answer. The estimation would proceed as if the samples were non-random by using grouping of data points as described in the paper.

E. Bresler: To best estimate the variogram by a least squares procedure one needs a sufficient number of pairs to cover the whole range of lags from zero to the maximum. In cases where the number of data points is small (30-50 points) random sampling might be preferred over regular sampling to get estimates at larger lags.

For estimating variograms and the trend, maximum likelihood methods can be useful if the data is from a multivariate normal population.

J. Hendrickx: What is an appropriate method to validate kriging results?

A. Gutjahr: A method often used for this purpose is the "leave-out-one data point" or jackknife procedure described in the paper.

B. Schuh: What are the effects of improper selection of the variogram in kriging?.

A. Gutjahr: In the stationary case the important features are generally the sill and the scale or range One can also try different variogram models and see what effects occur on the kriging estimates and kriging variances.

J. Allen: How can a priori patterns of spatial variability (like soil forming factors) be incorporated into geostatistical methods to account for spatial anisotropy?

A. Gutjahr: Here one might try to use modeling studies to account for known factors or try to put in features as added constraints. In general there is no all encompassing answer and one needs to study the specific situation.

C. Wang: In a kriging study we did with 50 columns of soil samples (150 samples within each column) the semi-variogram for each column varied

greatly in shape. How can this information be used to study a similar land form outside the study area?

A. Gutjahr: It is difficult to give an answer to this without looking at the data - kriging and semi-variogram studies are not always automatic. However it may be the case that you have a large degree of anisotropy in the region which yields widely differing semi-variograms.

I. Muraka: Could you elaborate on the "size of landscape" to which geostatistical analysis techniques would be limited to?

A. Gutjahr: Again I would refer to some later comments by Lynn Gelhar on problems of scales. There may be several scales and we need to take the scale of the problem into account when doing this kind of study - that often is information obtained not within the geostatistical study but from other sources.

F. Whisler: When will someone write a book with underlying theory and detailed examples and with programs adaptable to P.C.S.?

A. Gutjahr: Probably when someone has a lot of time and is assured they can make lots of money by doing so! Seriously, though, it is a subject still evolving - as more researchers use it and develop it I suspect texts will be forthcoming that meet the desired objectives. I would be careful, though, in looking at texts to be sure that they emphasize both what can't be done as well as what can be done, and that include good doses of common sense and thought - the procedures and methods aren't panaceas and can't be applied without thought. They should not be considered as "methods looking for a problem."

Time series in the soil sciences: is there life after Kriging?

R.H. Shumway, Division of Statistics, University of California, Davis, USA

1. Introduction

The development of statistical techniques for analyzing data in the soil sciences has traditionally followed along the lines laid out by early disciples of R.A. Fisher (e.g., Snedecor and Cochran (1967)) who assumed that observations obtained in the field were independent and identically distributed. The recent shift away from this early methodology, termed "aggie statistics", for example, by Nielsen et al. (1983) has been fueled by the realization that data collected in agricultural field trials and in the soil sciences are inherently spatially correlated. It is physically more reasonable to expect that measurements of soil parameters such as temperature or electrical conductively should be correlated when they are measured at adjacent points in space or time.

The resulting trend in the soil sciences has been to lean more heavily on geostatistics by which is meant the smoothing of experimental data using Kriging techniques developed by Matheron (1963). Such applications to the soil science are well documented in papers by Nielsen et al. (1982), (1983), Vieira et al. (1983) and Valdin et al. (1983) and by other participants in this workshop (see also Ripley (1981) or Journel and Huijbregts (1978)). The advantage of Kriging or CoKriging techniques is that one can do smoothing with a very sparse collection of observed data points whereas conventional time series techniques require that one collect relatively equally-spaced data from the random field.

The continued development of remote sensing devices and other systems of instrumentation, however, will soon enable research workers to bring to bear a number of alternative techniques for analyzing multidimensional random fields. Early techniques proposed by Whittle (1953), (1954) can be used to develop approaches to the problems of modelling and fitting data using stochastic partial differential equations. The use of spectral methods, suggested by McBratney and Webster (1981), or Nielsen et al. (1983) can be extended to embrace old fashioned aggie concepts using spectral analysis of variance (cf. Shumway (1970), (1971), Brillinger (1979)). The use of lagged regression models (cf. Brillinger (1975), Priestley (1981)) can be considered in order to develop input-output models relating various measured soil parameters. Problems involving non-stationarity and missing data

in transects can now be approached using the state-space approach for smoothing and signal extraction (cf. Parzen (1984)).

The purpose of the following discussion is to give some examples which demonstrate some of the kinds of questions which can be answered using the above techniques. The basic thrust of all of the methods is to identify models for the underlying processes and then to use standard statistical procedures based on maximum likelihood to estimate parameters and test hypotheses. In many cases, the models express the original unobserved series as solutions to stochastic differential equations driven by white noise. This links the statistical approach to realistic physical models which have been used to describe the dynamic interactions of soil science parameters. The emphasis will be on signal extraction as opposed to the best linear unbiased criterion used by the Kriging method.

2. Signal Extraction

A very versatile model can be developed when it is suspected that some underlying phenomenon of interest satisfies a first or higher order differential equation. The general form of the state-space model assumes that some unobserved underlying pxl vector signal of interest $\underline{x}(s) = (\underline{x}_1(s),\ldots,\underline{x}_p(s))'$ can only be observed through the qxl observation equation

$$\underline{y}(s) = M(s)\underline{x}(s) + \underline{v}(s) \qquad (1)$$

for $s=1,2,\ldots,n$ where $\underline{y}(s) = (\underline{y}_1(s),\ldots,\underline{y}_q(s))'$ denotes the observed vector at the spatial point s, $M(s)$ is a known qxp measurement matrix and $\underline{v}(s)$ is a qxl zero-mean vector noise process with qxq covariance matrix, $\text{cov}(\underline{v}(s)) = R$. Although the signal process $\underline{x}(s)$ is unobserved, it is assumed to satisfy the first-order difference (differential) equation

$$\underline{x}(s) = \Phi\,\underline{x}(s-1) + \underline{w}(s) \qquad (2)$$

where Φ is a pxp transition matrix and $\underline{w}(s) = (w_1(s),\ldots,w_p(s))$ has zero-mean and covariance matrix Q. Equation (2) is called the state equation and describes the evolution of the state-vector $\underline{x}(s)$ through space or time. The beginning value $\underline{x}(0)$ is assumed to have mean $\underline{\mu}$ and initial covariance Σ. The process $\underline{x}(s)$ can be stationary or non-stationary depending on the specification of the parameters $\Phi, Q, R, \underline{\mu}$ and Σ.

The model in the above form is partially identified, but there is still the problem of estimating the parameters. This is usually accomplished using various nonlinear optimization techniques to maximize the likelihood function (see papers by Kohn & Ansley, Harvey, Jones and Shumway in Parzen volume (1984)). The model identification phase generally makes use of the Akaike Information Criterion, AIC, defined as (Akaike (1974))

$$\text{AIC} = -2\ \log(\text{Likelihood}) + 2(\text{no. of parameters}) \qquad (3)$$

where one chooses the model for which AIC is a minimum.

A second important problem is that of extracting the signal $\underline{x}(s)$ from the data for given values of the parameters. The problems of developing the state-space model and estimating the state vector $\underline{x}(s)$ were solved in the landmark papers by Kalman (1960) and Kalman & Bucy (1961), who gave simple recursive solutions for the minimum mean square estimators; the procedure is now referred to as Kalman filtering and smoothing. Rather complete expositions of the basic principles involved can be found in Anderson & Moore (1979) or Jazwinski (1970). The advantage of the recursions is purely computational, since the ordinary linear minimum mean square estimator for x(s) involves inverting npxnp or nqxnq matrices whereas the Kalman filter-smoother procedure involves inverting n pxp or qxq matrices.

Before turning to an example, it is useful to relate the model given above to the one used in Kriging. The ordinary Kriging model writes the univariate (p=q=1) version of equation (1) without the noise term, say as

$$y(s) = x(s),$$

where $\underline{x}(s)$ is assumed to be stationary with constant mean value μ (Universal Kriging assumes a general mean $\underline{\beta}'\underline{z}(s)$ where $\underline{\beta}$ and $\underline{z}(s)$ are qxl vectors) and covariance function $\text{cov}(x(s_1),x(s_2)) = C_x(s_1-s_2)$. The Kriging estimator at $s=s_0$ is the linear unbiased estimator for $y(s_0)$ which has minimum variance. The form of the covariance is specified on a-priori grounds from a reasonable class of covariance functions using the variogram.

$$\gamma_x(m) = \frac{1}{2} E(x(s+m) - x(s))^2 \qquad (4)$$

$$= C_x(0) - C_x(m)$$

as a guideline. The main differences between Kriging and Kalman filtering using the state-space model are as follows:

1. Computational: Kriging requires inverting the (n+1)x(n+1) augmented covariance matrix. For p=q=1, Kalman filtering requires no matrix inversions.

2. Modelling: Kriging chooses from a class of stationary covariance functions. The state-space model specifies the first order model (2) and may be non-stationary.

3. Estimation: Kriging uses ad-hoc analysis of the variogram and noiseless prediction. The state-space procedure uses maximum likelihood estimation of the parameters and signal extraction under noise.

It should be noted that the signal extraction approach for the stationary signal plus noise model

$$\underline{y}(s) = \underline{x}(s) + \underline{v}(s), \tag{5}$$

where all vectors are pxl, originated in the work of Kolmogoruv (1941) and Wiener (1949) who showed, using the spectral approach, that the optimal minimum mean square solution could be reduced to inverting pxp spectral matrices. This requires that one know the form of the spectra for the noise and signal processes. It would seem to be a promising direction in which to move if the series are multidimensional, that is, they depend on the vector parameter $\underline{s} = (s_1, s_2, \ldots, s_d)'$ so that the Kalman recursions will not work. If the observations are regularly observed over a reasonably large grid, the spectral approximations are valid and a considerable computational simplification results over direct brute force Kriging or CoKriging. We do not give details here but the reader is referred, for example, to Priestley (1981). An example in the one-dimensional case where the signal and noise spectra are estimated by maximum likelihood is given in Shumway (1984).

As an example, of the state-space methodology, consider the data in Figure 1 taken from the study done by Morkoc et al. (1984) giving the mean values (over five transects) of yield and water and salt content at intervals of one five transects) of yield and water and salt content at intervals of one meter. The sprinkling system was arranged to distribute more salt (and more water) along the right-hand side of the transect. Morkoc et al. (1984) consider jointly modelling salt and water content using a p=q=2 dimensional version of (1) and (2).

In order to illustrate the versatility of the state-space approach, we consider a simple smoothing model for a single series which is related to spline smoothing (cf. Erh (1972), Kimball (1974)), Wecker & Ansley (1983) and has been proposed in another context by Kitagawa (1981) and Kitagawa & Gersch (1984). Assume for a single one of the series that we observe, say

$$y(s) = x(s) + v(s) \tag{6}$$

where v(s) has variance σ_v^2, which can be interpreted as observation noise. The signal x(s) is assumed to satisfy a second-order difference (differential) equation of the form

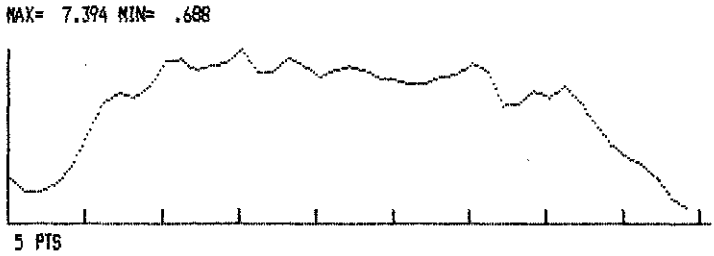
$$\nabla^2 x(s) = w_1(s) \tag{7}$$

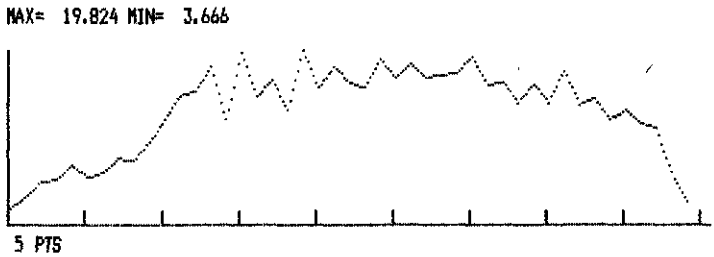
where

$$\nabla x(s) = x(s) - x(s-1) \tag{8}$$

so that

$$\nabla^2 x(s) = \nabla(\nabla x(s)) = x(s) - 2x(s-1) + x(s-2)$$





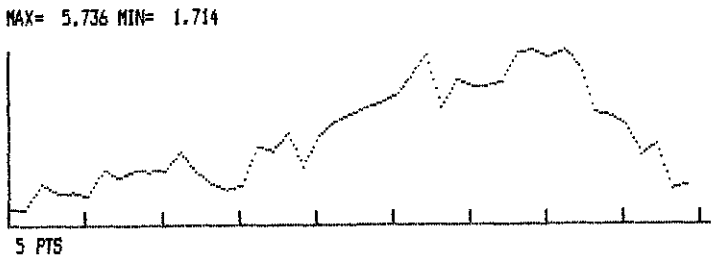


Fig. 1 Average over five transects of yield and associated water and salt content (1 pt = 1 m). (Morkoc et al. (1984))

Now, by defining the the state vector as $\underline{x}(s) = (x(s), x(s-1))'$, we may write the above model in the state-space form

$$y(s) = (1, 0)\begin{pmatrix} x(s) \\ x(s-1) \end{pmatrix} + v(s) \tag{9}$$

with

$$\begin{pmatrix} x(s) \\ x(s-1) \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x(s-1) \\ x(s-2) \end{pmatrix} + \begin{pmatrix} w_1(s) \\ 0 \end{pmatrix} \tag{10}$$

The identification allows us to use the maximum likelihood procedure for estimating the parameters σ_v^2 and σ_w^2 (Φ is known in this case) and the mean of x(0).

The EM algorithm of Dempster et al. (1978) as developed in Shumway & Stoffer (1982) was used to estimate the parameters and gave the results shown in Table 1 below:

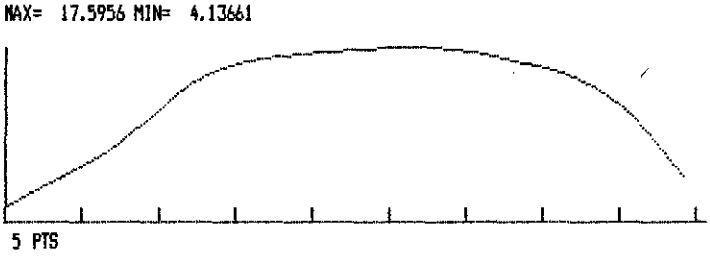
Table 1: Signal Extraction Parameters for Soil Data

	Observation σ_v^2	Model σ_w^2	Std. Dev. of Predicted Value
Yield	.109	.070	.19
Water	2.550	.079	.62
Salt	.102	.021	.16

The values for the water content appear to have the largest observation error which leads to a larger standard error for the predicted value.

The smoothed values as computed by the Kalman filtering-smoothing recursions (see Jazwinski (1970)) are shown in Figure 2, and they seem to do an excellent job of capturing the non-stationary trend behavior without smoothing out critical components.

The special form considered here is obviously not the only model which can be treated under the state-space framework. One might want to add another component into the measurement equation which satisfies a first-order difference equation of the autoregressive moving average type (see Box and Jenkins (1970) or Harvey (1981)).



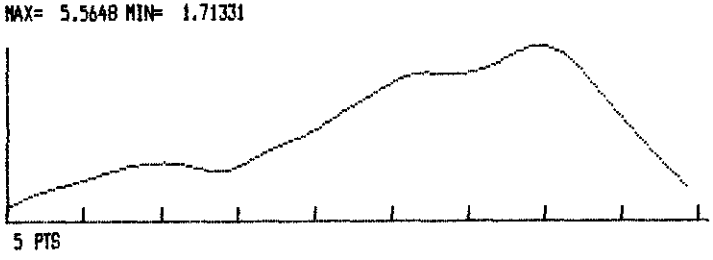


Fig. 2 Signals extracted from Figure 1 using the state-space model in equations (9) and (10) with parameters as specified in Table 1.

3. Spectral Analysis of Variance, Lagged Regression

The fact that data are autocorrelated over space does not preclude one from using analysis of variance or regression techniques as long as a collection of independent spatial series, say, along independent transects, can be identified. Then, a collection of transects with, say, characteristic A, can be compared with a collection of transects with characteristic B, keeping in mind the fact that observations made within a given transect are still highly autocorrelated. It is well known that when the spatial series are stationary, the notion of variance is expressed in terms of the separate kinds of cyclical variation found in the series. The variance measured as a function of spatial oscillations called frequencies is called the power spectrum. Nielsen et al. (1983) have identified furrows, tractor compaction and pre-plant irrigation as possible causes for cyclic variation in soil and have predicted the behavior to be expected in the power spectrum from such causes.

In order to illustrate the possibilities along these lines, we consider some rather ancient and well-worn data on wheat yields due to Mercer & Hall (1911).

The data, shown as a rough contour plot in Figure 3, are grain yields in lbs. per plot, recorded over an acre which was divided into 500(20x25) plots 2.5m wide and 3.3m long (length corresponded to the East-West direction). The strongest characteristic of the observed yields is an apparent periodicity running down columns. The mean row profile averaged over the 25 columns is shown in Figure 1 and also exhibits this cyclical behavior. One way to analyze this periodic behavior over columns is to calculate the two dimensional wavenumber spectrum as in Ripley (1981) or McBratney & Webster (1981) who attributed the periodicity to an earlier plowing of the area into ridges and furrows.

We discuss here an approach that seems to be particularly appropriate for transect samples or for two-dimensional fields where the phenomenon of interest seems to be occurring parallel to or orthogonal to rows. Kunsch (1982) has compensated for this tendency in the Mercer-Hall data by adjusting all values for the column means. To be specific, assume that the s^{th} yield in the j^{th} row or transect satisfies the model

$$y_j(s) = g(s) + v_j(s) \qquad (11)$$

$j=1,\ldots,N$, $s=0,\ldots,n-1$, where $g(s)$ is an unknown fixed mean signal and $v_j(s)$ $j=1,\ldots,N$ are autocorrelated but mutually independent identically distributed stationary processes.

The natural test to make in this case (cf. Shumway (1970), (1971)) is of the hypothesis that the mean signal $g(s)$ is absent and this should be done on a frequency dependent (cycle by cycle) basis.

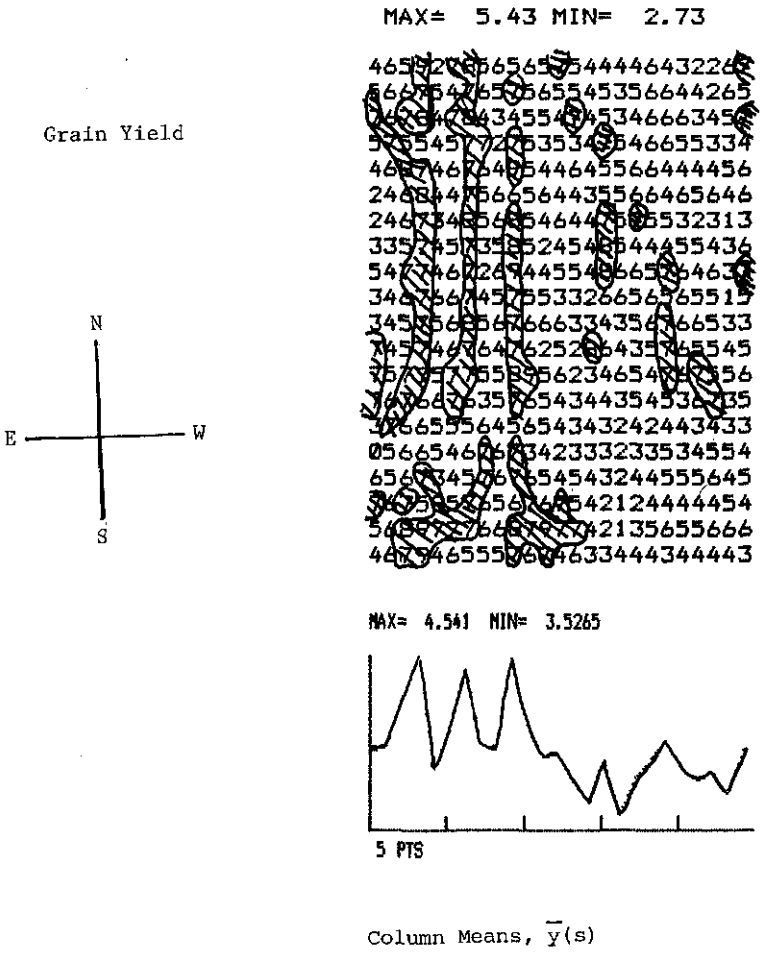


Fig. 3 Two dimensional plot of Mercer-Hall (1911) wheat data and column means. Yield in lbs. over 1 acre divided into 20 x 25 = 500 plots (3.3m x 2.5m).

The resulting spectral analysis of variance approach is completely analogous to what is done in the ordinary case except that the observed series is replaced by the transformed series

$$Y_j(k) = n^{-1/2} \sum_{s=0}^{n-1} y_j(s) \exp\{-i\lambda_k s\}, \qquad (12)$$

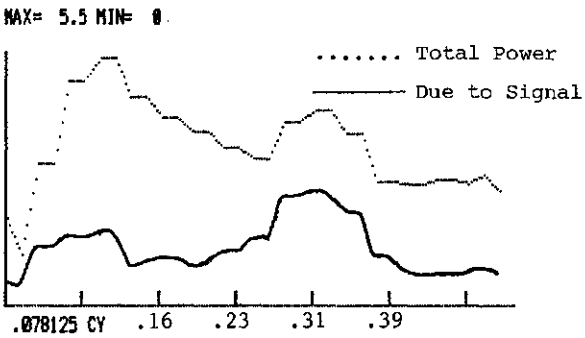
where $\lambda_k = 2\pi k n^{-1}$, $k=0,1,\ldots,n-1$ are the possible periodic components in the data. The transformation (12) is the discrete version of the usual Fourier transform and can be computed rather quickly by various algorithms (cf. Cooley & Tukey (1965), Bloomfield (1976)). The advantage is that for stationary process, the transformed values are approximately uncorrelated for different k for n reasonably large. In addition, the variance of the transformed error $v_j(s)$ is approximately the power spectrum at frequency λ_k.

We can illustrate the procedure for the special model given in equation (11) by specifying the spectral analysis of variance in a form analogous to that usually encountered in conventional experiments (see Shumway (1970), (1971), (1984), Brillinger (1979), (1983)). The F-test for detecting the signal g(s) is the ratio of two mean square power components and can be plotted as a function of frequency.

The series observed in Figure 3 are somewhat short for the asymptotics of the spectral approximations but periodicity is strong enough so that this data has been useful in the past for illustrating spectral methods. The spectral power components are more easily interpreted if they are plotted as a function of frequency and we do this in Figure 4. The frequency scale is in columns per cycle, and we note that there is an impressive component in the neighborhood of .31 cycles per column, corresponding to a period of 1/.31 = 3.2 columns or 8m. As mentioned before, this could be due to an earlier ridge and furrow plowing.

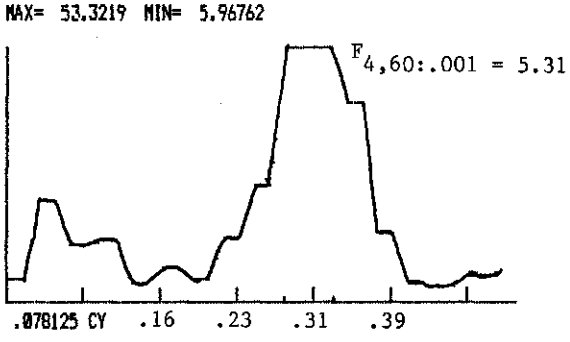
The example given above shows only a very special kind of analysis of variance situation. The procedure can be extended to testing equality of treatment means under various design configurations. A number of possible designs are covered in Brillinger (1979) Extensions to the vector and multidimensional cases are discussed in Shumway (1982) who also gives an example from seismology involving discriminating between waveforms generated by earthquakes and nuclear explosions.

Another question of interest in some soil science applications involves investigating possible lagged regressions which assume that the output series at transect j, say $y_j(s)$, is related to a collection of p input series $x_{j1}(t),\ldots,x_{jp}(t)$ through



Power Components

F STATISTIC
DF1= 4.6875 DF2= 89.0625



F Statistic

Fig. 4 Spectral analysis of variance (cycles per column) (1 col. = 2.5m).

$$y_j(s) = \sum_{m=1}^{p} x_{jm}(s) * \beta_m(s) + v_j(s) \qquad (13)$$

The notation * denotes the convolution, i.e.,

$$x_{jm}(s) * \beta_m(s) = \sum_{u=-\infty}^{\infty} x_{jm}(u)\beta_m(s-u) \qquad (14)$$

and the noise processes are assumed to be stationary processes with identical autocorrelations. The problems of interest relate to (1) determining which of the inputs are significant contributors, and (2) estimating the form of the impulse-response functions, $\beta_m(s)$.

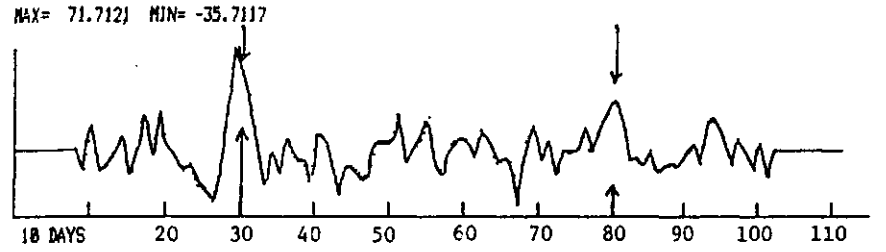
For example, Mechergui (1984) investigated the logarithm of hydraulic conductivity $y_j(s)$ as a function of the silt, sand and clay contents $x_{j1}(s)$, $x_{j2}(s)$ and $x_{j3}(s)$ over rows. As another example, Figure 5 shows three possible environmental series as inputs and cardiovascular mortality as a possible output. The series, measured daily for j=1,...,14 winters in London, are taken from Shumway et al. (1983).

The objectives in both of the above applications would be to determine what kind of smoothing filters are necessary to connect the inputs to the output. The basic regression computations for the stationary case can be found in Brillinger (1974) (see also Shumway (1970), (1983) for the case where series are repeated).

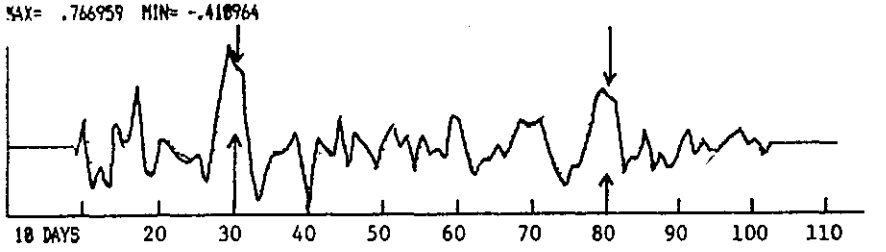
In order to identify which inputs are significant contributors to the output series under any given model of the form (13), one can look at an analog of multiple correlation defined over frequency. The resulting multiple coherence function and a spectral analysis of variance procedure was used in Shumway et al. (1982) to arrive at a final model which identified the sulfur dioxide (SO2) and temperature series as the primary contributors to cardiovascular mortality; the strongest coherence was at an approximate period of ten days.

The structural relation which is producing a significant coherence can sometimes be inferred from the estimated impulse-response functions, shown here in Figure 6. In this case, it can be seen that pollution acts positively and instantaneously whereas mortality has a negative effect at a lag of two days. This means that a decrease in temperature produces an additional number of cardiovascular deaths two days later.

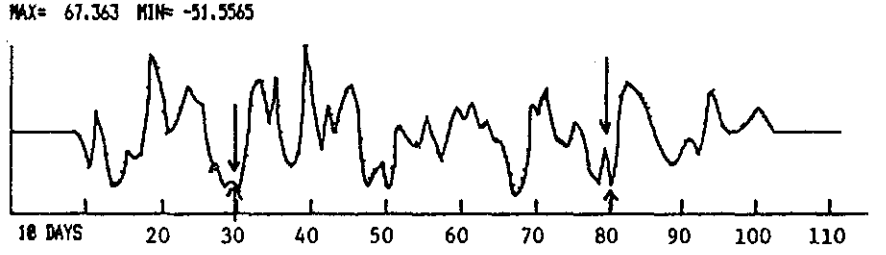
The regression approach can also be followed without transforming to the frequency or wavenumber domain. This makes the computations extremely involved, and one must either assume a form for the autocorrelation as in CoKriging or collect spatial series from many transects. An alternative which may be acceptable in some contexts is to assume that the error terms $v_j(s)$ are not



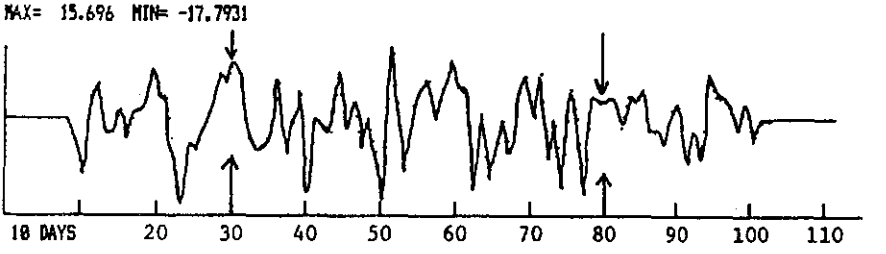
FILTERED CARDIOVASCULAR MORTALITY- LONDON WINTER(1962)



FILTERED SO2(LOG)- LONDON WINTER(1962)



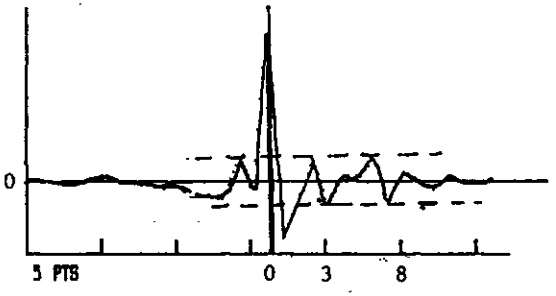
FILTERED TEMPERATURE- LONDON WINTER(1962)



FILTERED RELATIVE HUMIDITY- LONDON WINTER(1962)

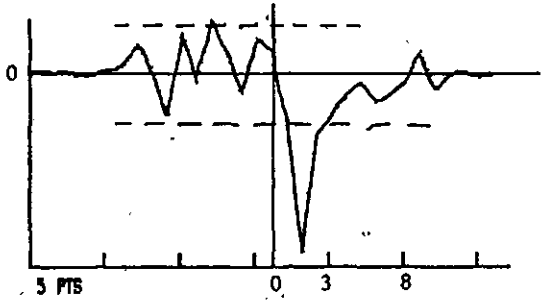
Fig. 5 Filtered, Detrended Cardiovascular Mortality and Associated SO_2 and Temperature Levels for the Winter of 1962.

MAX= 8.16033 MIN= -2.66635



TIME VERSION OF FILTER
CARDIOVASCULAR MORTALITY VS SO2 (PREFILTERED)

MAX= .0231278 MIN= -.0755022



TIME VERSION OF FILTER

CARDIOVASCULAR MORTALITY VS TEMPERATURE (PREFILTERED)

Fig. 6 Multiple impulse response functions relating cardiovascular mortality to daily pollution and temperature levels (dotted lines indicate ± 2 standard errors).

autocorrelated and use ordinary least squares.

4. Stochastic Differential Equations

In Section 2., it was noted that a common way of incorporating realistic physical assumptions into the signal extraction methodology is to assume that the underlying signal satisfies some stochastic differential equation. In that case, the assumption served as an alternative to assuming a parametric form for the autocorrelation as would be done in Kriging.

A number of considerations related to using stochastic and deterministic differential equations in the soil sciences were given in Nielsen et al. (1983). The use of the partial differential equation

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial s^2} - v \frac{\partial c}{\partial s} \qquad (15)$$

to model solute movement in soils where $c(s,t)$ is the solute concentration at time t for vertical depth s (D is the apparent diffusion coefficient, v is the pore water velocity) has been explored in Amoozegar et al. (1982), Nielsen et al. (1982), and Biggar & Nielsen (1976). The modelling of hydraulic conductivity as a function of space and time using Boussinesq's equation gives a similar second-order differential equation (cf. Gelhar (1974), Gutjahr et al. (1978) and Mechergui (1984)).

The approximation of such equations as (15) with finite difference equations has also been suggested by Wierenga et al. (1982) who were modelling heat transfer in a one-dimensional isotropic medium. This suggestion combined with the theoretical work of Gelhar (1974) and Gutjahr (1978) on stochastic differential equations sets the stage for models expressed as partial difference equations driven by white noise. Larimore (1977) has reviewed the early work of Whittle (1953), (1954), (1963) and developed computational techniques for modelling and fitting stochastic difference equations to data generated by stationary random fields. Applications of fitting techniques for random field data can also be found in Besag (1974), Besag & Moran (1975). Kunsch (1982) fits Markovian models to a number of agricultural field trial experiments including the Mercer-Hall data considered earlier in this paper. Cliff and Ord (1981) have given a description of the spatial Markovian model and have discussed the problem of fitting such models by maximum likelihood. A disadvantage of such procedures which do not transform to the frequency domain is that, like Kriging, they are computationally intensive and involve inverting large matrices.

As an example of a random field, Larimore (1977) gives the perspective plot in Figure 7 which is a realization of a random field generated at spatial coordinates s_1 and s_2 by an equation of the form

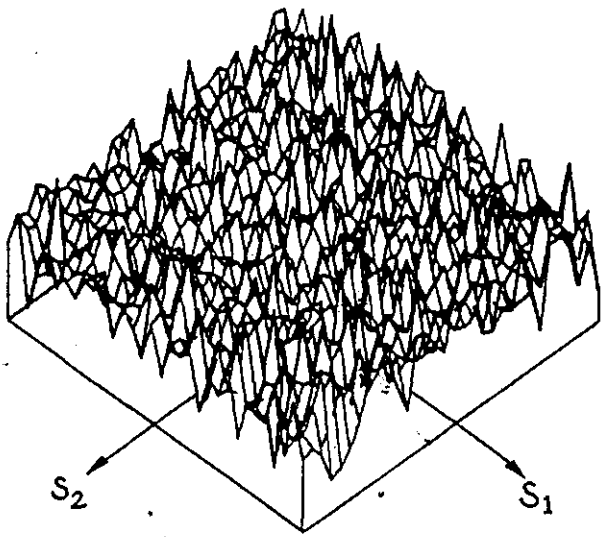


Fig. 7 Realization of the random field generated by Equation (17) (from Larimore (1977)).

$$x(s_1,s_2) + \alpha_1[x(s_1+1,s_2) + x(s_1-1,s_2)] + \alpha_2[x(s_1,s_2+1) + x(s_1,s_2-1)] + \alpha_3[x(s_1+1,s_2+1) - x(s_1-1,s_2+1) - x(s_1+1,s_2-1) + x(s_1-1,s_2-1)] = w(s_1,s_2) \tag{16}$$

which expresses the process in terms of adjacent values and $w(s_1,s_2)$, a white noise process with variance σ_w^2. The coefficients for this example were $\alpha_1 = .1$, $\alpha_2 = -.1$, and $\alpha_3 = .0867$. A compact notation is needed for equations of the above form and it is convenient to define

$$D_1^r\, x(s,t) = x(s+r,t) \tag{17}$$

and

$$D_2^r\, x(s,t) = x(s,t+r) \tag{18}$$

so that (17) can be rewritten as

$$[1+\alpha_1(D_1+D_1^{-1}) + \alpha_2(D_2+D_2^{-1}) + \alpha_3(D_1-D_1^{-1})(D_2-D_2^{-1})]\ x(s_1,s_2)$$

$$= w(s_1,s_2) \qquad (19)$$

This model can be compared, for example, with a white noise model of the form

$$x(s_1,s_2) = w(s_1,s_2) \qquad (20)$$

or perhaps with an approximately isotropic model of the form

$$[1 + \alpha_1(D_1+D_1^{-1}+D_2+D_2^{-1})]\ x(s_1,s_2) = w(s_1,s_2) \qquad (21)$$

One could also consider a simplification of the form

$$[1 + \alpha_1(D_1+D_1^{-1}) + \alpha_2(D_2+D_2^{-1})]\ x(s_1,s_2) = w(s_1,s_2) \qquad (22)$$

which specifies different coefficients for different directions. Larimore (1977) considers estimating the parameters in these several kinds of models using maximum likelihood in the frequency domain and computes the Akaike information criterion AIC (see equation (3)) for each of them. A partial listing of his results is given in Table 2 below.

Table 2 - Parametric models for random field in Figure 1 (taken from Larimore (1977)).

Model eq. no.	Shifts	Coeff. est.	AIC
(20)	---------		12,564
(21)	$D_1+D_1^{-1}+D_2+D_2^{-1}$	.00735	12,566
(22)	$D_1+D_1^{-1}$	.10253	12,486
	$D_2+D_2^{-1}$	-.09342	
(19)-True	$D_1+D_1^{-1}$	.1046	12,444
	$D_2+D_2^{-1}$	-.1045	
	$(D_1-D_1^{-1})(D_2-D_2^{-1})$	.0816	
Full Model	$D_1^iD_2^j$, i,j=0,±1	Given in Larimore (1977)	12,452

It is clear, in this case, that the correct non-isotropic model would be chosen by AIC and that the coefficient estimators are quite close to the true values of $\alpha_1 = .1$, $\alpha_2 = -.1$ and $\alpha_3 = .0867$.

Many theoretical developments relating to estimation and testing hypotheses for multidimensional autoregressive random fields still need to be explored. For example, the estimation and smoothing problems cannot be solved easily when the data are irregularly observed over the grid since the Kalman filtering and smoothing algorithm used in Section 2 is not available for the multidimensional case. Another area under development is the specification of hypotheses related to isotropy (see Solo (1984), for example). Alternative parameter specifications for the dependence structure as a function of distance such as the space-time autoregressive moving average STARMA models of Pfeifer and Deutsch (1980) may be of interest.

5. Discussion

The approach that one may decide to take in analyzing spatially correlated data depends mainly on (1) the theoretical models which are likely to be of use in describing the physical phenomena being studied; and (2) the sampling configurations over which the data are observed.

Theoretical models may be relatively non-specific, for example, such as simply requiring that certain components be non-stationary or stationary. More often than not it seems that the phenomena of interest satisfy certain systems of differential equations and these assumptions can be incorporated into parametric models and smoothing procedures. The additive 'signal plus noise' approach seems to be useful for both stationary and non-stationary data. This pertains to the material in Section 2 and the first example of Section 3. For stationary series, the input-output model of Section 3 may lead to a procedure for understanding the transfer mechanisms between various soil science parameters. By contrast, the Kriging procedure uses an assumed autocorrelation structure to smooth the data using a noiseless model.

The sampling procedure also puts restrictions on the kinds of methodology which can be applied to any given data set. For example, the state-space methodology in Section 2 would be limited practically to one-dimensional spatial series sampled along transects although there can be data points missing from the series. The stationary methods of Section 3 and the partial difference equation approach of Section 4 apply to the multidimensional case but require that data is equally spaced, either along a transect or over a grid. The stationary methods of Section 3 also require a reasonably large number of points in all directions and hence, are basically large sample techniques. By contrast, Kriging can be used when observations are sparse as well as irregularly observed in space or time.

6. Acknowledgements

I am indebted to Professor Don Nielsen of the Department of Land Air and Water Resources for introducing me, through his students, to research in the soil sciences. The paper was typed by Cathy DiSano of the Division of Statistics.

References

Akaike, H., 1974. A new look at statistical model identification. IEEE Trans. Aut. Control AC-19: 716-723.

Anderson, B.D.O. and J.B. Moore, 1979. Optimal Filtering. Prentice Hall, Englewood Cliffs.

Amoozegar-Fard, A., D.R. Nielsen and A. Warrick, 1982. Soil solute concentration distributions for spatially varying pore water velocities and apparent diffusion coefficients. Soil Sci. Soc. Am. J. 46: 3-9.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. J. Roy. Statist. Soc. B36: 192-236.

Besag, J.E. and P.A.P. Moran, 1975. On the estimation and testing of spatial interaction in Gaussian lattice process. Biometrika 62: 555.

Biggar, J.W. and D.R. Nielsen, 1976. Spatial variability of the leaching characteristics of a field soil. Water Resources Research 12: 78-84.

Bloomfield, P., 1976. Fourier Analysis of Time Series: An Introduction. Wiley, New York.

Box, G.E.P. and G.W. Jenkins, 1970. Time Series Analysis, Forecasting and Control. Holden Day, San Francisco.

Brillinger, D.R., 1979. Analysis of variance and problems under time series models. In: Analysis of Variance, Chapter 8, Editor: P.R. Krishnaiah. Handbook of Statistics, Vol. 1. North Holland, New York. 237-277.

Brillinger, D.R., 1975. Time Series Data Analysis and Theory. Holt, Reinhart and Winston, New York.

Cliff, A.D. and J.K. Ord, 1981. Spatial Processes, Chapter 6, Methuen, New York.

Cooley, J.W. and J.W. Tukey, 1965. An algorithm for the machine computation of complex Fourier series. Math. Comput. 19: 297-301.

Dempster, A.P., N.M. Laird and D.B. Rubin, 1977. Maximum likelihood estimation from incomplete data via the EM algorithm. J. Roy. Statist. Soc. B39: 1-38.

Erh, K.T., 1972. Applications of the spline function to soil science. Soil Sci. 114: 333-338.

Gelhar, L.W., 1974. Stochastic analysis of phreatic aquifers. Water Resources Research 10: 539-545.

Gutjahr, A.L., L.W. Gelhar, A.A. Bakr and J.R. McMillan, 1978. Stochastic analysis of spatial variability in subsurface flows, 2. Evaluation and application. Water Resources Research 14: 953-959.

Harvey, A.C., 1981. Time Series Models. Halsted Press, New York.

Jazwinski, A.H., 1970. Stochastic Processes and Filtering Theory. Academic Press, New York.
Journel, A.G. and C.J. Huijbregts, 1978. Mining Geostatistics. Academic Press, New York.
Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Trans. ASME J. of Basic Eng. 8: 35-45.
Kalman, R.E. and R.S. Bucy, 1961. New results in linear filtering and prediction problems. Trans. ASME J. of Basic Eng. 83: 95-108.
Kimball, B.A. 1974. Smoothing data with cubic splines. Agron. J. 68: 126-129.
Kitagawa, G., 1981. A nonstationary time series model and its fitting by a recursive technique. J. of Time Series Anal. 2: 103-116.
Kitagawa, G. and W. Gersch, 1984. A smoothness priors modeling of time series with trend and seasonality. J. of Am. Statist. Assoc. 79: 378-389.
Kolmogoruv, A.N., 1941. Interpolation und extrapolation von stationary zufallig folgen. Bull. Acad. Sci. USSR 5: 3-14.
Kunsch, H., 1982. Statistical analysis of uniformity trials, Research Report No. 32, Fachgruppe fuer Statistik, Eidgenoessische Technische Hochschule, 8092 Zurich, Switzerland.
Larimore, W.E., 1977. Statistical inference on random fields. Proc. IEEE: Special Issue on Multidimensional Systems 65: 961-970.
McBratney, A.B. and R. Webster, 1981. Detection of ridge and furrow pattern by spectral analysis of crop yield. International Statist. Rev. 49: 45-52.
Mechergui, M., 1984. Stochastic modeling of the water table in the vicinity of drainage tiles. Univ. of California, Davis, Ph.D. Dissertation.
Morkoc, F., J.W. Biggar, D.R. Nielsen and D.E. Rolston, 1984. Analysis of surface soil water content and surface soil temperature using a state-space approach. Submitted: Soil Sci. Soc. Am. J.
Matheron, G., 1963. Principles of geostatistics. Econ. Geology 58: 1246-1266.
Mercer, W.B. and A.D. Hall, 1911. The experimental error of field trials. J. Agric. Sci. 4: 107-132.
Nielsen, D.R., J.W. Biggar and P.J. Wierenga, 1982. Nitrogen transport in soils. Nitrogen in Agricultural Soils, Agron. Mon. No. 22, ASA, CSSA, SSSA. Madison, Wisconsin, 53711.
Nielsen, D.R., P.M. Tillotson and S.R. Vieira, 1983. Analyzing field-measured soil properties. Agriculture Water Management 6: 93-109.
Parzen, E., 1984. Time Series Analysis of Irregularly Spaced Data. Proceedings, College Station 1983. Springer-Verlag, New York.
Pfeifer, P.E. and S.J. Deutsch, 1980. Identification and interpretation of first-order space-time ARMA models. Technometrics 22: 35-47.

Priestley, M.B., 1981. Spectral Analysis and Time Series, Vol. 1: Univariate Time Series, Vol. 2: Multivariate Series, Prediction and Control. Reprinted in paperback, 1982, Academic Press, New York.

Ripley, B.B., 1981. Spatial Statistics. Wiley, New York.

Shumway, R.H. and W.C. Dean, 1968. Best linear unbiased estimation for multivariate stationary processes. Technometrics 10: 523-534.

Shumway, R.H., 1970. Applied regression and analysis of variance for stationary time series. J. of Am. Statist. Assoc. 65: 1527-1546.

Shumway, R.H., 1971. On detecting a signal in N stationarily correlated noise series. Technometrics 10: 523-534.

Shumway, R.H., 1982. Discriminant analysis for time series. In: Classification, pattern recognition and reduction of dimensionality, Chapter 1. Editor: P.R. Krishnaiah & L.N. Kanal. Handbook of Statistics, Vol. 1. North Holland, New York. 1-46.

Shumway, R.H. and D.S. Stoffer, 1982. An approach to time series smoothing and forecasting using the EM algorithm. J. of Time Series Anal. 3: 253-264.

Shumway, R.H., 1983. Replicated time series regression: An approach to signal estimation and detection. In: Handbook of Statistics, Vol. 3. Time Series in the Frequency Domain, Chapter 18. Editor: D. Brillinger & P.R. Krishnaiah. North Holland New York. 383-408.

Shumway, R.H., Y. Pawitan, R. Tai, and L. Tai, 1983. Multiple time series modeling of daily London mortality and associated weather and pollution effects. Final Report, Contract A1-154-33, California Air Resources Board.

Shumway, R.H., 1984. Some applications of the EM algorithm to analyzing incomplete time series data. In: Time Series Analysis of Irregularly Observed Data, Proceedings, College Station, 1983 (ed. E. Parzen), 290-324. Springer-Verlag, New York.

Solo, V., 1984. Modelling of 2D random fields by parametric cepstrum. In: Statistical Analysis of Time Series, Proceedings of U.S.A.-Japan Joint Seminar, 140-155. Institute of Statist. Math., Tokyo.

Valdin, M., S.R. Vieira, G. Vachaud and D.R. Nielsen, 1983. The use of cokriging with limited field soil observations. Soil Sci. Soc. Am. J. 47: 175-184.

Vieira, S.R., J.L. Hatfield, D.R. Nielsen, J.W. Biggar, 1983. Geostatistical theory and application to variability of some agronomical properties. Hilgardia 51: 1-75.

Vieira, S.R., D.R. Nielsen and J.W. Biggar, 1981. Spatial variability of field-measured infiltration rate. Soil Sci. Soc. Am. J. 45: 1040-1048.

Wecker, W.E. and C.F. Ansley, 1983. The signal extraction approach to nonlinear regression and spline smoothing. J. of Am. Statist. Assoc. 78: 81-89.

Whittle, P., 1953. The analysis of multiple stationary time series. J. Roy. Statist. Soc. B15: 125-139.

Whittle, P., 1954. On stationary processes in the plane. Biometrika 41: 434-449.
Whittle, P., 1963. Stochastic processes in several dimensions. Bull. Int. Statist. Inst. 40: 974-994.
Wiener, N., 1949. The Extrapolation, Interpolation and Smoothing of Stationary Time Series. Wiley, New York.
Wierenga, P.J., D.R. Nielsen, R. Horton and B. Kies, 1982. Tillage effects on soil temperature and thermal conductivity. Predicting tillage effects on soil physical properties and processes. ASA,SSA, Madison, Wisconsin, 53711.

Discussion

K. Cassel: How do you decide if a data set is stationary or non-stationary? What different options exist for converting a non-stationary data set into a stationary data set?

R. Shumway: A plot of the original data will sometimes exhibit non-stationary behavior due to trends or changes in the variance of the series. The autocorrelation function may decay very slowly if there are non-stationary trends. The time-varying power spetrum (cf. Priestly (1981)) may fluctuate if there are more subtle non-stationarities present. The methods for adjusting non-stationary data are (1) detrending using linear or non-linear regression, (2) differencing, (3) transformation (e.g., log), and (4) linear filtering (e.g., high-pass).

N. Safaya: How do you determine the size adequacy of the sampling grid or transect?

J. Hendrickx: What is the best kind of sampling design? Possibilities of interest are a grid with regular or random assignment of points and a grid with locations around it at smaller distances.

A. Warrick: With regard to necessary number of samples, could you distinguish between the problem of sample numbers for estimating a variogram and number of values for a given precision, given a variogram which is already determined? My own opinion is that, the second aspect is more easily addressed (cf. recent papers by Webster and colleagues) and the former is addressed primarily in a pragmatic fashion.

R. Turner: Soil solutions move through soils at irrigation intervals in response to irregular precipitation events. Is there a time series technique that would allow interpretations of lysimeter solution concentrations and flows collected at irregular intervals? How does one make interpretations of lysimeter data collected at regular intervals (e.g., monthly) when the solutions they collected actually were collected at irregular intervals?

R.H. Shumway: An adequate grid should cover a large enough area to capture the phenomenon of interest. The

size of the mesh should be such that no important periodic variations or frequencies in the data are distorted or aliased by the sampling scheme. Again, it is important to match the sampling interval with the characteristic soil parameters of interest. Time domain methods require at least 30 points along a transect. Frequency domain techniques require at least 50 point spaced uniformly along a transect or on a 50x50 grid.

I am opposed to random sampling of design points whether they are located on a grid or not. For example, the selection of design points determines the variance of kriging estimators and it makes sense to choose these systematically to reduce the mean square error. Of course, if nothing is known about the variability of the soil measurements over the region of interest, it may make sense to construct several uniform grids or transects at different spacings.

If one is willing to assume a form for the autocovariance or variogram, the configuration of observed data completely determines the mean square error of the predicted value at each point on the surface. The determination of a sampling scheme for estimating the variogram depends on the "pragmatic" considerations mentioned above plus the parametric restrictions which one is willing to assume for the variogram (e.g., isotropic, exponential decay, etc.).

Time series methodology generally requires that observations be regularly spaced although some progress is being made on the irregular case using the state-space model (cf. Parzan (1984) and Sections 2 and 5 of my paper).

R. Kachanoski: Is the state-space model (Kalman filter) a model for a non-stationary dynamic series and not a stationary series?

R. Shumway: The state-space model can be either stationary or non-stationary depending on the structure of the transition matrix ΦD in equation (2) of my paper. The example in Section 2 is a non-stationary model. A simple example of a stationary model would be to choose p=1 and $|\Phi|<1$ in (2), in which case the signal process is simple first-order autoregression with an exponentially decaying autocorrelation function.

A number of examples of various stationary and non-stationary state-space models can be found in Harvey (1981).

W. Jury: Do you see any way in which state-space analysis could be used in conjunction with a variogram analysis to help determine drift components which may be biasing the variogram.

R. Shumway: The example in Section (2) has a drift component which could be eliminated by using the residuals (observed minus predicted) to do the variogram analysis.

C. Topp: Noise or observation error and uncertainty arise from both the method of measurement and from the variability of the signal. What effect does the magnitude of measurement error have on the application of state-space or time series analysis vs. kriging?

R. Shumway: The state-space model provides separately for estimating both the variance of the measurement error, R, in Equation (1) and the signal or model error, Q, in Equation (2). Values for these two components for some soil data are given in Table 1 of Section 2.

A. Gutjahr: Is there an order in 2 or 3 dimensions which could allow one to use "submatrices" in the Kalman Filter?

R.H. Shumway: Unfortunately, there are no convenient Kalman recursions for more than one dimension. The answer in the higher dimensional case probably involves sampling over a regular grid and transforming to the frequency domain (see the discussion below Equation (5) of my paper).

D. Myers: The support of the samples and the support of the region for which an average is determined can be incorporated into the kriging equation. How are these incorporated into the state-space modelling?

R. Shumway: This can be done using the "measurement" matrix, M(s), in the observation Equation (1) of my paper. For example, in Equation (a), one might choose $M(s) = (1,1)'$ if the observation y(s) is really an average.

G. Holmoren: Does bootstrapping have any possible application to reducing the number of samples?

R. Shumway: Bootstrapping is a resampling method for estimating the variance of an estimator. Its primary purpose would not be for reducing the sample size requirements. A discussion of the "bootstrap" as well as the related "jacknife" and "cross validation" techniques is in Diaconis, P. and B. Efron (1983), Computer-Intensive Methods in Statistics, Scientific American. May, 116-130.

D. Mulla: Is there any statistical information that can be obtained from the imaginary part of the Fourier transform?

R. Shumway: The imaginary part of the transform is used to determine the phase, which gives information about the relative arrival times of different frequencies on a single record or about the lag relations of different frequencies when there are two records. A good discussion of some of these questions can be found in Bloomfield (1976) or in Priestley (1981).

Spatial variability of soil properties

Goro Uehara, B. B. Trangmar and R. S. Yost, Department of Agronomy and Soil Science, University of Hawaii, Honolulu, Hawaii, USA.

Introduction

One aim of soil science is to establish the cause and effect relationship between soil properties and soil behavior so that users of soil resources can predict the performance and behavior of soils. The prediction is made by matching the requirements of a specific use to the characteristics of the soil. To rectify mismatches in spatially variable soils, it is necessary to know the number, magnitude, and whereabouts of the mismatches. It is the inability to deal with spatial variability that prevents soil users from accurately matching soil use requirements to soil characteristics and, therefore, from predicting soil performance and behavior.

The purpose of this paper is to illustrate how geostatistics can be used to deal with soil spatial variability.

Sources of data

Four sets of data were used to illustrate soil spatial variability. Information about the data sets is summarized in Table 1. The areal extent of the four study areas ranged from 2.6 million hectares in Rwanda, Africa, to an experimental plot of less than one-tenth hectare in area in Sitiung, West Sumatra. The closest sampling distance was nearly four kilometers in Rwanda and about one meter apart in the experimental plot. Sampling in the small experimental plot was designed for geostatistical analysis, but in the three other locations sampling was made at a time when geostatistics was still unknown to the individuals collecting the samples. It is likely that many similar data sets exist today in files of soil survey institutes around the world. They represent a rich source of information that can be used to study soil spatial variability.

Data analyses

The four data sets described in Table 1 are analyzed separately to illustrate how scale, sampling distance, sampling direction, and kinds of soil affect the range, sill, and nugget variance of the semi-variogram, and therefore the kriged results.

Table 1. Sampling location and characteristics of the data sets

Sampling location	Size of study area (hectares)	Closest sampling distance (meters)	Sampling pattern	Number of sampling sites
Rwanda, Africa	2,633,800	3700	random	119
Kenana, Sudan	34,318	1000	grid	254
Sitiung, West Sumatra	106,650	400	random	88
Sitiung, West Sumatra	0.078	1.0	grid	137

Analysis of Rwanda data set

Between 1979 and 1980, Peter Vander Zaag, a potato agronomist with the International Potato Center stationed in Rwanda, collected soil samples from 0-15 cm and 30-45 cm depths at intervals of several kilometers along the major highways of the country, as shown in Figure 1. His aim was to obtain an overall picture of the fertility of the country's soils. Fortunately, the sample location was clearly marked on a road map.

Semi-variograms of soil pH, soil calcium, soluble silicon, and extractable ammonia nitrogen are shown in Figures 2-5. What is most striking about the semi-variograms is the long range over which the soil properties are related. The results were surprising because Campbell (1978), for example, obtained a pure nugget effect for soil pH collected 10 m apart. Table 2 summarizes the main features of the variograms for pH, exchangeable calcium (me/100g), effective cation exchange capacity (ECEC) (me/100g), extractable silica (ppm), and extractable ammonia nitrogen (ppm). The Rwanda data summarized in Table 2 suggest that soil properties are spatially related over long distances. A more detailed account of the Rwanda data set is contained in a paper by Vander Zaag et al. (1981).

Analysis of Kenana, Sudan data set

When the result of the Rwanda data set was presented at the Fourth International Soil Classification Workshop in June 1981 in Kigali, Rwanda, a Sudanese soil scientist recommended that soil data collected to assess land for sugarcane production in Kenana, Sudan, be analyzed geostatistically. The data set consisted of

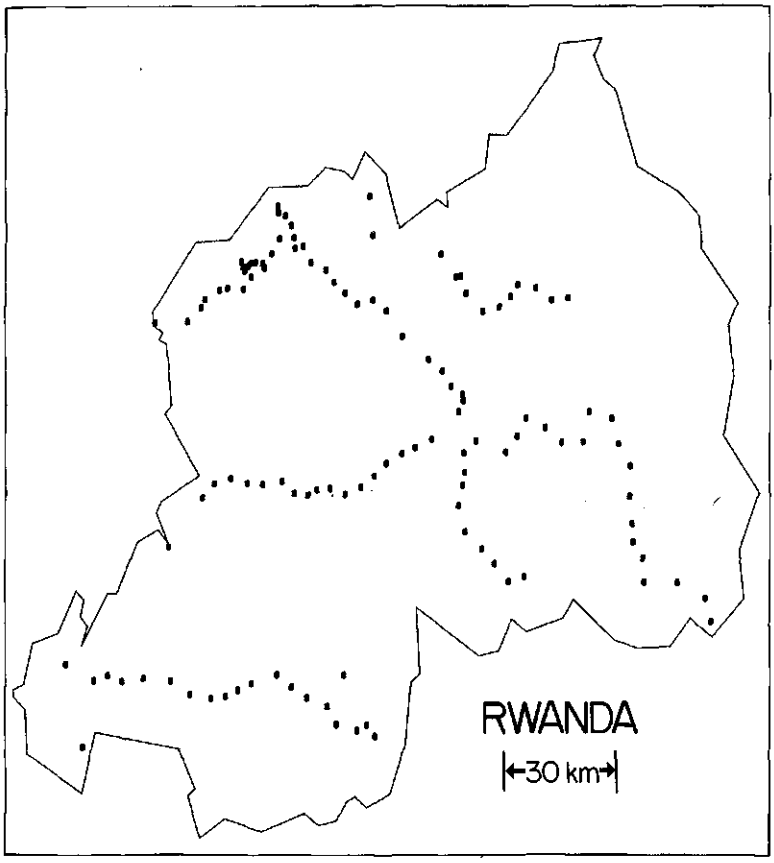


Fig. 1 Location of sampling sites.

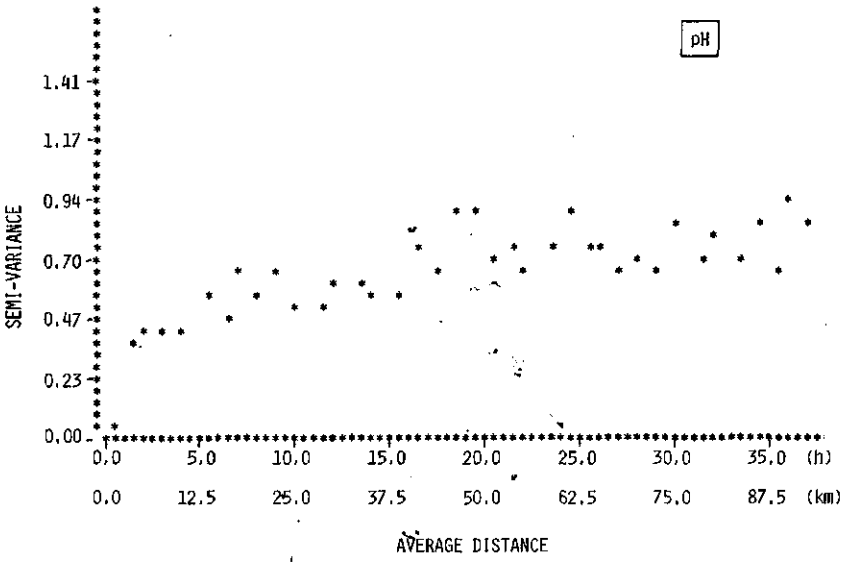


Fig. 2 Isotropic semi-variogram of soil pH.

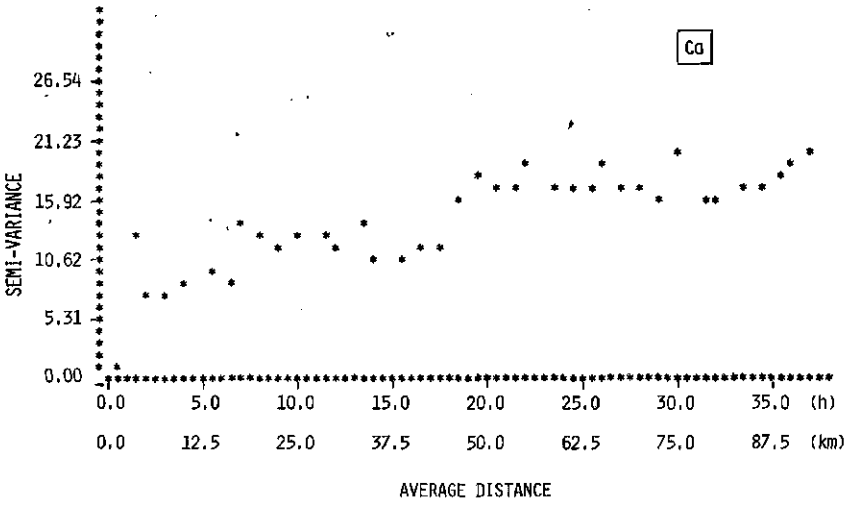


Fig. 3 Isotropic semi-variogram for exchangeable Ca.

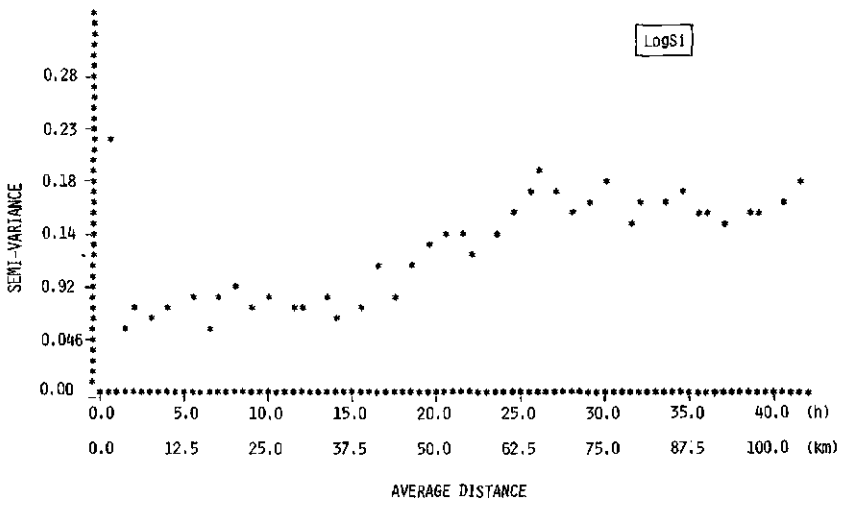


Fig. 4 Isotropic semi-variogram for log Si.

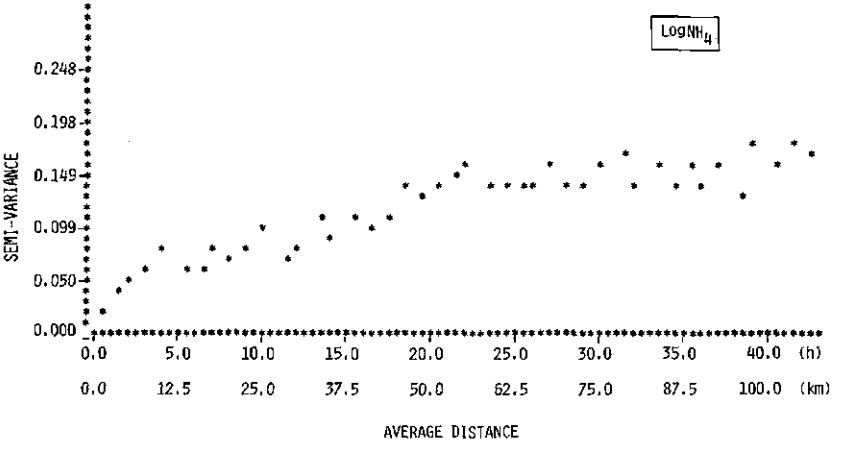


Fig. 5 Isotropic semi-variogram for log NH_4.

Table 2. Summary of variograms. The re-expressed mean and variance for Si and NH_4 were calculated from the log data using the equations: $M = \exp(\mu_y + \sigma_y^2/2)$ and $V = \mu_x^2[\exp(\sigma_y^2) - 1]$ from Haan (1977), where M is the re-expressed mean, μ_y is the general mean of the data logarithms, σ_y^2 is the general variance of the data logarithms, V is the re-expressed variance, and μ_x is the mean of the original data values.

Variable	Intercept (C_o)	Range h	Range km	General mean	General variance	Re-expressed mean (M)	Re-expressed variance (V)
pH	0.0	15	37.5	5.71	0.783	--------	--------
Ca	6.29	20	50	7.14	17.69	--------	--------
ECEC	14.61	20	50	11.35	31.30	--------	--------
log Si	0.0198	25	62.5	2.22	0.153	9.96	16.4
log NH_4	0.0296	22	55	4.41	0.165	89.6	1430

auger samples collected from the 0-25 cm and 75-100 cm depths at 254 locations on a one-kilometer by one-kilometer grid (Figure 6). At 18 of the 254 sampling locations, a pit was excavated next to the auger hole and sampled for complete soil characterization data in order to classify the soils.

As indicated in the Soil Survey Report (Adam, 1976), exchangeable sodium percentage (ESP) was judged to be the soil property most likely to limit sugarcane performance in the Kenana area. Taking this cue, the ESP was examined in two ways. First, we wanted to know if there was structure in the variance of ESP, and second, we wanted to know if information about the spatial distribution of ESP could have been obtained with fewer samples.

The semi-variogram (Figure 7) indicates structure in the variance but also suggests that considerable more information about the structure could have been obtained by collecting samples between zero and one-kilometer distance.

In order to obtain an estimate of the nugget variance, the semivariance of the 18 pairs of adjacent auger and pit samples was calculated. Since the auger and pit samples were collected a few meters apart, the semivariance appears as the nugget variance in Figure 7. In Table 3, the exchangeable sodium percentages of adjacent auger and pit samples are compared to the kriged value for the same location.

As a second exercise, the semi-variogram and kriged map of ESP were constructed using 100, 75, 55, and 47 percent of the data points. The purpose of this exercise was to ascertain the extent to which kriging would compensate for reduced sample size.

The sample number was first reduced by deleting every third row of data points in the grid as shown in Figure 8. The semi-variogram constructed from 75 percent of the data is shown in Figure 9.

The sample size was further reduced to 56 percent of the original by deleting every third row and column from the data set except those on the borders as shown in Figure 10. The semi-variogram for this data set is shown in Figure 11.

The effect of reducing sampling intensity can be judged by comparing the map based on 100 percent of the measured data (Figure 12) against the kriged maps constructed from 254 (100 percent), 192 (75 percent), and 143 (56 percent) of the data as shown in Figures 12, 13, 14, and 15. The estimation variance resulting from reducing sample size is summarized in Table 4. A more complete analysis of the data is presented by Trangmar et al. (1982).

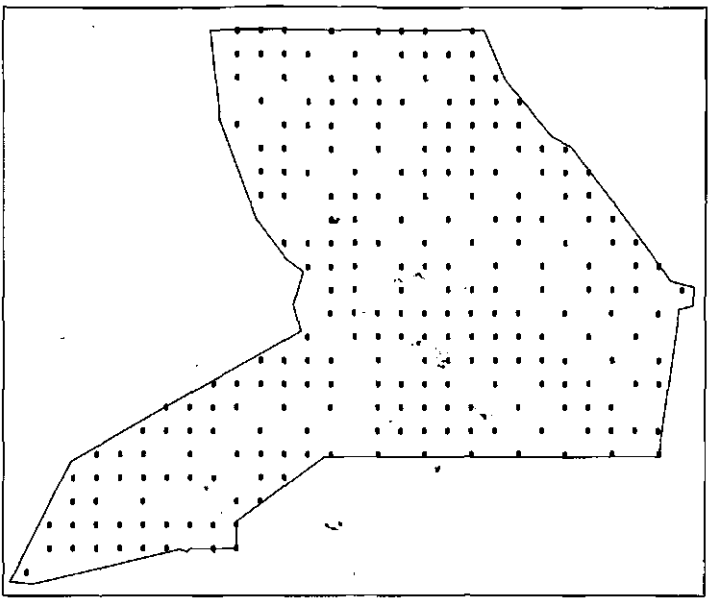


Fig. 6 Location of soil sampling sites, Kenana Sugar Project.

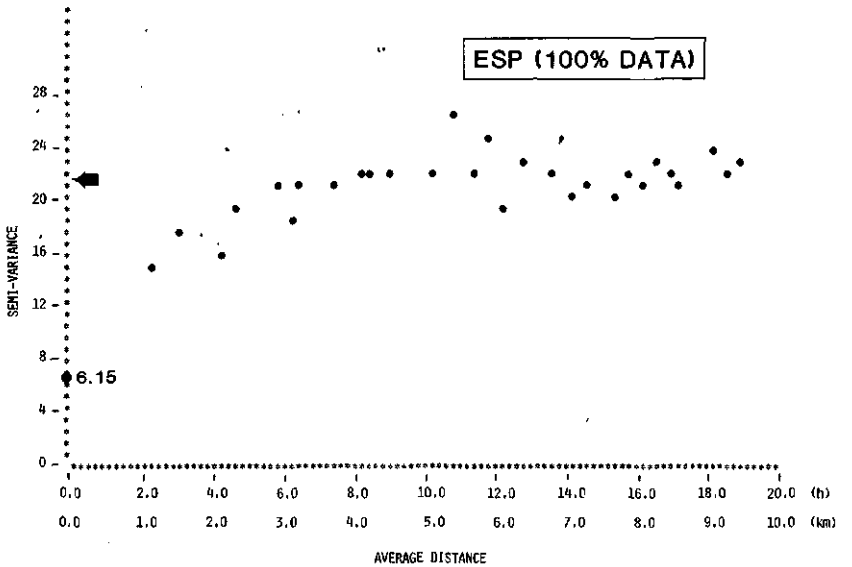


Fig. 7 Semi-variogram for ESP determined from 100% (254) of samples.

Table 3. Comparison of exchangeable sodium percentage (0-25 cm soil depth) from adjacent pit and auger samples with kriged values for the same sites. Pit and auger values derived from Table 8, Adam 1976.

Sample site	Pit	Auger	Kriged	Estimation variance
KSPA 01	19.4	13	11.0	11.4
02	14	11	10.7	16.6
03	14.8	8	8.8	12.3
04	9	12	6.9	12.1
05	12	6	11.1	13.2
06	7.2	7	10.4	11.9
07	6.2	7	6.1	11.8
08	11.2	8	9.7	11.4
09	1	24	13.5	13.8
10	2	4	4.3	12.3
11	6.4	6	8.1	11.4
12	14.8	8	8.6	12.4
13	11.6	11	9.9	12.3
14	12.8	13	10.7	11.4
15	9.6	10	10.5	11.4
16	10.4	9	10.6	11.5
17	13	11	10.2	13.0
18	10	10	13.6	12.5

Table 4. Effect of reducing sample size on the semi-variogram and mean estimation variance of kriged values.

Percent of samples used in analysis (n = 254)	Semi-variogram		Mean estimation variance
	Intercept	Sill	
100	6.15	21.6	10.5
75	6.15	22.1	12.2
56	6.15	24.0	13.1

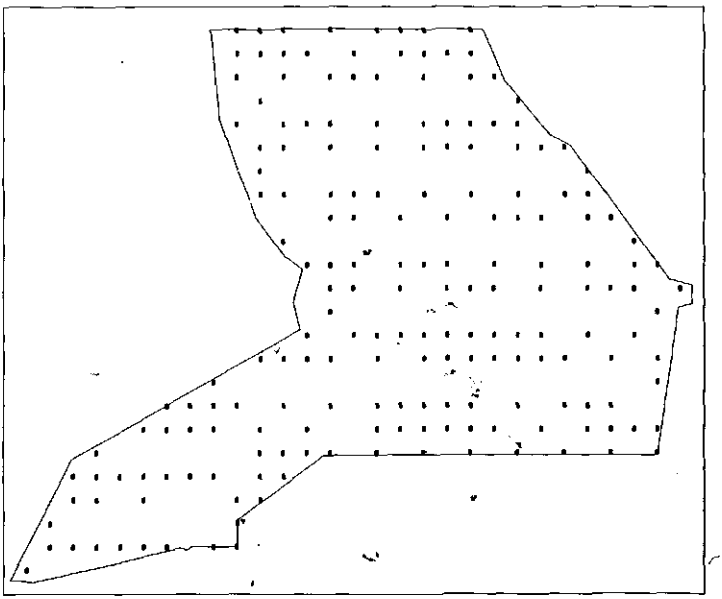


Fig. 8 Location of sampling sites for calculations based on 75% of data collected.

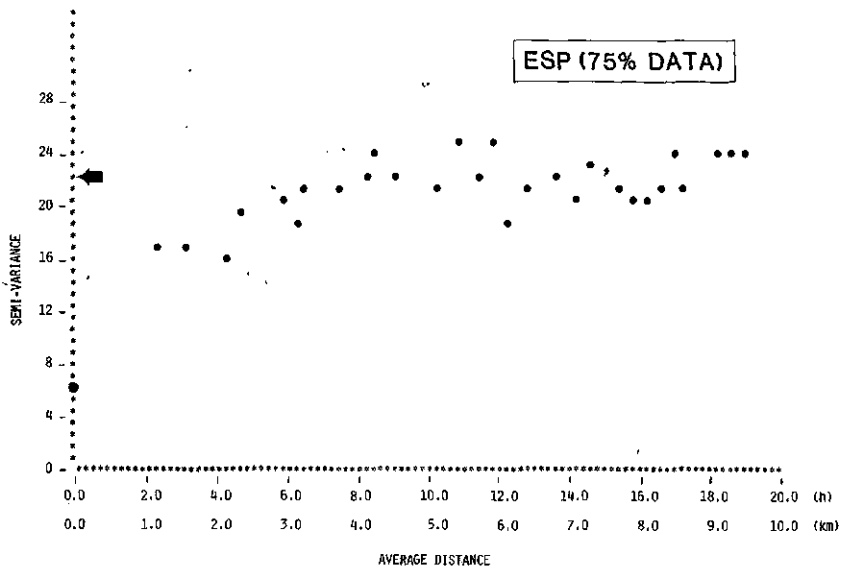


Fig. 9 Semi-variogram for ESP determined from 75% (192) of samples.

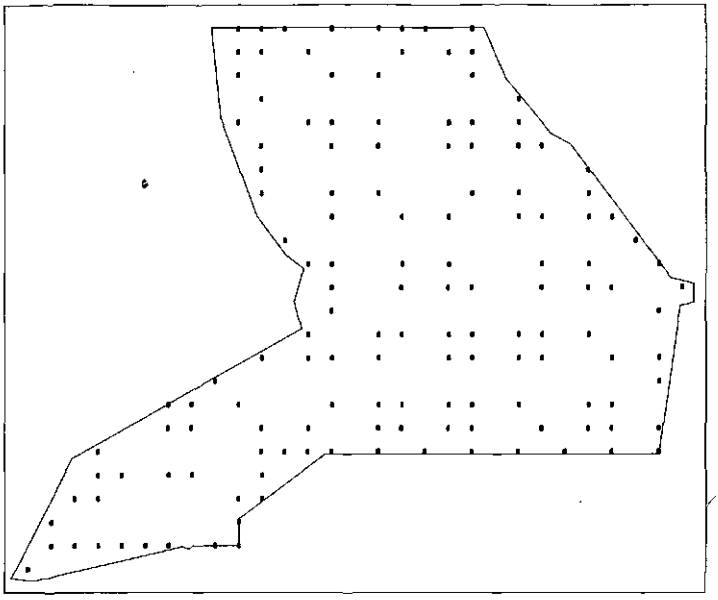


Fig. 10 Location of sampling sites for calculations based on 56% of data collected.

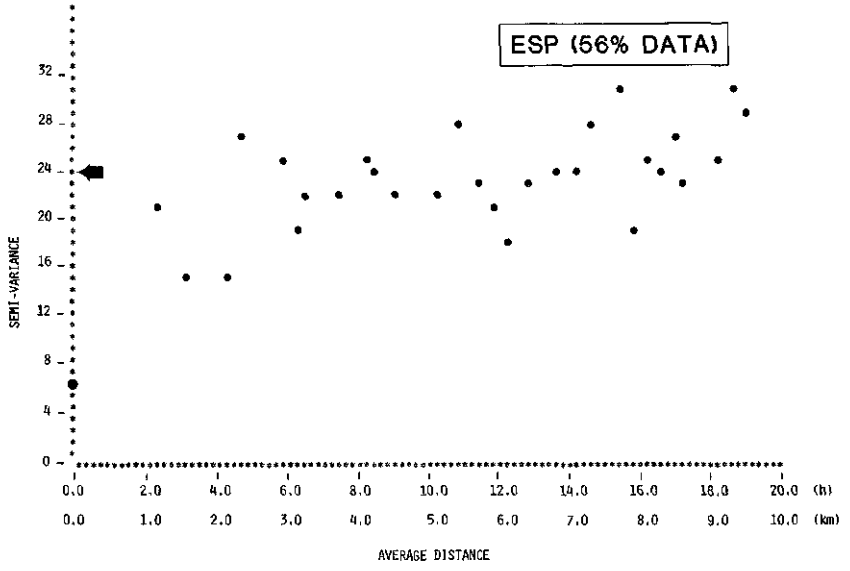


Fig. 11 Semi-variogram for ESP determined from 56% (143) of samples.

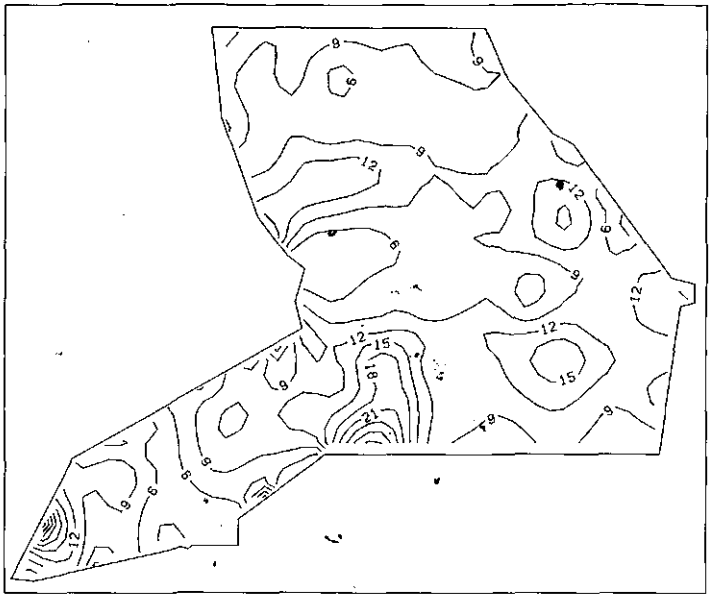


Fig. 12 Map of ESP determined from 254 sample values, Kenana Sugar Project.

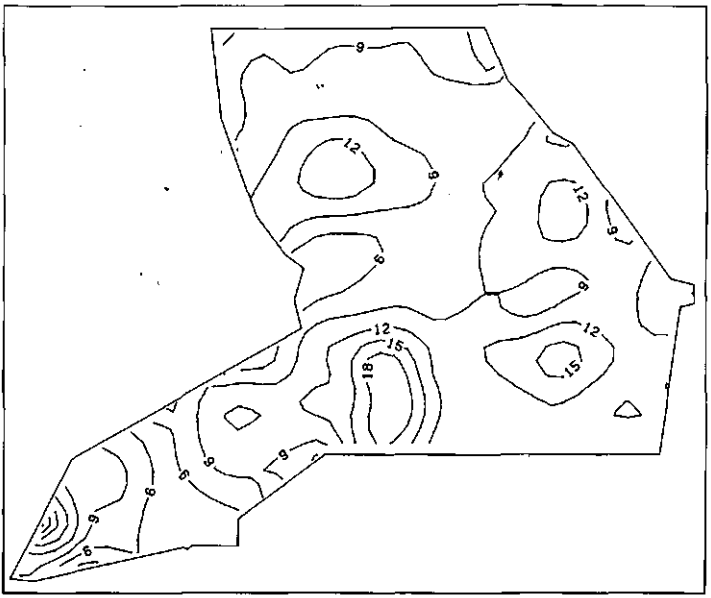


Fig. 13 Map of kriged ESP values determined from 100% (254) of samples.

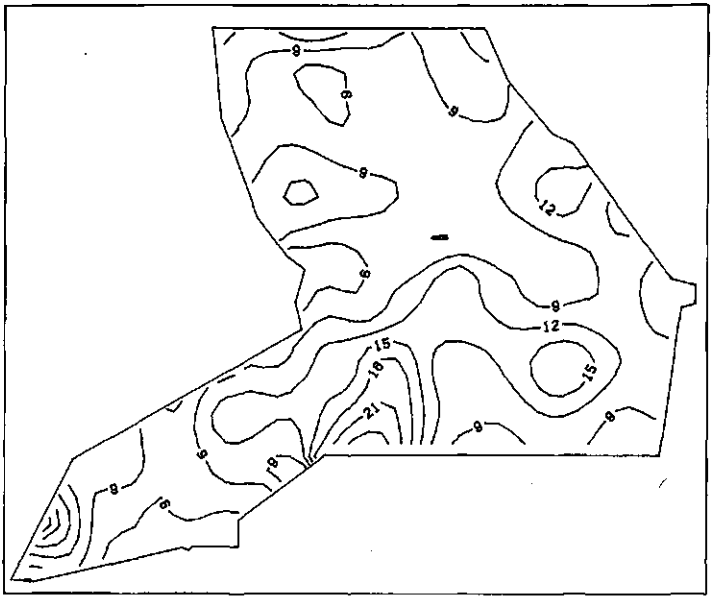


Fig. 14 Map of kriged ESP values determined from 75% (192) of samples.

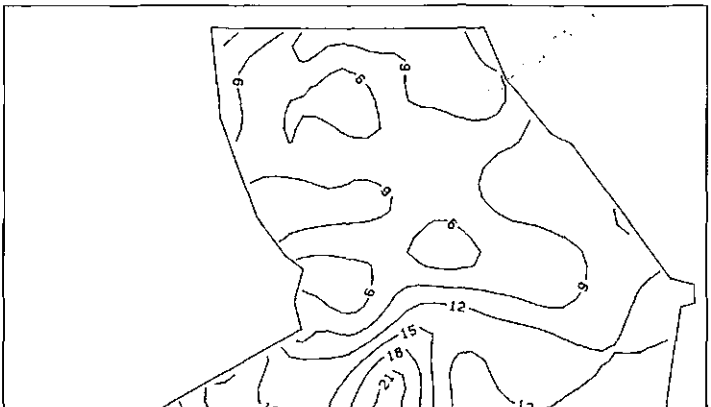


Fig. 15 Map of kriged ESP values determined from 56% (143) of samples.

Analysis of the Sitiung data set

Sitiung serves as a settlement area for an Indonesian transmigration program. Several thousand transmigrant families from Java have settled in a 100,000 hectare area along the Batanghari River. The soils of the region range in quality from moderately fertile Inceptisols on the subrecent terraces of the river to the highly leached and impoverished Oxisols and Ultisols of the dissected peneplain to the southwest. Mean annual rainfall is 2800 mm and mean annual air temperature is 26 Celcius.

In preparation for the arrival of the transmigrants, complete soil characterization data were collected from 88 locations in the 100,000 hectare area. A map of the area showing sampling locations is shown in Figure 16, and a summary of statistics of soil properties subjected to geostatistical analysis is given in Table 5.

The semi-variograms for sand content and pH illustrate the degree to which soil properties in the Sitiung area are spatially related (Figures 17 and 18). In order to display the spatial variability of soil properties on a map, each property was kriged for 268 locations as shown in Figure 19.

Since soil phosphorus levels were low in this area, the Indonesian scientists had measured $NaHCO_3$ extractable phosphorus in 52 locations in the area. They had also measured phosphorus extracted by hydrochloric acid as a parameter for soil classification. In order to assess the spatial distribution of phosphorus deficient soils, a map of the $NaHCO_3$ extractable phosphorus was prepared. An attempt was made to increase the accuracy of the map by co-kriging. The cross- and auto-semivariogram for phosphorus extracted by $NaHCO_3$ and HCl is shown in Figure 20. The improvement in the map of the $NaHCO_3$ extractable phosphorus is shown in Figure 21 by displaying the reduction in the estimation variance. Co-kriging resulted in the greatest reduction in the estimation variance where the sampling density of $NaHCO_3$ extractable phosphorus was lowest.

This data set also enabled anisotropic kriging to be computed for sand content. The semi-variograms for sand computed in four directions are shown in Figure 22. Anisotropic kriging resulted in marked reduction in the estimation variance as shown in Figure 23. Details of the co- and anisotropic kriging are presented in a dissertation submitted to the Graduate Division of the University of Hawaii (Trangmar, 1984).

Analysis of data from a small experimental plot in Sitiung

The spatial variability of soil quality in a typical farm in Sitiung is marked by differences in plant growth that range from bare spots to lush green strips. The bare spots are sterile subsoil exposed by the bulldozer, and the green strips correspond

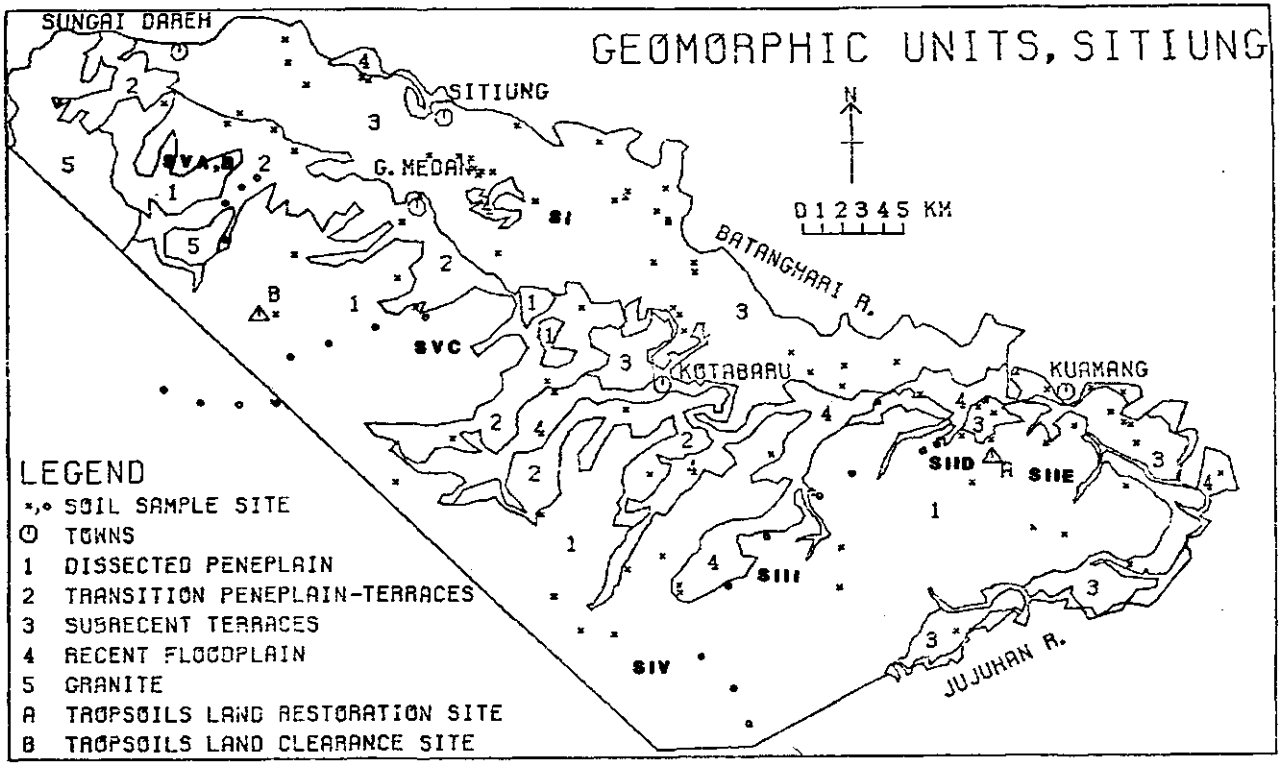


Fig. 16 Main geomorphic units and sample locations, Sitiung. SI-SVC marks location of transmigration settlements.

Table 5. Mean, range, variance and number of samples for soil properties analysed using geostatistics, Sitiung region.

Soil property	Depth (cm)	Mean[a]	Range	Variance	Number of samples
Sand (%)	0-15	16	0-59	213.9	108
	15-45	14	0-64	157.8	109
Silt (%)	0-15	26	4-69	213.5	109
	15-45	23	2-67	181.1	109
Clay (%)	0-15	58	17-90	201.4	109
	15-45	64	32-90	149.5	109
pH	0-15	4.5	3.7-6.0	0.19	109
	15-45	4.7	4.0-6.3	0.15	109
Organic C (%)	0-15	3.14	0.37-20.43	2.60	109
	15-45	1.29	0.37-21.89	0.39	109
Total N (%)	0-15	0.25	0.07-1.20	0.01	107
	15-45	0.13	0.04-1.24	0.01	107
HCl-P (ppm)	0-15	107	4-391	8645.8	107
	15-45	67	6-335	5018.6	107
$NaHCO_3$-P (ppm)	0-15	8	2-33	34.6	52
	15-45	2	0-8	5.3	44
Exch. cations (meq/100g)					
Ca	0-15	0.6	0.1-23.0	2.16	109
	15-45	0.3	0.1-8.8	0.20	108
Mg	0-15	0.3	0.1-3.1	0.12	109
	15-45	0.2	0.1-1.9	0.03	108
K	0-15	0.2	0.0-1.2	0.02	109
	15-45	0.1	0.1-0.7	0.01	108
Sum of cations (meq/100g)	0-15	1.3	0.3-27.4	2.17	109
	15-45	0.8	0.3-10.6	0.37	108
Exch.Al (meq/100g)	0-15	3.6	0-9.3	3.43	99
	15-45	3.3	0-9.9	3.34	99
ECEC (meq/100g)	0-15	5.1	0.8-27.4	6.50	99
	15-56	4.3	0.6-11.0	5.08	98
Al satn. (%)	0-15	73	0-94	504.2	99
	15-45	78	0-96	332.5	98
Cu (ppm)	0-15	3	1-8	4.04	56
	15-45	3	1-8	3.78	54
Zn (ppm)	0-15	5	1-55	47.9	56
	15-45	5	0-30	55.3	54
Lime req. (tonnes/ha)	0-15	4.0	0-10.4	4.34	99

a. Means and variances for organic C, N, HCl-P, $NaHCO_3$-P, exch. Ca, Mg, sum of cations, ECEC, Cu and Zn were determined on log transformed values and re-expressed in terms of the original data using equations of Haan (1977).

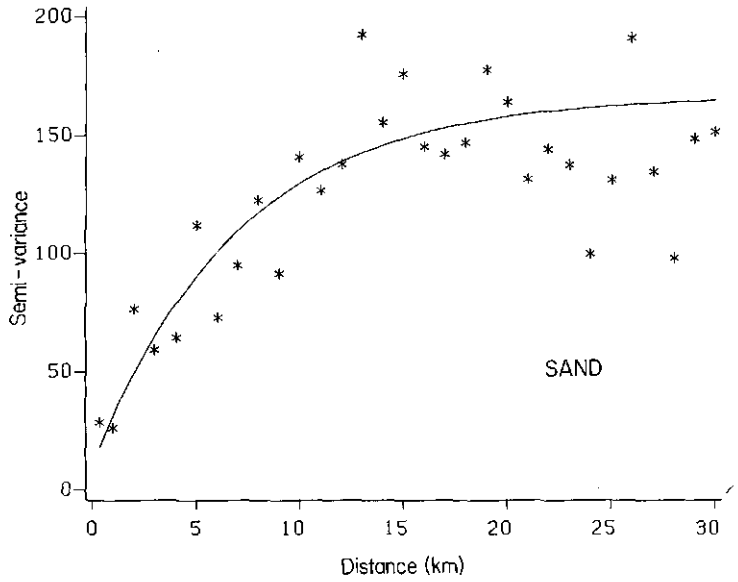


Fig. 17 Isotropic semi-variogram for sand content, 0-15 cm depth; Sitiung.

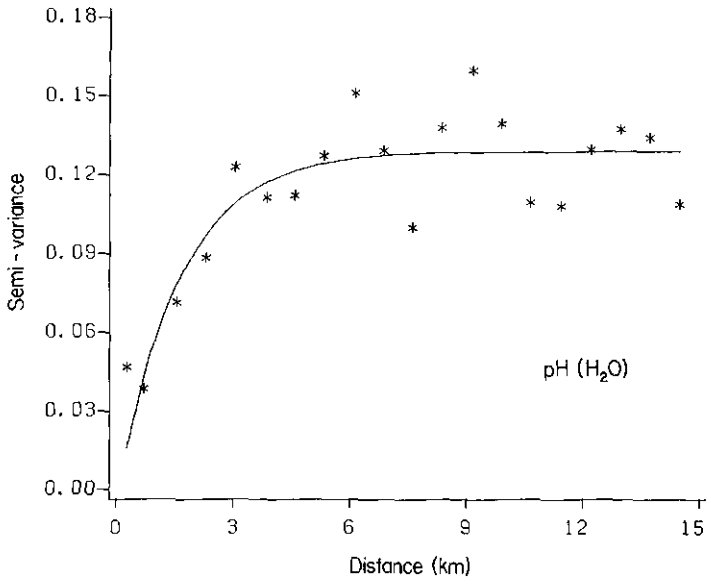


Fig. 18 Isotropic semi-variogram for pH(H_2O), 0-15 cm depth; Sitiung.

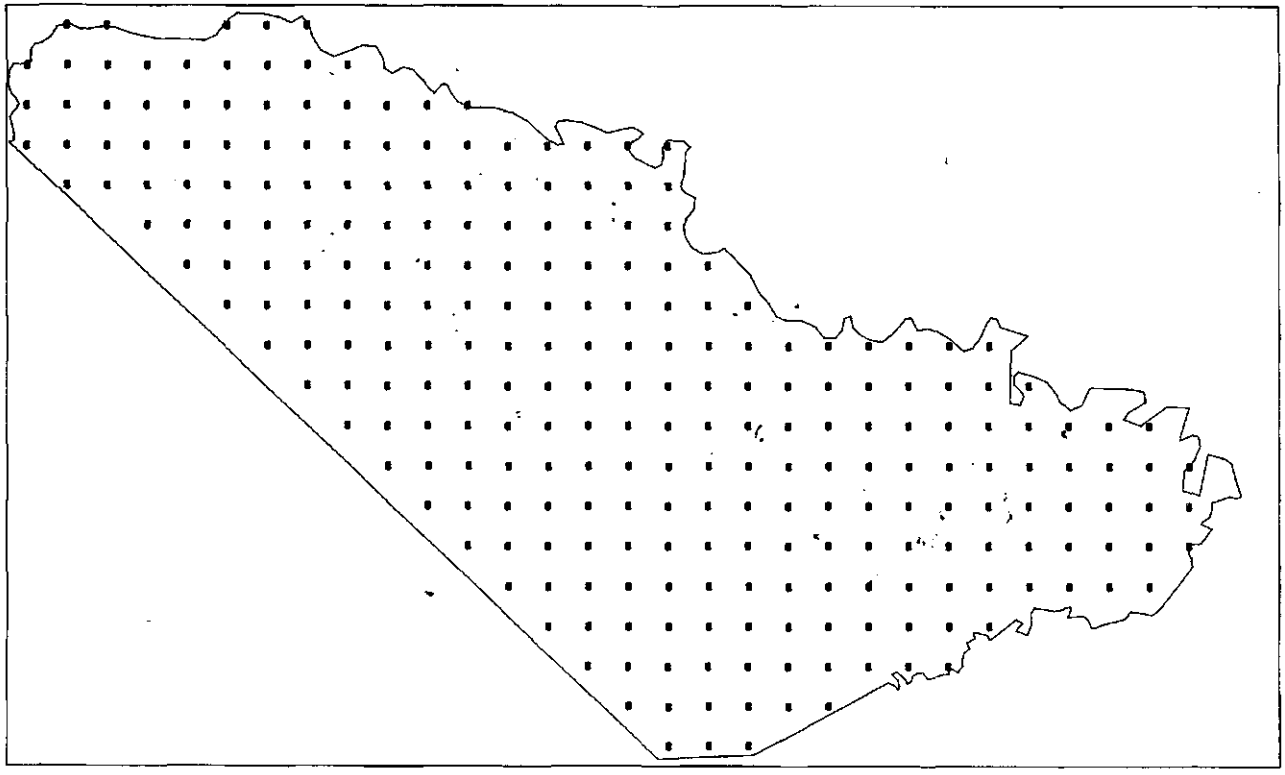


Fig. 19 Location of 268 kriged points; Sitiung. Each point indicates centroid of a 4 sq. km cell.

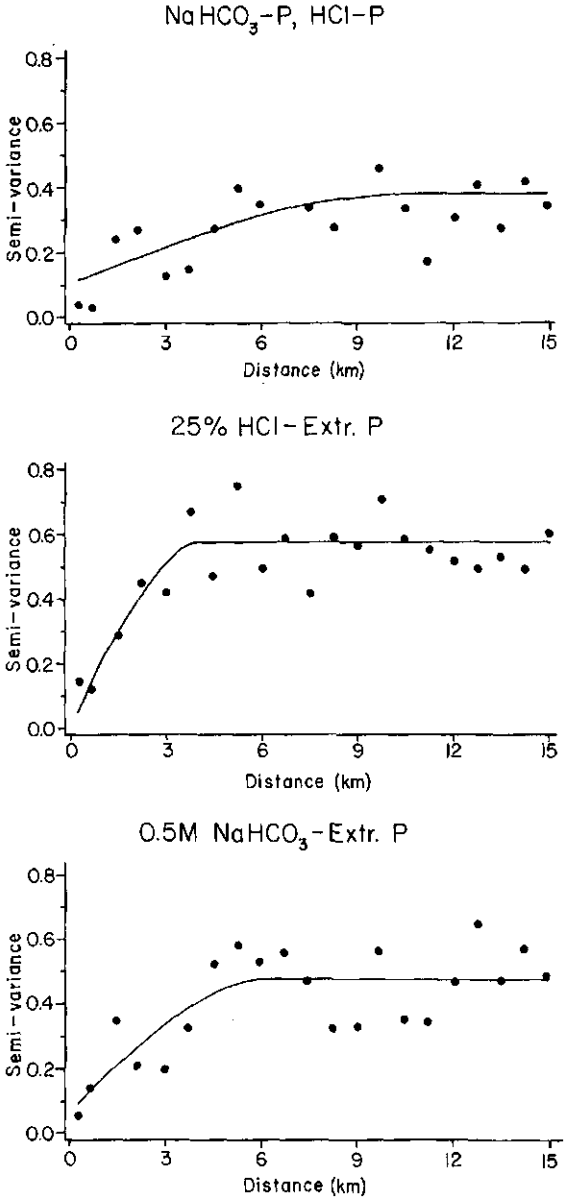


Fig. 20 Isotropic cross- and auto-semi-variograms for log transformed values (ppm) of 25% HCl-P and 0.5M $NaHCO_3$-P, 0-15 cm depth, Sitiung.

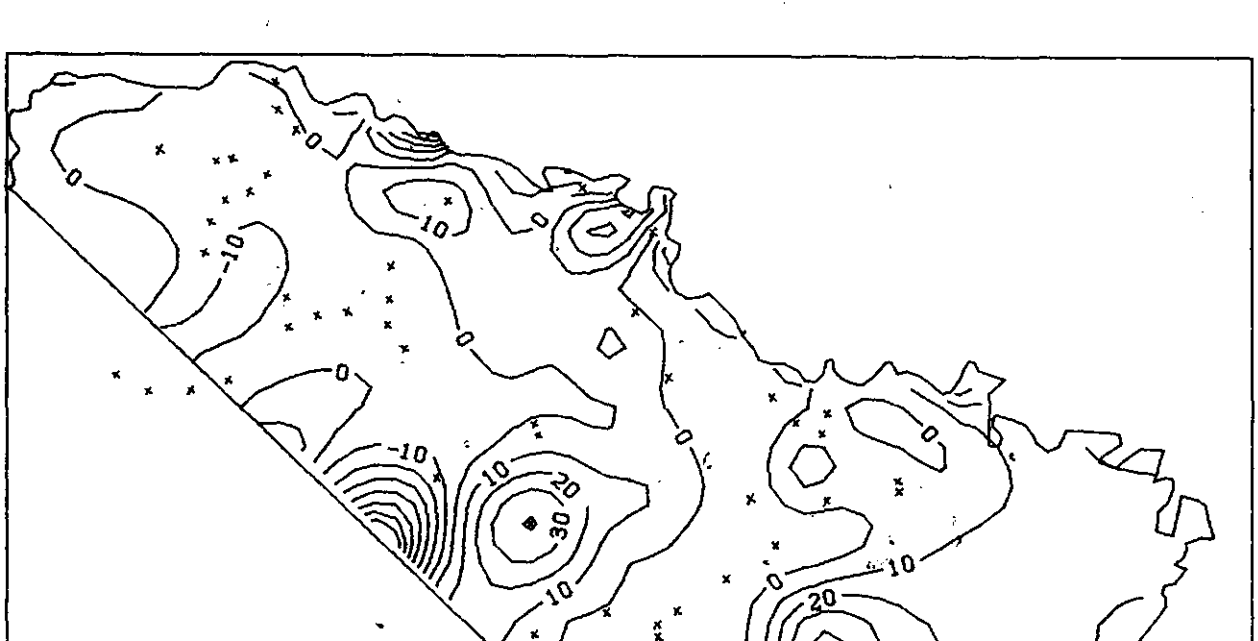


Fig. 21 Percent reduction in estimation variance by co-kriging topsoil 0.5M $NaHCO_3$-extr. P relative to auto kriging. Crosses indicate sample locations of $NaHCO_3$-P.

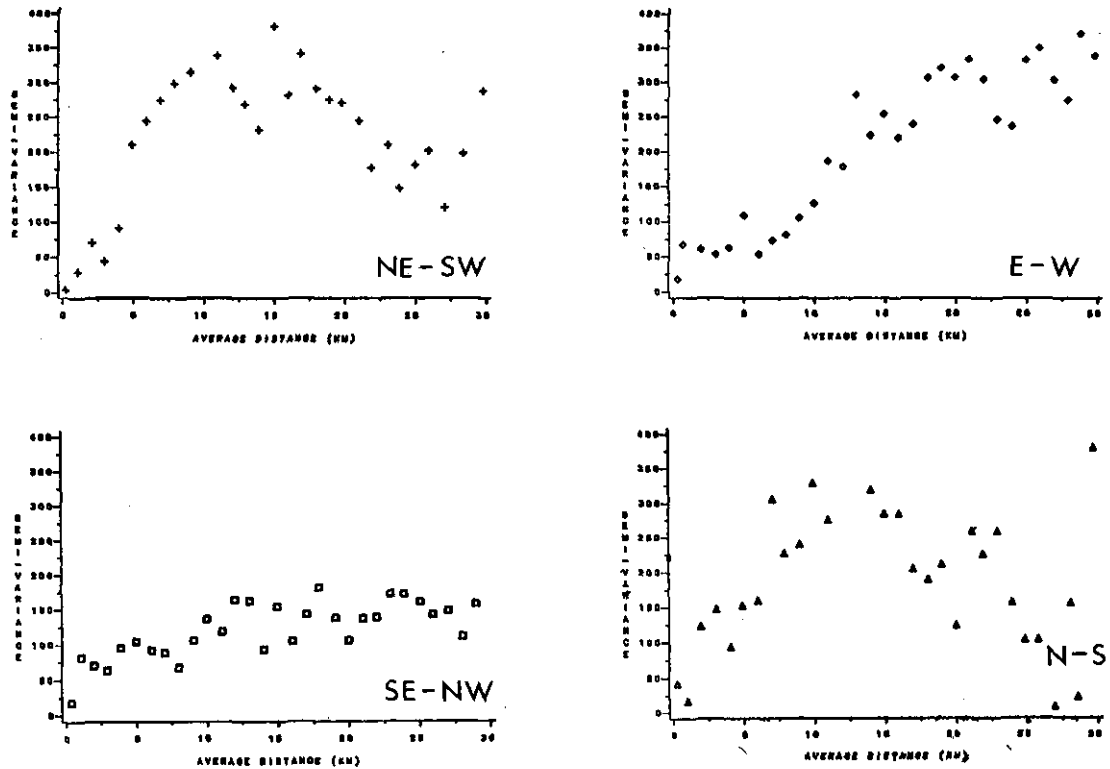


Fig. 22 Semi-variograms for sand calculated in the NE-SW, E-W, SE-NW and N-S sectors, 0-15 cm depth, Sitiung.

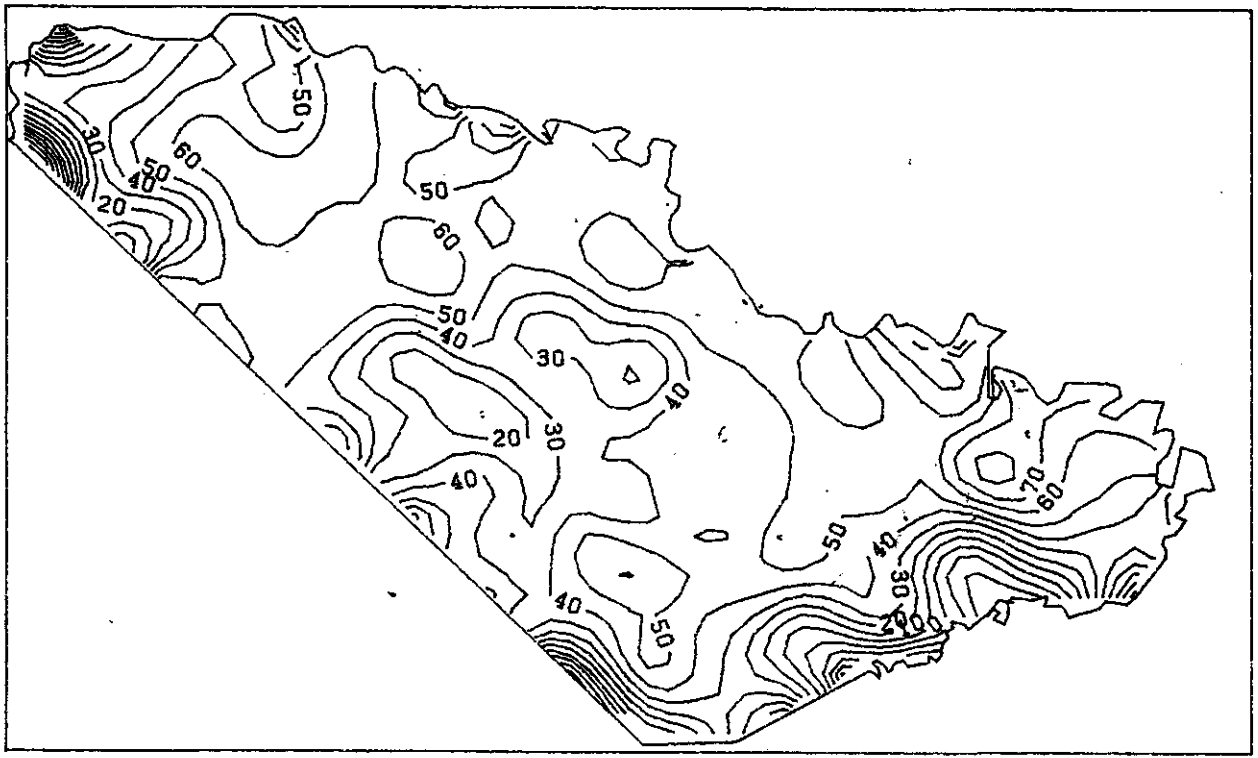


Fig. 23 Percent reduction in estimation variance for topsoil sand obtained by anistropic punctual kriging relative to that obtained by isotropic punctual kriging, Sitiung.

to the ash lines of burnt trees. This variability created during the land clearing operation and subsequent burning presents serious problems for the farmer and researcher. In farmers' fields, the bare spots are most prone to erosion. Bare spots produce nothing and therefore are neglected. Erosion feeds on neglect and the land is eventually abandoned. The researcher needs variability, but wants a field with minimum natural variability so that the effects of the treatment variable will be clearly expressed. In Sitiung the range of natural soil variability is almost always as large as the imposed treatment range. The aluminum saturation, for example, ranges from zero to 90 percent. This is the range a researcher would impose in a liming experiment to assess the effect of aluminum toxicity on crop performance.

To assess the short range variability in a small plot, sampling was conducted on a 28x28 meter plot in the manner indicated in Figure 24. A picture of the visible kinds of variability encountered in the plot is shown in Figure 25.

High soil acidity and the resulting release of toxic aluminum severely constrains crop performance in Sitiung. Isotropic semivariograms for aluminum saturation and soil pH are shown in Figure 26. A three dimensional diagram of aluminum saturation for the plot is shown in Figure 27. The effect of aluminum toxicity is reflected in grain yield of rice as shown in Figure 28. Semi-variograms of plant height at 60 days and grain yield (Figure 29) show that the range of spatial dependence for crop parameters are longer than for soil properties that are thought to affect crop performance.

Table 6 summarizes the significance of selected soil properties on crop parameters, and Table 7 shows the effect of various terrain units on crop performance. These results show that natural soil variability can be exploited to answer agronomic questions.

Summary and conclusion

Many commonly measured soil properties are spatially related for long distances. The same property measured a meter apart or one kilometer apart shows structure in the variance but shows different ranges of spatial dependence for different scales. It is likely that different kinds of heterogeneity are sampled as the sampling scale is varied. Different soil forming factors operating simultaneously at a point with different degrees of intensity over differing scales may be the cause of the scale-dependent spatial variability.

Geostatistics is a useful tool to study soil genesis and can also be used to develop special purpose maps for specific soil interpretation needs. For example, a map showing the spatial distribution of soil aluminum can be converted into a map showing the lime requirement of an area.

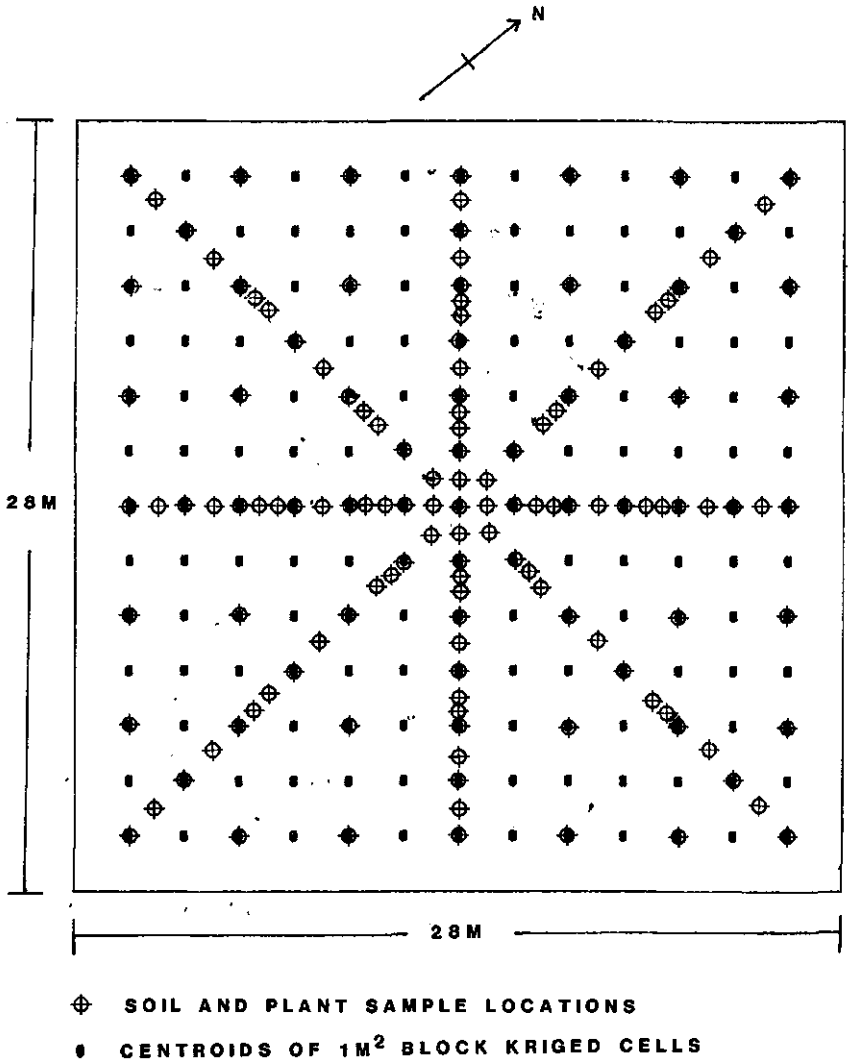


Fig. 24. Location of soil, plant samples and centroids of $1m^2$ block kriged cells, soil variability trial.

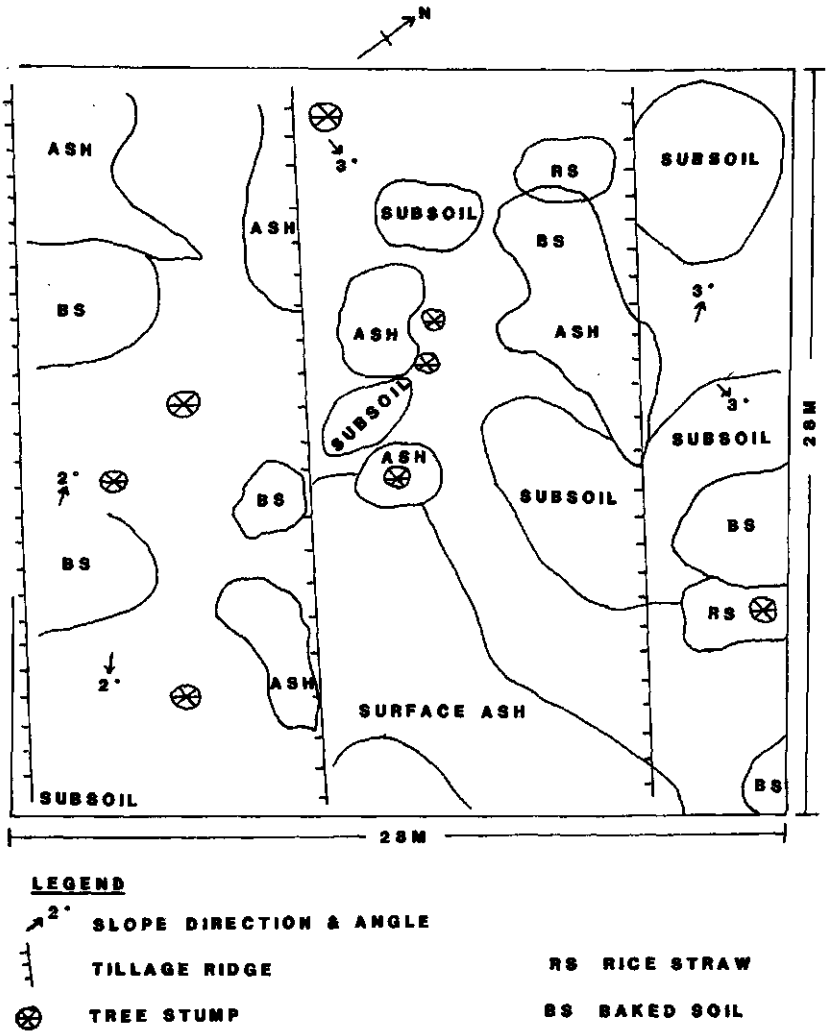


Fig. 25 Map of surface features, soil variability trial, Sitiung IIE. Scale 1:200.

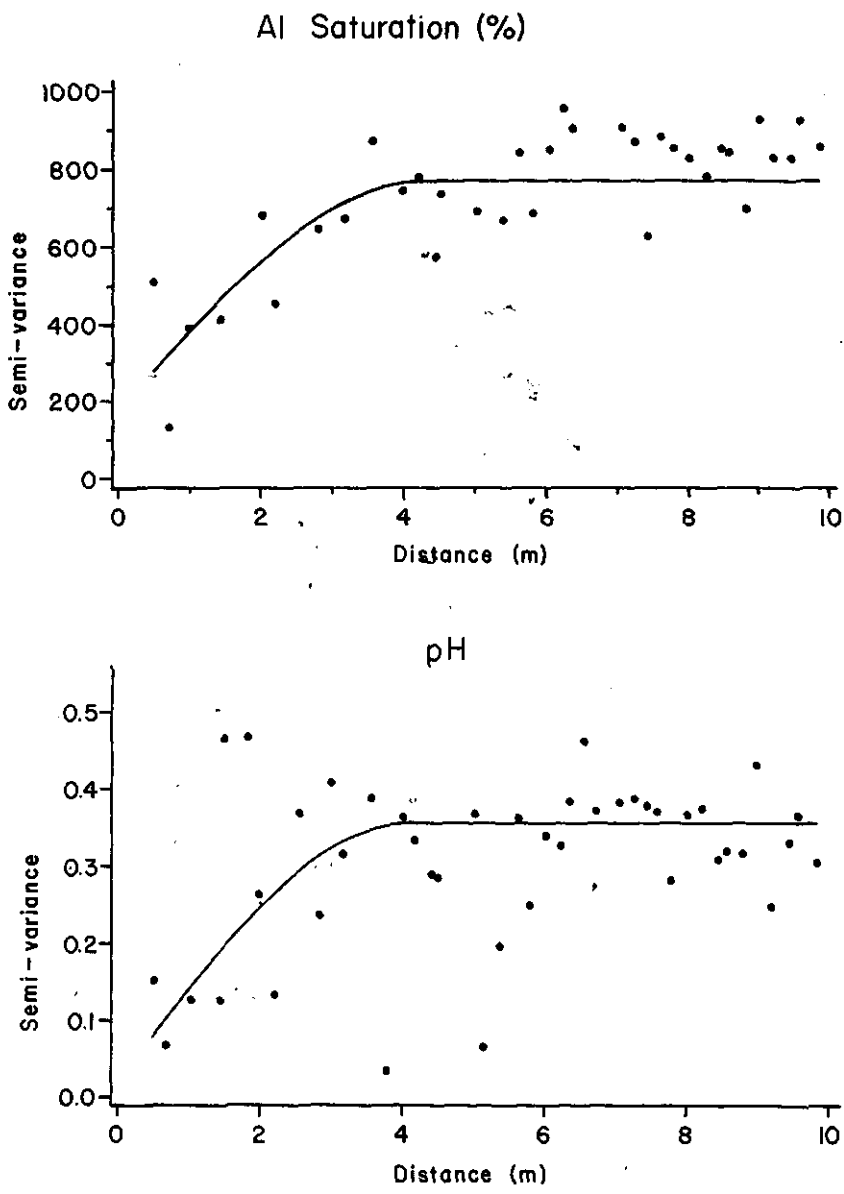


Fig. 26. Isotropic semi-variograms for Al saturation and pH, 0-20 cm depth, soil variability trial.

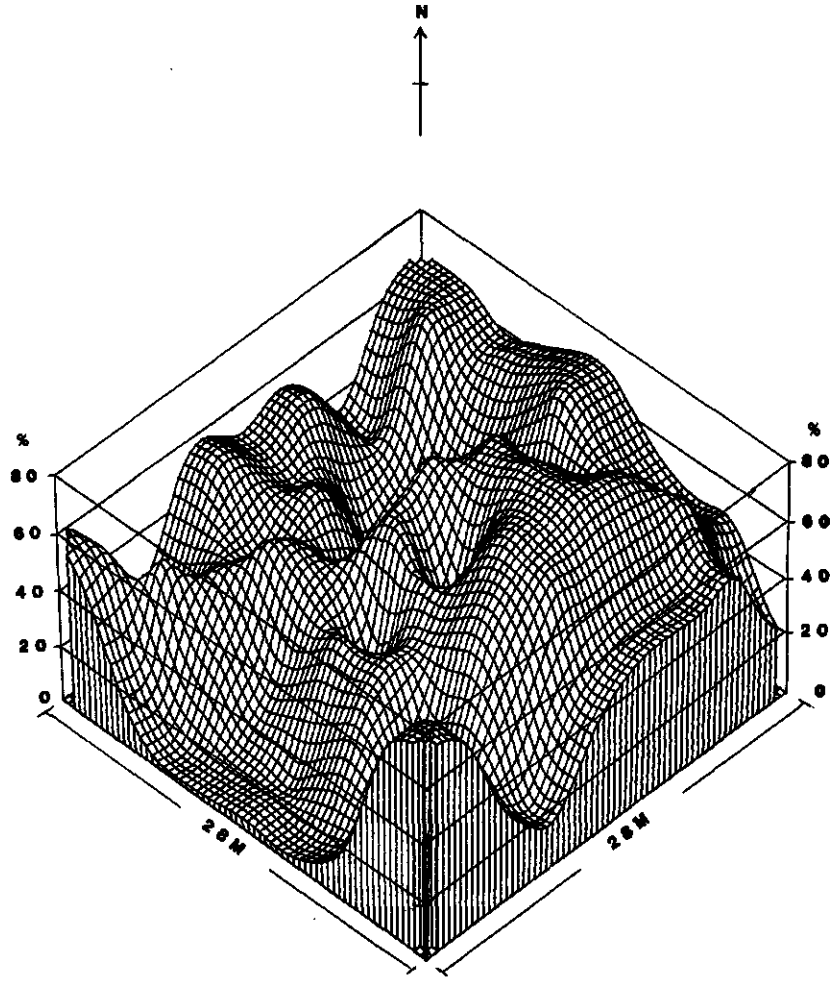


Fig. 27 Three dimensional diagram of Al saturation (%) block kriged over $1m^2$ cells, 0-20 cm depth, soil variability trial.

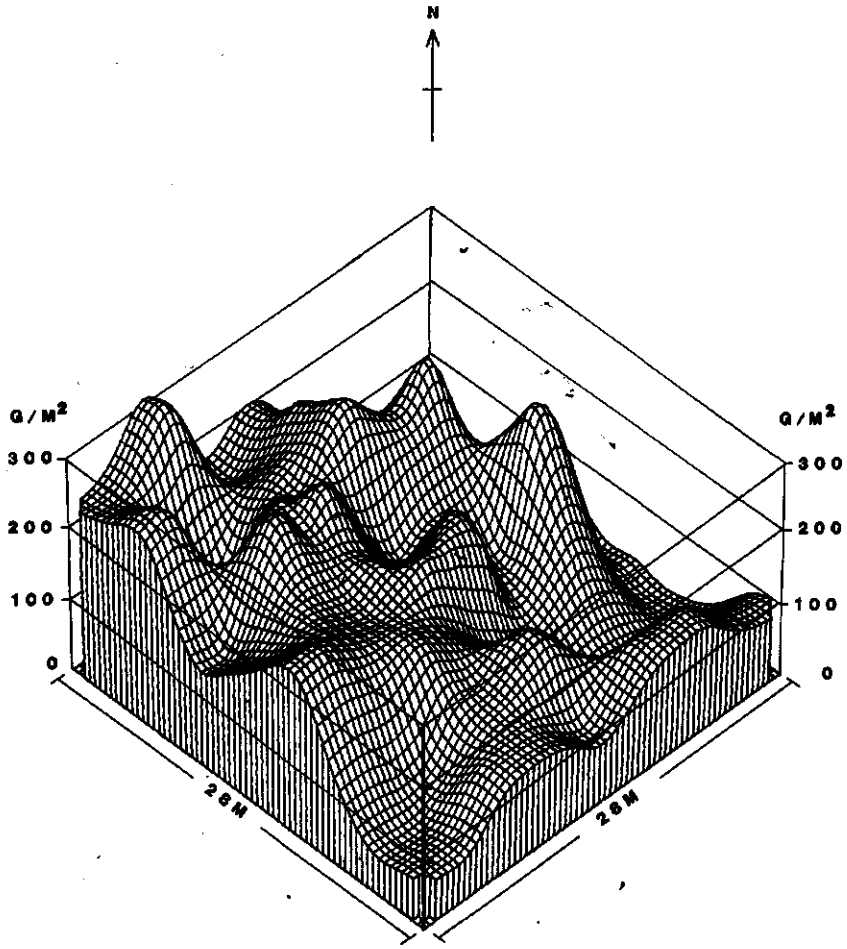


Fig. 28 Three dimensional diagram of rice grain yield (g/m^2) block kriged over $1m^2$ cells, soil variability trial.

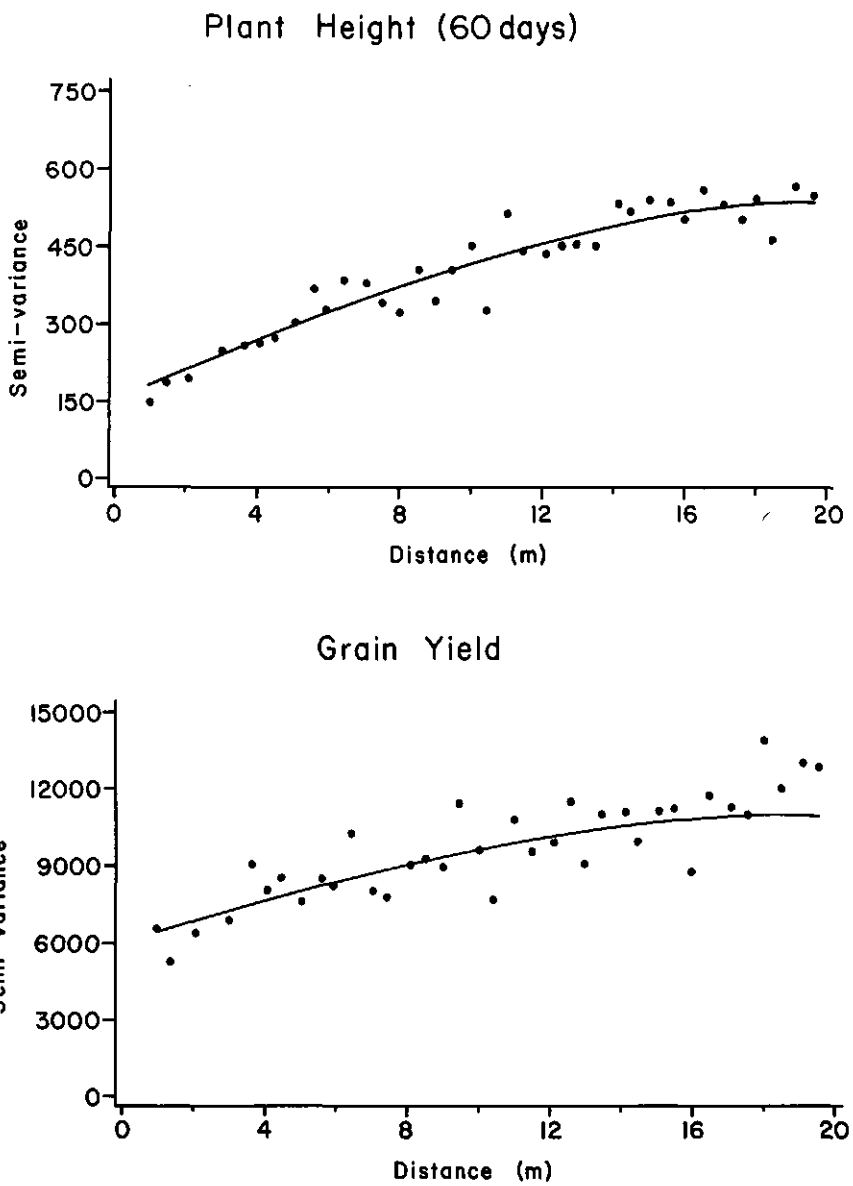


Fig. 29 Isotropic semi-variograms for plant height at 60 days and grain yield of upland rice, soil variability trial.

Table 6. Coefficients of correlation (r)[a] of soil chemical properties with plant height and yield components of upland rice soil variability trial.

Plant height Soil property	Stover dry at 60 days (cm)	Grain dry weight (g/m^2)	weight (g/m^2)
pH	0.51	0.44	0.38
Organic C (%)	-0.07^{ns}	-0.02^{ns}	0.02^{ns}
Total N (%)	0.09^{ns}	0.11^{ns}	0.14^{ns}
$NaHCO_3$-P (ppm)	0.06^{ns}	0.11^{ns}	0.11^{ns}
Exch. cations (meq/100g)			
Ca	0.56	0.55	0.49
Mg	0.52	0.45	0.40
K	0.41	0.38	0.46
Exch. Al (meq/100g)	-0.55	-0.47	-0.41
Al satn. (%)	-0.56	-0.51	-0.41
Cu (ppm)	0.07^{ns}	0.09^{ns}	0.12^{ns}
Zn (ppm)	0.42	0.34	0.33

a. All r values significant at $P < 0.01$ unless otherwise indicated.

ns = nonsignificant at $P = 0.05$.

Spatial variability of soil properties affect soil performance as demonstrated by Warwick and Gardner (1983). A uniform application of soil amendment in a spatially variable soil results in over application in some parts of the field and under application in others. If the range of soil variability is much larger than the means employed to apply the amendment, it may be possible to vary the rate according to needs.
Finally, spatial variability in soil properties that restrict soil use for dwellings, septic tank adsorption fields, small reservoirs, sewage lagoons, and so forth can be mapped so that marginal zones in a spatially variable field can be avoided.

Table 7. Mean comparisons[a] of plant height, stover and grain yield of upland rice among terrain units, soil variability trial.

Terrain unit	Plant height at 60 days (cm)	Stover dry weight (g/m^2)	Grain dry weight (g/m^2)
Burn sites (n=24)	92.5a	328.3a	227.1a
Surrounding soil (n=74)	78.5b	233.8b	166.5b
Exposed subsoil (n=23)	71.2b	192.7b	146.0b

a. Means within a column followed by the same letter are not significantly different at P = 0.05 according to Student's t test.

References

Adam, A. I., 1976. Semi-detailed soil survey report of Kenana Sugar Project (S.E. Ext.), White Nile Province. Soil Survey Administration, Sudan.

Campbell, J. B., 1978. Spatial variation of sand content and pH within single contiguous delineations of two soil mapping units. Soil Sci. Soc. Am. J. 42:460-464.

Haan, C. T., 1977. Statistical methods in hydrology. Iowa State University Press, Ames.

Trangmar, B. B., 1984. Spatial variability of soil properties in Sitiung, West Sumatra, Indonesia. Ph.D. dissertation, Dept. of Agronomy and Soil Science, University of Hawaii, Honolulu, Hawaii, USA.

Trangmar, B. B., U. Singh, R. S. Yost, and G. Uehara, 1982. Application of geostatistics for soil survey and land evaluation. 5th Int. Soil Classification Workshop, Nov. 2-11, 1982, Wad Medani, Sudan.

Trangmar, B. B., R. S. Yost, M. Sujadi, M. Soekardi, and G. Uehara, 1984. Regional variation of selected topsoil properties in Sitiung, West Sumatra, Indonesia. Research Series 026, Hawaii Inst. of Trop. Agr. and Human Resources, College of Trop. Agr. and Human Resources, University of Hawaii, Honolulu. Hawaii, USA.

Vander Zaag, P., R. L. Fox, R. S. Yost, B. B. Trangmar, K. Hayashi, and G. Uehara, 1981. Spatial variability of selected properties of Rwanda soils. 4th Int. Soil Classification Workshop, June 2-12, 1981, Kigali, Rwanda.

Warrick, A. W. and W. R. Gardner, 1983. Crop yield as affected by spatial variations of soil and irrigation. Water Resources Res. 19:181-186.

Discussion

M. Nash: How did you take the average distance?

G. Uehara: If by average distances you mean lags, they are selected to provide an adequate number of pairs of observation to compute a reliable semivariance. Large average distances give large numbers of pairs but few lags. Shorter lags give fewer pairs. In samples collected in a grid, the lag is generally the distance between the closest pair. The printout should include the number of pairs for each instance so that the reliability of each semivariance corresponding to each lag can be judged.

I. Murarka: For any of four samplings that you described, did you make a prediction for a spatial point based on your kriging results and then go into the field and observe the value? How did the predicted (calculated) value compare with the measurement(s)?

G. Uehara: It is probably simpler and cheaper to validate your results by jack-knifing. This procedure simply involves deleting an observed value from the analysis repeated for each observed value.

Remember that the kriging variance gives you a picture of the reliability of the predicted value.

P. Germann: Have you moved the jack-knifing across the field in order to get a "kriging variance"?

G. Uehara: Jack-knifing generally refers to the procedure of deleting one observation at a time. However, if one wishes, one can delete rows, columns or cluster of points to test the degree to which the deleted information can be recovered by kriging.

K. Lang: Given several hundred acres of regraded soil data collected on a 300 ft. grid basis, we wish to justify a reduction in the sampling intensity by using the jack-knife method. Should sample points be eliminated on a row basis or on an alternate sample point basis?

G. Uehara: The rule of thumb is to maintain a uniform density of observations over the area and to preserve sufficient number of closely spaced

observations to construct a useful semivariogram. My response to the previous question applies here.

D. Clay: How were the semivariance models fit, and were the models evaluated statistically?

G. Uehara: The selection of a model to fit the semivariogram is somewhat subjective. However, the slope and intercept of a linear model, for example, can be objectively determined by statistical means. Remember however, that each point on the semivariogram carries different weights because they are calculated from different number of pairs. The list of acceptable semivariogram models can be obtained from the literature.

B. Luxmoore: Have the analyses for Rwanda helped potato production?

G. Uehara: Not directly. Geostatistics helped the potato agronomist obtain a more accurate picture of the soil fertility status of the soils of Rwanda. This helped the agronomist make better choices in his work and this in turn probably improved his capability to help growers increase potato production.

A. Ruellan: The soil has a morphology. If we wish to emphasize our effort in terms of fertility, we have to take the samples as a function of the morphology of soils, and we have to study crop production in relation to morphology.

I think if we have a good understanding of soil morphology we can take less and better samples and achieve better statistics. The variability of the soil is normal. The problem is to understand this variability.

G. Uehara: It is not our intent to replace soil morphology with geostatistics. Class, grade and type of soil structure, color, consistency, horizon boundaries are important soil characteristics that cannot be analyzed geostatistically because they are qualities that cannot be expressed as continuous numerical values. This may be more a deficiency of soil morphology than a problem with geostatistics.

L.P. Wilding: Do you have any ideas or knowledge of how well the geostatistical kriged data sets can be extrapolated from one area of a given soil to the next area of like soil? Geostatistics is an interpolative tool but soil surveying is an extrapolative process?

G. Uehara: If we can demonstrate through geostatistics that similar kinds of spatial relations exist among similar named kinds of soils, regardless of where they occur, we will have solved one of the major soil science problems facing us today. This is a hypothesis that needs to be tested.

Stochastic models of fluid flow in heterogeneous media

Leslie Smith, Department of Geological Sciences, University of British Columbia, Vancouver, Canada.

Introduction

Significant advances have been achieved in the past ten years in our understanding of how the spatial variability of porous media influences groundwater flow. This work has led to the development of stochastic simulation techniques. In this approach, input parameters in the groundwater flow equation are assumed to be random variables with an associated probability distribution at each point in the flow domain. The spatial dependence between neighboring values of the random variable is defined in terms of an autocorrelation function, or a related measure of the spatial continuity. Solution of the flow equation leads to probability distributions on the output variables such as hydraulic head that can be interpreted in terms of the uncertainty in model prediction. The objectives of this paper are to review some basic elements of fluid flow in heterogeneous media, and to discuss current research directions.

The concepts discussed in this paper can be outlined by considering the cartoon shown in Figure 1. Assume the volume of water seeping beneath the embankment is to be estimated. Because of the spatial variability in hydraulic conductivity, there will be a degree of uncertainty in the discharge prediction. Even if considerable data were available, the exact patterns of spatial variation could never be uncovered. With a conventional deterministic model, calibration and validation techniques would be used during model parameterization in an attempt to improve the reliability of the discharge estimate. Uncertainty in the model output is determined, at best, in a subjective manner. The stochastic approach has been introduced to quantify that uncertainty. Central issues in this approach can be expressed by the following set of questions:

1. What are the patterns of the spatial variability in hydraulic conductivity, how is the spatial variability described in probabilistic terms, and what techniques are available to estimate model parameters?
2. What mathematical techniques can be used to form estimates of the probability distributions on the seepage volume?
3. Given a probabilistic model for the heterogeneity in hydraulic conductivity, what is the probabilistic structure of

the hydraulic potential and the fluid discharge?
4. How does the design of a data measurement network influence the magnitude of the uncertainty in a model prediction?
5. What is the probability that the system fails (ie., unacceptably high seepage), and what are the tradeoffs to be considered in a measurement program when assessing the probability of failure?

Questions 1, 2 and 3 deal with the mechanics of stochastic analyses and the physics of fluid flow in heterogeneous media. Considerable progress has been achieved in this area, with the theoretical foundations now essentially in place. Less is known about the concepts expressed in questions 4 and 5, which integrate model analysis with decisions on data worth. These latter topics identify an active field of current research. In the sections that follow, each of these questions is discussed in sequence.

Spatial Variability of Porous Media

Stochastic simulations of seepage beneath the embankment shown in Figure 1 would require that a number of assumptions be made in order to obtain a workable problem. The variations in hydraulic conductivity at the site are assumed to represent a realization of a stochastic process which is statistically homogeneous. Statistical homogeneity requires that hydraulic conductivity have the same mean and variance at every point and that the autocorrelation between hydraulic conductivities at any two points depend only on the vector separating those points and not their absolute position. Field data indicate that hydraulic conductivity can be represented by a lognormal probability distribution. The standard deviation in log hydraulic conductivity (base 10) estimated for data sets from a variety of geologic media ranges in value from 0.10 in homogeneous systems to 1.50 in very heterogeneous media (see an early review by Freeze, 1975). The integral scale, defined mathematically as the area beneath the autocorrelation function, is commonly used as a measure of the average distance over which hydraulic conductivity is correlated. Correlation lengths for hydraulic conductivity vary with the scale of the analysis. Horizontal correlations on the order of 1 to 10 m have been observed in samples distributed over distances of tens to a few hundred meters. In large-scale regional aquifers, correlation distances can be as large as 10 or 20 km (Hoeksema and Kitanidis, 1984). Layered heterogeneities on a scale similar to that of the flow domain would, if present, be identified as distinct units described by their own set of statistical parameters.

Several research groups are carrying out work to identify possible relationships between the nature of a depositional environment and the probabilistic models describing the resulting spatial variability in hydraulic conductivity. The hope here is that correlation structures and length scales can be identified which are typical of, for example, channel-fill sands, glacial outwash or alluvial fan deposits. As an example,

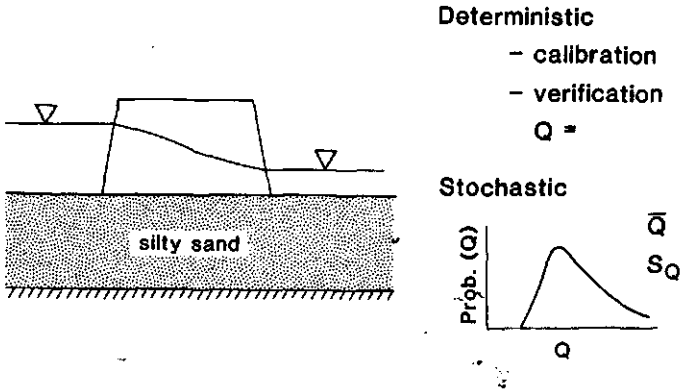


Figure 1. Schematic diagram differentiating between deterministic and stochastic solutions in predicting the seepage volume (Q) beneath the embankment.

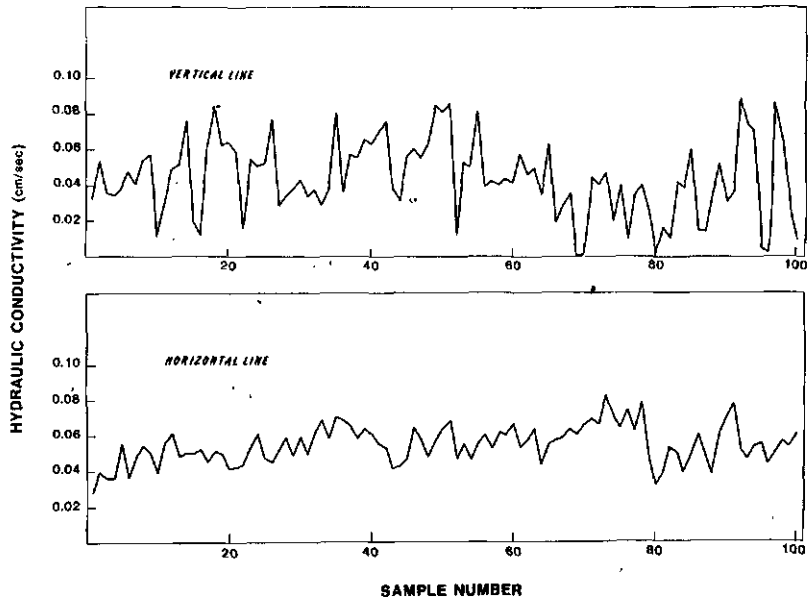


Figure 2. Spatial variations in hydraulic conductivity along horizontal and vertical line transects, Quadra Sands (from Smith, 1981).

consider the data shown in Figure 2. The two plots illustrate variations in hydraulic conductivity along two line transects from the Quadra Sand where it is exposed in a series of cliffs near Vancouver, Canada (Smith, 1981). This unit is a well-sorted, medium grained, horizontally stratified sand with relatively few silt or gravel interbeds. Identifiable horizontal beds are up to tens of meters in length and up to a meter thick. The Quadra Sand was formed as reworked distal outwash ahead of advancing glaciers to the north. The line transects consist of 100 samples on a spacing of 0.30 m. One line sample is parallel to the bedding, the other is directed up the face of the cliff.

The differing character of the variations within these two series is apparent. The heterogeneity is much greater in the vertical line sample. This result could be expected in a stratified medium whenever the length of the horizontal sample is of the same order as the scale of the bedding, while the vertical sample crosses the sedimentary sequence. The standard deviation is log hydraulic conductivity (base 10) is 0.09 for the horizontal line sample, and 0.36 for the vertical line sample.

Sample autocorrelation functions for these line transects are plotted in Figure 3. Several points can be noted. First, the extent of the autocorrelation between neighboring values is a function of the orientation of the line transect in the stratigraphic section. Second, the functional form of the autocorrelation may also depend upon orientation. Because the length of the horizontal transect is of the same order as the scale of the layering, the sample autocorrelation function reflects the point-to-point correlation within that layer. The vertical line sample is directed across the stratification, traversing a series of layers. As such, the sample autocorrelation function reflects structures within the sequence of layers, as well as that within each layer (in the vertical direction). The possiblity of anisotropy in the autocorrelation must be considering in any measurement program designed to estimate the parameters of the statistical model describing the heterogeneity. Knowledge of the geologic environment and its associated depositional features will be of extreme value in the interpretation of the spatial patterns of the hydraulic conductivity variations within a porous medium.

Solution Techniques

A number of techniques have been used in solving stochastic equations of groundwater flow. These approaches differ in their mathematical formulation and in the way in which the spatial variability in hydraulic conductivity is incorporated in the analysis. Summarized below are several of the techniques which have been used to model fluid flow in heterogeneous media. In the Monte Carlo approach, a deterministic boundary value problem is solved repetitively using a discrete representation of the flow domain, obtained by dividing that domain into a

number of equisized blocks. Single hydraulic conductivity values are assigned to each block by generating values from the statistical model describing the spatial variability. If warranted, values assigned to single blocks can be generated as spatial averages with appropriate modification of the probability model. Each realization is one member of the ensemble of all possible realizations which could be generated. Numerical techniques are used to solve for hydraulic head within each of the realizations. The output from a large number of such trials can be analyzed to obtain estimates of the probability distributions on hydraulic head and on other measures such as fluid discharge.

One of several different approaches can be used in generating realizations of a heterogeneous porous medium. Smith and Freeze (1979) use a stochastic process model to develop a set of linear equations linking hydraulic conductivity values in each of the blocks to neighboring values in adjacent blocks. Variability is introduced through a perturbation term added to each equation. An algorithm which operates on the spectral distribution function characterizing the hydraulic conductivity variations has been used by Mejia and Rodriguez-Iturbe (1974) and Freeze (1980). The turning bands method (Delhomme, 1979; Dettinger and Wilson, 1982), which is based on simulating a two-or three- dimensional random field as the sum of a series of one-dimensional realizations, provides an efficient technique for generating hydraulic conductivity variations in domains of large extent. These latter two techniques are favored because of greater flexibility in representing the form of the autocorrelation function.

An important class of stochastic problems have been solved using spectral analysis techniques (Bakr et al., 1978; Gutjahr et al., 1978; Mizell et al., 1982). In this approach, the parameters of the flow equation are written in terms of a mean and a perturbation. After subtraction of the terms describing the mean flow, an equation is developed connecting the perturbation in hydraulic head to that in hydraulic conductivity. From this equation, an expression is derived relating the power spectrum of the hydraulic conductivity perturbation to the resulting power spectrum in hydraulic head. A specific solution is developed once the form of the autocorrelation function in hydraulic conductivity is assigned. This spectral approach has the advantage of yielding a closed form solution for the statistical properties of hydraulic head and of representing the spatial structure in a continuum form. Analytic solutions developed in the space domain have been presented by Dagan (1979, 1981).

An alternative numerical technique to Monte Carlo simulation was presented by Sagar (1978) and extended by Dettinger and Wilson (1981) and Townley (1984). Their approach introduces variability in the medium properties after the formulation of the matrix equations which result from the finite difference or finite element approximations to the groundwater flow equation. The technique is based on a Taylor series expansion for the

hydraulic head, from which expressions for the expected value and variance of the solution vector can be computed if higher order terms are neglected. This approach has been used to advantage in investigating transient flow problems.

An important distinction can be made among the various stochastic models between those based on unconditional and conditional simulations. In an unconditional simulation, each realization is generated using only the parameters of the probabilistic model describing hydraulic conductivity, without regard to the location of data points. The standard deviation in hydraulic conductivity, and the uncertainty in the estimate of a hydraulic conductivity value at some point in the flow domain, are viewed as equivalent terms. The primary focus of studies adopting this representation has been to describe the physics of flow in heterogeneous media. In a second approach, the input parameters reflect a measure of the information on the spatial variation in hydraulic conductivity, with the standard deviation interpreted in terms of an estimation error. Emphasis here is placed on the use of hydrogeologic models as predictive tools. Using the techniques of kriging and conditional simulation (Delhomme, 1979), it is possible to preserve field data at measurement points in each realization. In this case, the ensemble is limited to include only those realizations that satisfy the available field data. Because of spatial autocorrelation, known values of hydraulic conductivity exert an influence not just at the measurement point but over a surrounding neighborhood. In solving the stochastic flow equation for hydraulic head, the standard deviation in hydraulic conductivity is effectively reduced in the neighborhood of each measurement point. As a result, the variability in hydraulic head may also be reduced in the vicinity of the measurement points.

Conditional simulation can be extended by also incorporating hydraulic head measurements in the stochastic model (Clifton and Neuman, 1982). Here the stochastic model is viewed in the framework of a parameter estimation problem. Clifton and Neuman use a statistical inverse simulation, coupled with kriging, to refine the estimate of the hydraulic conductivity and the standard deviation in that estimate in a number of subregions within the flow domain. Examples have been presented to show that the variance in the prediction of hydraulic head can be reduced markedly if the model is also conditioned on a set of hydraulic head measurements.

Stochastic Output

To illustrate several features of steady state fluid flow in a heterogeneous porous medium, consider the domain shown in the inset diagram of Figure 4. The domain is 20 m long and 10 m wide. The inflow boundary has a hydraulic head of 1.0, with a value of 0.0 on the outflow boundary. The upper and lower boundaries are impermeable, leading to a mean hydraulic gradient describing one-dimensional flow. The results that follow are

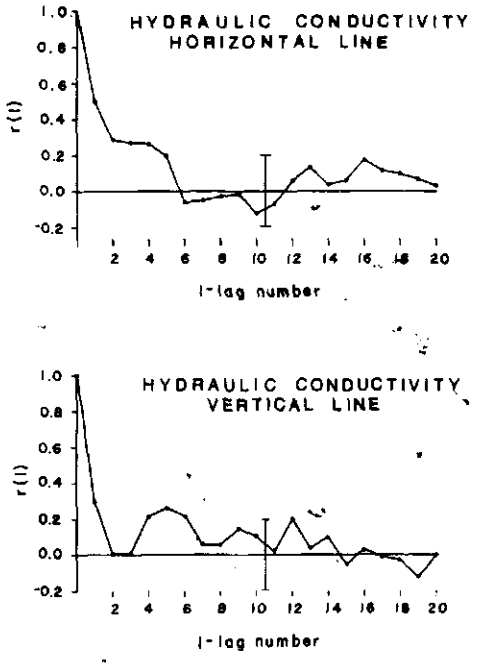


Figure 3. Sample autocorrelation functions for hydraulic conductivity variations along the line transects, Quadra Sands (from Smith, 1981).

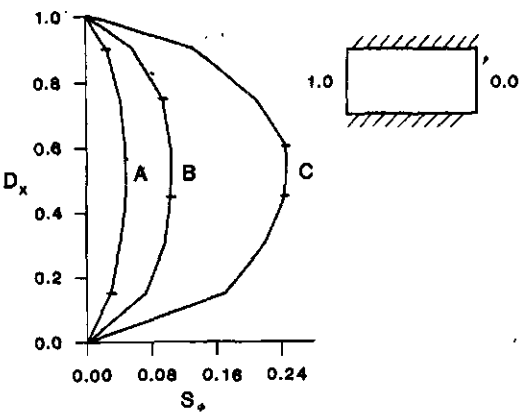


Figure 4. Influence of the standard deviation in hydraulic conductivity on the variability in hydraulic head for a case with unidirectional mean flow. Curves A, B, and C correspond to media with a standard deviation in log hydraulic conductivity of 0.21, 0.43, and 0.91; respectively. See text for details. (from Smith and Freeze, 1979)

based on unconditional simulations. The first example illustrates the influence of the standard deviation in hydraulic conductivity on the variability in hydraulic head. The spatial continuity of the variations, described by the integral scales in the two coordinate directions, are held constant with values approximately equal to 1.2 m. Results are shown along a line parallel to the x axis, midway between the upper and lower impermeable boundaries. In a more heterogeneous medium, the standard deviation in hydraulic head is larger. With greater hydraulic conductivity contrasts possible in a more heterogeneous medium, the deviation of hydraulic head at any point from its mean value at that point can increase. The parabolic shape of the curves reflects the statistical nonhomogeneity in hydraulic head because of the constant head values on the boundaries of the flow domain.

Figure 5 shows the effects of a nonuniform mean hydraulic gradient on the variability in hydraulic head. The boundary value problem is shown in the upper diagram, together with the deterministic solution for hydraulic head assuming the medium is homogeneous. The lower two diagrams are contour plots of the estimated standard deviation in hydraulic head for media with differing degrees of spatial continuity in hydraulic conductivity. In plot b, the integral scales are approximately 1.25 m; in plot c, 2.1 m. In both cases, the standard deviation in hydraulic conductivity is 0.43. Several conclusions can be formed. First, the standard deviation in hydraulic head is greatest in the region of flow domain where the mean hydraulic gradients are the highest, yet the region is removed from the constant head boundaries where the possible range of head values is constrained by the fixed head values there. Second, the variability in hydraulic head increases in a medium with a greater spatial continuity in hydraulic conductivity. This response occurs because if zones with hydraulic conductivities either above or below the mean value are likely to be more extensive, the hydraulic heads in those regions can deviate further from their mean values. Comparison of this example with the previous one indicates that the variability in hydraulic head depends upon both the heterogeneity in hydraulic conductivity and the nature of the flow system.

Analytic solutions for steady state fluid flow in an unbounded domain indicate that the perturbations in hydraulic head are more strongly correlated than the input perturbations in hydraulic conductivity (Bakr et al., 1978; Mizell et al., 1982). In addition, hydraulic head perturbations are correlated over a greater distance in a direction at right angles to the mean hydraulic gradient. Thus, the statistical structure of the hydraulic head variations is anisotropic, even though the autocorrelation in hydraulic conductivity may be isotropic.

Massmann (unpublished data, 1984) has carried out several simulations that demonstrate the manner in which the variability in hydraulic head depends upon the dimension of the flow domain. The simulations model unidirectional mean flow through a rectangular volume. The domain was divided into 300 square

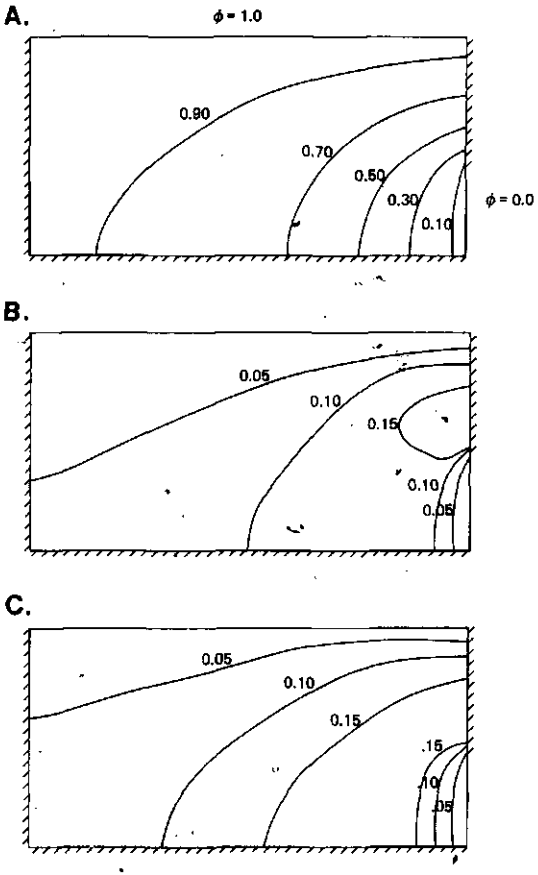


Figure 5. Influence of the spatial autocorrelation in hydraulic conductivity on the variability in hydraulic head, nonuniform mean gradient. Plot A is the deterministic solution for a homogeneous medium; plots B and C show contours of the standard deviation in hydraulic head for media with differing degrees of spatial continuity in hydraulic conductivity. See text for details. (from Smith and Freeze, 1979)

elements (10x6x5), with the long axis parallel to the mean flow direction. Realizations of the three-dimensional random field were generated using a turnings band technique, with an isotropic exponential autocorrelation function. For each realization, flow was first modeled in three dimensions, and then along the central plane parallel to the long axis of the domain, and along the center line of that plane. Such an experiment will show that the variability in hydraulic head is smaller in a three-dimensional analysis than in either a one or two-dimensional simulation. The greatest difference will occur between the one and two-dimensional simulations. The variability in hydraulic head depends upon the dimension of the stochastic model because individual zones of low or high hydraulic conductivity are of less importance in perturbing the flow field in multidimensional models. Similar effects are noted in the analytic solutions of Bakr et al (1978) and Dagan (1979). Realistic assessment of the sensitivity of hydraulic head or discharge predictions to the unknown patterns of spatial variability in hydraulic conductivity must consider the three-dimensional nature of groundwater flow.

Network Design and Risk Assessment

There is considerable potential in adopting stochastic techniques in the solution of problems involving fluid flow through porous media. The primary advancement to date has been the development of better physical models describing fluid flow (and transport) in heterogeneous media. This work sets the stage for the future use of stochastic models in a design mode to optimize the planning and implementation of field sampling programs. A central concept here is to consider the location of measurement sites in relation to the use of that data in model prediction, rather than as a problem in mapping the spatial variations in hydraulic conductivity within some region of interest. Stochastic models also hold promise as a central component in a risk analysis of an engineering design which is based, in part, on a hydrogeologic prediction. Here the uncertainties in model prediction are evaluated in light of an associated probability of system failure and the (remedial) costs following a failure.

Two different concerns arise when discussing data requirements for a stochastic simulation. First, there is uncertainty in estimating the parameters of the statistical model characterizing the spatial variability. The estimation error for these parameters will be a function of the number of sample points, the geometry of the sample grid, the autocorrelation between neighboring hydraulic conductivity values, the size of the domain, and the variance of the hydraulic conductivity distribution. Second, there is the impact of measured values in constraining the unknown patterns of spatial variation in hydraulic conductivity. Work by Smith and Schwartz (1981) suggests that given a realistic number of data points, unknown patterns of spatial variation in hydraulic

conductivity are a greater source of uncertainty in model prediction than are errors in estimating the mean and standard deviation of the hydraulic conductivity distribution. A determination has yet to be made of the importance of errors in estimating the autocorrelation function. Analytic solutions for various flow problems indicate that the output distributions can be sensitive to the form of the autocorrelation function. In this case, estimation of the autocorrelation function will be an important consideration in data collection.

Conditional simulation provides a mechanism for investigating how the design of a data measurement network influences the uncertainty in model prediction. Delhomme (1979) has shown that the locations of sample points relative to boundary locations and the hydraulic conditions on those boundaries are important in determining the degree of uncertainty in hydraulic head predictions. Uncertainties in hydraulic conductivity in key regions of the flow domain can result in large uncertainties in hydraulic head throughout most of the flow domain. Conversely, there will be regions where the uncertainty in hydraulic head is insensitive to the variability in hydraulic conductivity. Recognition of these 'key regions' in the flow domain will be an integral part of the efficient design of a measurement network that exploits the variable sensitivity of hydraulic head to the spatial variability in hydraulic conductivity.

With recent advances in the design of multiple-port piezometers, it is now feasible to collect in situ data on the spatial variability in hydraulic conductivity in a three-dimensional framework. This instrumentation allows many independent sampling points to be established in a single borehole. Consider the goundwater flow system shown schematically in Figure 6. Two statistically homogeneous units are present in the system. The upper boundary is the water table. The face which is shaded is assigned a constant value of hydraulic head. All other boundaries are taken to be impermeable. Given the objective of predicting the volume of discharge across the constant head boundary, guidelines are required on which to base decisions on the number of boreholes needed, their location, and the depth and number of observation points at each location. For example, choices can be made in deciding where to site multiple-port piezometers to collect detailed vertical information, and where a larger set of spatially-distributed, single-port piezometers may be more effective. Similarly, the number of observation points in each of the two geologic units should be selected is some optimum manner. The variable sensitivity of hydraulic head perturbations to the spatial variability in hydraulic conductivity is an important factor in addressing these kinds of questions.

A framework is required to examine tradeoffs in selecting among potential sampling sites. Several approaches are possible. A sampling strategy may be preferred if it leads to smaller estimation errors in model prediction. Sensitivity analyses using the estimation error as a measure of the

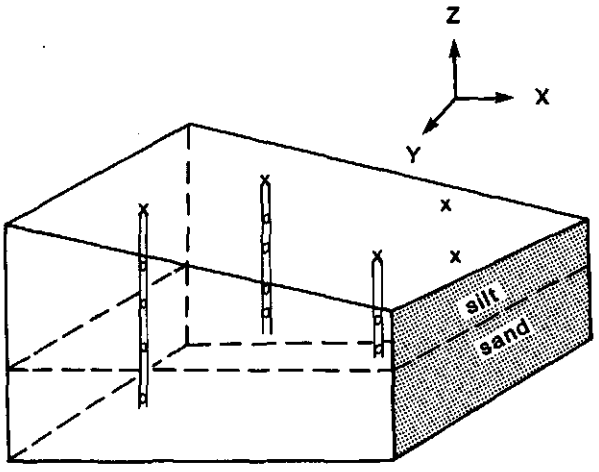


Figure 6. Schematic plot of a sampling network using multiple-port piezometers. Guidelines are required on which to base decisions on the number of boreholes needed, their location, and the depth and number of observation points at each location.

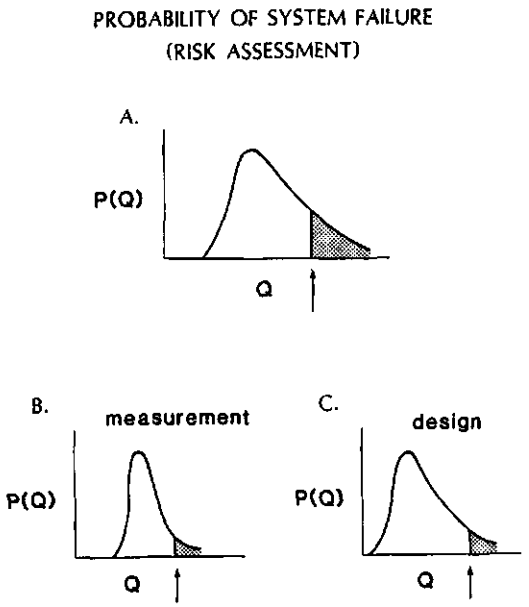


Figure 7. Schematic plot of the probability of excessive seepage (shaded region) beneath the embankment shown in Figure 1. Plots b and c refer to strategies for reducing the probability of failure by either collecting more data in key regions of the flow domain, or by more conservative design.

efficiency of a sample grid can be carried out to establish qualitative guidelines for site selection. However, this approach neglects an important variable, the cost of data acquisition. The use of the terms such as 'effective' and 'optimal' generally infers the definition of an objective function with which to judge the merits of alternative designs. Using techniques of decision theory, a sampling strategy can be assessed under the goal of minimizing the estimation error in model prediction, within the constraint of the total funds available for site investigation.

Integration of stochastic flow analyses with decisions on data worth must eventually be considered in light of the purpose for which the prediction is developed. The significance and impact of uncertainties in model prediction originating from the heterogeneous nature of porous media can be evaluated in terms of a probability of system failure. Consider again the cartoon shown in Figure 1. A prediction is required of the volume of water seeping beneath the embankment. The system is assumed to have failed if the discharge exceeds a maximum permissible level. By carrying out an unconditional simulation, it is possible to estimate the probability that a particular alignment of high-permeability elements leads to such a failure (Figure 7a). The extent to which the standard deviation of this distribution can be reduced by collecting more hydraulic conductivity data can be determined using conditional simulation techniques (Figure 7b). Note that constraining the patterns of spatial variation in hydraulic conductivity beneath the embankment will provide a better estimate of the probability of failure, but not necessarily a reduction in that probability. As an alternative to further sampling, modifications in the design of the embankment may be incorporated to decrease the expected value of the seepage volume and thereby reduce the probability of failure (Figure 7c). Selection among these strategies (more data in key regions of the flow domain, more conservative design) must also include consideration of the remedial costs if the seepage volume exceeds the maximum allowable. The determination of an acceptable level of uncertainty in model prediction is tied to the costs of further site investigation, the design of measurement networks, and the consequences of system failure. Research is ongoing to incorporate stochastic groundwater models in such risk assessments.

Summary

The effects on fluid flow of the spatial variability within a porous medium can be evaluated quantitatively using stochastic modeling techniques. The primary focus of research to date has been on the development of methodologies for the solution of the stochastic flow equations and on developing a better understanding of the physics of flow in heterogeneous media. The potential exists for using stochastic models in the planning and evaluation of data measurement networks where the locations

of sample sites are considered in relation to the use of that data in model prediction. The efficient design of a measurement network should exploit the variable sensitivity of hydraulic head to the unknown patterns of spatial variation in hydraulic conductivity. Stochastic models are likely to also be incorporated as an essential element in risk assessments which require hydrogeologic predictions.

Acknowledgements

The author has benefitted from discussions with Joel Massmann and Allan Freeze.

References

Bakr, A. A., L. W. Gelhar, A. L. Gutjahr, and J. R. MacMillan, 1978. Stochastic analysis of spatial variability in subsurface flows, 1, Comparison of one- and three-dimensional flows. Water Resour. Res., 14(2), 263-271.

Clifton, P. M., and S. P. Neuman, 1982. Effects of kriging and inverse modeling on conditional simulation of the Avra Valley Aquifer in Southern Arizona. Water Resour. Res., 18(4), 1215-1234.

Dagan, G., 1979. Models of groundwater flow in statistically homogeneous porous formations. Water Resour. Res., 15(1), 47-63.

Dagan, G., 1981. Analysis of flow through heterogeneous random aquifers by the method of embedding matrix, 1, Steady flow. Water Resour. Res., 17(1), 107-121.

Delhomme, J. P., 1979. Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach. Water Resour. Res., 15(2), 269-280.

Dettinger, M. D., and J. L. Wilson, 1981. First order analysis of uncertainty in numerical models of groundwater flow, Part 1, Mathematical development. Water Resour. Res., 17(1), 149-161.

Freeze, R. A., 1975. A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. Water Resour. Res. 11(5), 725-741.

Freeze, R. A., 1980. A stochastic-conceptual analysis of rainfall-runoff processes on a hillslope. Water Resour. Res., 16(2), 391-408.

Gutjahr, A. L., and L. W. Gelhar, 1981. Stochastic models of subsurface flow: Infinite versus finite domains and stationarity. Water Resour. Res., 17(2), 337-350.

Gutjahr, A. L., L. W. Gelhar, A. A. Bakr, and J. R. MacMillan, 1978. Stochastic analysis of spatial variability in subsurface flows, 2, Evaluation and application. Water Resour. Res., 14(5), 953-959.

Hoeksema, R. J. and P. K. Kitanidis, 1984. Analysis of the spatial structure of properties of selected aquifers. in press, Water Resour. Res.

Mantoglou, A., and J. L. Wilson, 1982. The turning bands method for simulation of random fields using line generation by a spectral method. Water Resour. Res., 18(5), 1379-1394.

Mejia, J., and I. Rodriguez-Iturbe, 1974. On the synthesis of random fields from the spectrum: An application to the generation of hydrologic spatial processes. Water Resour. Res., 10(4), 705-711.

Mizell, S. A., A. L. Gutjahr, and L. W. Gelhar, 1982. Stochastic analysis of spatial variability in two-dimensional steady groundwater flow assuming stationary and unstationary heads. Water Resour. Res., 18(4), 1053-1067.

Sagar, B., 1978. Galerkin finite element procedure for analyzing flow through random media. Water Resour. Res., 14(6), 1035-1044.

Smith, L., 1981. Spatial variability of flow parameters in a stratified sand. Mathematical Geology, 13(1), 1-21.

Smith, L., and R. A. Freeze, 1979. Stochastic analysis of steady state groundwater flow in a bounded domain, 2, Two-dimensional simulations. Water Resour. Res., 15(6), 1543-1559.

Smith, L., and F. W. Schwartz, 1981. Mass transport, 3, Role of hydraulic conductivity data in prediction. Water Resour. Res., 17(5), 1463-1479.

Townley, L.R., 1984. Second order effects of uncertain transmissivites on predictions of piezometric heads. Fifth International Conference on Finite Elements in Water Resources, Vermont.

Discussion

B. Overmas: Have you accounted for anisotropy in your model?

L. Smith: In heterogeneous media, there is the possibility of anisotropy in both the spatial autocorrelation function (statistical anisotropy) and in the local mean values of hydraulic conductivity (hydraulic anisotropy). It is possible to define probability distributions for both the horizontal and vertical hydraulic conductivity. Presumably these random variables would be correlated. Both factors can be included in the numerical models in a straightforward manner.

B. Overmas: Does the 3-dimensional approach indicate that we are dealing with too much variation compared with a 1-dimensional approach?

L. Smith: A one-dimensional simulation leads to greater variability in the output distributions than does a 3-dimensional analysis. Individual zones of low or high hydraulic conductivity are of less importance in perturbing the flow field in multidimensional models. Realistic assessment of the sensitivity of the model output to spatial variability should account for the 3-dimensional nature of groundwater flow, even in the case where the mean flow is unidirectional. In that sense, a 1-dimensional analysis will overestimate the variability in the output variables.

Spatial variability of soil-water properties in irrigated soils

P. J. Wierenga, Department of Crop and Soil Sciences, New Mexico State University, Las Cruces, NM 88001, USA.

Introduction

A number of reviews have recently been written on spatial variability of soils (Peck, 1983; Warrick et al., 1985). These reviews have presented geostatistical techniques used in the analysis of field data and summarized much of the field data on spatial variability of soil water properties. As is evident from these review articles, the use of geostatistical methods has gained increasing support among soil scientists. Techniques that appear to have been used most frequently are based on the theory of regionalized variables (Matheron, 1965), or on time series analysis theory. With either of these techniques, the spatial dependence of soil properties is investigated through variograms or correllograms. Next the information on spatial dependence is used, in conjunction with kriging, to obtain information at locations other than where the original date were obtained, as for mapping. A sizeable number of soil properties were investigated using geostatistical methods. These include soil texture (Webster and Cuanalo, 1975), saturated hydraulic conductivity (Russo and Bresslet, 1981), infiltration rate (Vieira et al., 1981), water content (Gajem et al., 1981) and electrical conductivities of saturation extracts (Hajrasuliha et al., 1980). These studies have shown that there is spatial dependence of many soil properties, invalidating the use of statistics based on independent samples. However, there appears to be no clear trend in the degree of spatial dependence. The degree of this dependence, or the correlation length, has been found to vary for each soil variable and study area and may also be a function of time. For example in the studies mentioned above, correlation distances vary from 230 m for soil texture to 5 m for water content. Furthermore, the degree of spatial dependence seems to depend on the sampling distance, with greater spatial dependence found for larger sampling distances (Gajem, 1981; Kies, 1982; Jury, 1984).

In view of this uncertainty and considering the potential of geostatistical techniques for interpreting soils data, it seems appropriate to examine additional field data. In particular, we want to look at spatial variability of tension, of water content, and of the infiltration rate of irrigated soils. Three field studies will be discussed in which soil water tension, water

content and hydraulic conductivity were measured in considerable detail. In the first study tension was measured along a trickle irrigated row crop. The purpose of the second study was to measure the variability in tension and water content along a 91 m transect of irrigated bare soil. In the third study the hydraulic conductivity was measured over a 100 ha farm and compared with hydraulic conductivity values obtained from a detailed soil survey. The third study was included to demonstrate the importance of soil survey information in characterizing physical properties of field soils and vice versa.

Variation in soil water tension along a trickle line

Ninety-nine tensiometers were placed 60 cm apart along a trickle irrigated row of chile peppers (Saddiq, et al., 1985). All tensiometers were first installed with their tips at 15 cm below soil surface and, following a number of readings at this depth, pushed down so their tips were at the 30 cm depth. Near the end of the growing season, the row with tensiometers was flood irrigated from both sides and a set of tensiometer readings was taken. All tensiometer readings were made with a handheld pressure transducer (Marthaler et al., 1983).

Spatial dependence between tension data was calculated with equations (1) and (2) as follows:

$$r(h) = \frac{cov(T(x) \cdot T(x+h))}{\sqrt{var(T(x))} \cdot \sqrt{var(T(x+h))}} \qquad (1)$$

$$\gamma(h) = \tfrac{1}{2} \cdot var(T(x)-T(x+h)) \qquad (2)$$

where r(h) is the autocorrelation, h the lag or distance between tensiometers, T(x) the soil water tension at point x and $\gamma(h)$ the semivariance (Rendu, 1978). Cumulative frequency distributions and fractile diagrams (Vieira et al., 1981) were calculated for the tension data at several times after irrigation or rain.

Results showed that, after flood irrigation or after rain, soil water tension data was normally distributed. After trickle irrigation there was a much greater spread in the data with more high and low values and no clearly defined statistical distribution.

In Fig. 1 the semivariance of soil-water tension is plotted versus distance for the 30 cm depth 2 and 13 days, respectively, after flooding the row with 26 cm water. The data clearly show spatial correlation, with a range of dependence of 3 to 5 m, similar to what was observed after rainfall. The presence of the nugget in the variogram (Fig. 1B), which was not observed for data taken 2, 4 and 6 days after flooding, indicates greater variation in soil water tension as the soil dries out. There

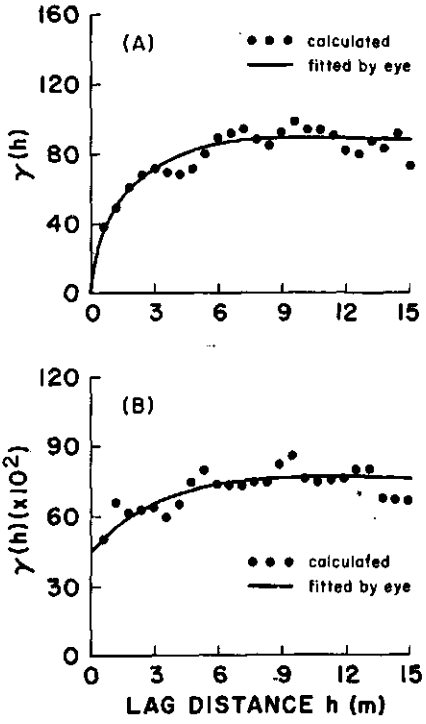


Fig. 1 Semivariograms for soil-water tension at 30 cm taken two days (A) and thirteen days (B), respectively, after flooding with 26 cm water.

was little or no spatial structure in the data after trickle irrigation, even though the amounts of water added per trickle irrigation varied between 1 and 7 cm. Even after 7 cm of water was applied through the trickle line, the distance of dependence was only 1.2 m. In contrast, a rainfall of 2.5 cm resulted in a spatial dependence of 5 to 6 m. Thus, it appears that the presence or absence of spatial dependence in soil water tension is closely related to the method of water application. These results imply that tests for uniform application of water over a field could be designed based on correlation distance of soil water tension. Thus, a large correlation length of soil water tension right after irrigation would mean uniform water application.

Table 1 summarizes the results of the various experiments. These data (Saddiq et al., 1985) show that, as the soil dries out and its tension increases, the variance increases. This agrees with the results of Yeh et al. (1984) who predicted such behavior on the basis of a stochastic analysis of unsaturated

Table 1 Mean, variance, coefficient of variation and distance of dependence for soil water tension at 30 cm as a function of time after flood and trickle irrigation.

Water application method	amount	Time after	Mean	Variance	CV	Distance of dependence
Date	(cm)	(days)	mbar	$mbar^2$	%	m
Flood	26.0	2	65	105	16	5.4
10/05/81		4	82	210	18	4.2
		6	112	436	19	4.8
		9	152	1589	26	3.6
		11	215	5092	33	5.4
		13	266	9085	36	1.2
Trickle	1.6	1	143	11290	74	<0.6
8/30/81		2	167	25130	95	<0.6
		3	208	36300	92	<0.6
		4	304	35240	62	<0.6
		5	461	39820	43	1.2
Trickle	7.0	1	47	116	23	1.2
1/29/82						

water flow through heterogeneous soil. These data further show that the variance and the coefficient of variation of soil water tension are generally much higher in trickle irrigated soil than in flood irrigated or in rainfed soil (not shown). While CV's ranged between 20 and 40% in flood irrigated or rainfed soil, they were between 40 and 100% in trickle irrigated soil at comparable mean soil water tensions. This compares with CV's of between 10 and 30% for water contents of core samples held at tensions of 0 to 200 cm (Nielsen et al., 1973, Gajem et al., 1981; Gumaa, 1978). No data were found in the literature on coefficients of variation of soil water tension in field soils. However, it is well known that many researchers have experienced problems in using tensiometers for scheduling irrigation. These problems are at least partly the result of the considerable spatial variability of soil water tension in field soils.

Variation in tension and water content along a 91 m transect

Neutron access tubes were installed along the center of a 3 m wide by 94 m long field plot of clay loam over sand, located adjacent to the trickle irrigated chile plot discussed above (Nash, 1984). A total of 91 access tubes were placed at 1 m intervals to a depth of 150 cm. Five tensiometers with their tips at 30, 60, 90, 120 and 150 cm below soil surface were installed at each neutron access tube along a line perpendicular to the transect. The 60 and 120 cm tensiometers were placed 30

and 60 cm to the right of each tube and the 30, 90 and 150 cm tensiometers at 30, 60 and 90 cm to the left of each tube. A total of 455 tensiometers were used. The plot was flooded with 14.3 cm water and measurements of water content and tension were taken at frequent time intervals.

The complete results of this study are presented elsewhere (Nash, 1984). A summary of the results will be presented here. Fig. 2 presents the variation in tension along the transect at 30 cm and at 120 cm, 1 day after flood-irrigation. The data in figure 2 show considerable variation in tension with distance along the transect. This was true for all measurement periods after irrigation. Table 2 summarizes the results for soil water tension during the first 44 days after flooding.

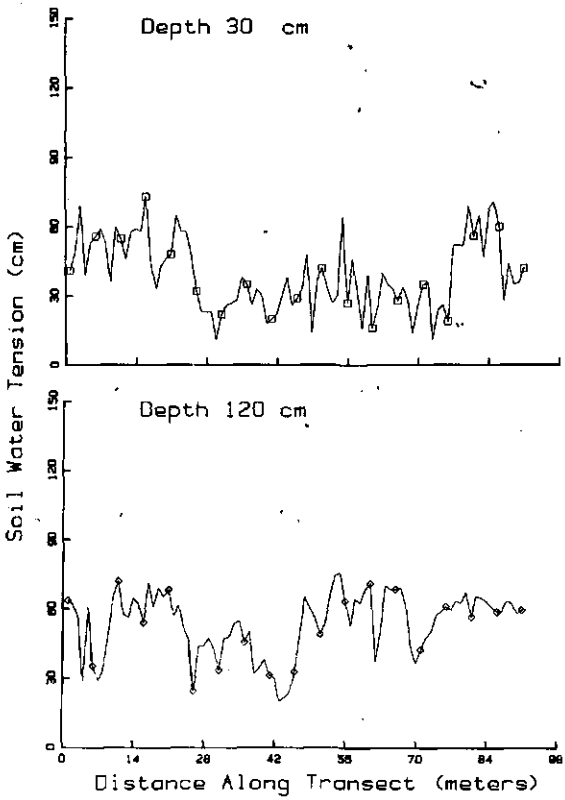


Fig. 2 Soil water tension (cm) at 30 and 120 cm as a function of distance along transect one day after flooding.

The data in table 2 show the change in tension with time after flooding for the 30, 60 and 120 cm depths. Note first that it takes more than 14 days for the tension at 30 cm to reach a value greater than 0.1 bar. In fact, from a plot of tension versus time Nash (1984) concluded that it took 25 days for the average tension at 30 cm to reach 0.1 bar, while it took 44 days for the average tension at 60 cm to reach this value. Thus, if

Table 2 Mean, variance, coefficient of variation, and distance of dependence for soil water tension at 30, 60 and 120 cm as a function of time after flooding.

Depth	Time after flooding	Mean	Variance	CV	Distance of dependence
m	days	mbar	$mbar^2$	%	m
0.3	1	39	241	40	15
	2	54	234	28	--
	4	62	454	35	--
	8	78	624	32	--
	14	92	518	25	22
	44	118	700	22	27
0.6	1	38	771	72	17
	2	45	800	62	--
	4	55	1031	59	--
	8	65	1156	53	--
	14	77	1083	43	19
	44	106	792	26	13
1.2	1	72	387	25	15
	2	69	307	25	--
	4	62	260	26	--
	8	66	303	26	--
	14	68	115	16	12
	44	77	238	20	--

field capacity is defined as the water content three days after irrigation or a heavy rain, then field capacity in this soil is between 0.05 and 0.06 bar, rather than 0.1 bar, as commonly assumed in the literature. Note also that the variance in soil water tension at 30 cm increases with tension. The same trend may be observed for the 60 cm depth. This agrees with Yeh et al., 1985, who predicted that the variance of soil water tension should increase as the soil dries out. The coefficients of variation in tension tend to decrease as the tension increases. Apparently, with time there is some redistribution of water in both horizontal and vertical directions, so that the variability in tension becomes less. However, even after 44 days of redistribution the coefficients of variation are still 20% or higher.

Comparing the data in Table 2 with those in Table 1, one sees that the trickle irrigated soil dries out faster, while the coefficient of variation of the tension in trickle irrigated soil increases rather than decreases with time as for the flood irrigated soil. The trickle irrigated soil was not covered with plastic, resulting in water loss by surface evaporation. Furthermore, although the tension measurements were taken late in the season, there may have been some water uptake by the chile plants, which increased variability in tension in this soil and resulted in larger values for the coefficient of variation. The last column in table 2 shows the distances of dependence, which varied between 12 and 27 meter (distances of dependence for days 2, 4, and 8 were not calculated). Thus, even though the variability in tension is large, especially at the 60 cm depth, there is considerable spatial dependence between adjacent tension values.

Changes in water content with distance along the transect were also large, especially for the subsoil. Fig. 3 shows spatial

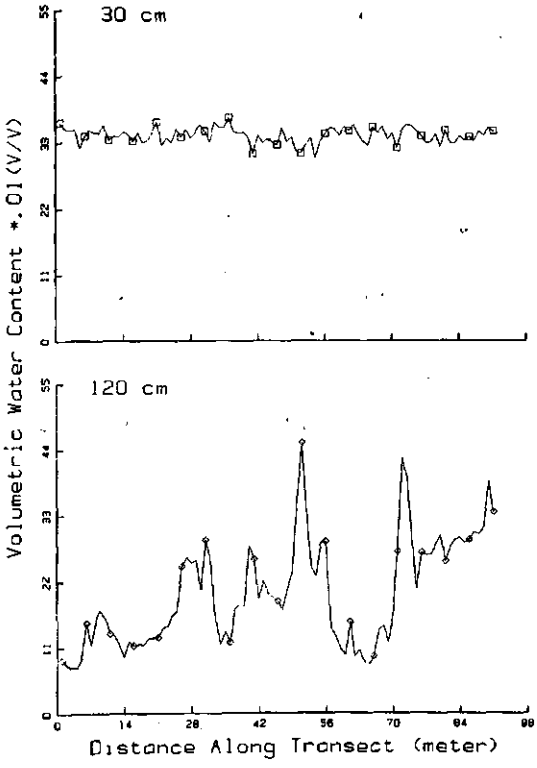


Fig. 3 Water content (cm^3/cm^3) at 30 and 120 cm as a function of distance along the transect one day after flooding.

variations in water content for the 30 and 120 cm depths one day after flood-irrigating the transect with 13.4 cm water. The variations in subsoil water content are especially large. This large variation at the deeper depths is the result of variations in texture.

Table 3 presents the statistical parameters for the water content at depths 30, 60 and 120 cm at various times after flood-irrigation. Note that the water content at 120 cm is much lower than the water content at 60 cm or above. This is due to the lower clay content at 120 cm. The lower clay content at 120 cm also causes the soil at this depth to drain faster as is evident from the 7% decrease in water content, over 44 days versus a much smaller decrease in water content with time at the upper two depths. The greater variation in water content with distance along the transect at 120 cm as compared to the variation in water content at 30 and 60 cm is also evident from the larger CV's at 120 cm. While the coefficients of variation in water content at 30 cm are less than 13%, they have values between 40 and 50% at the 120 cm depth.

Table 3 Mean, variance, coefficient of variation and distance of dependence for water content at 30, 60 and 120 cm, as a function of time after flood-irrigation.

Depth	Time after flood-irrigation	Mean	Variance	CV	Distance of dependence
m	days	cm^3/cm^3	$(cm^3/cm^3)^2$	%	m
0.3	1	.342	.0002	4.1	-
	2	.346	.0001	3.3	-
	4	.356	.0003	4.5	-
	8	.348	.0002	3.7	-
	14	.359	.001	6.8	20.0
	44	.358	.002	12.8	8.0
0.6	1	.403	.003	13.6	21.0
	2	.393	.004	16.2	-
	4	.391	.006	20.0	-
	8	.369	.010	27.1	-
	14	.363	.011	30.1	22.0
	44	.377	.011	27.8	19.5
1.2	1	.200	.008	44.3	12.4
	2	.193	.007	42.7	-
	4	.193	.008	45.1	-
	8	.151	.005	45.9	-
	14	.146	.005	45.9	9.0
	44	.130	.004	47.3	11.0

The last column in Table 3 shows the distance of dependence for water contents. They range between 8 and 22 m when there is spatial dependence. Distances of dependence of this magnitude are in agreement with distances of dependence found by others with closely spaced neutron probe measurements.

Fig. 4 shows the variation in tension and water content along the transect on day 14. This figure is a clear illustration of the interdependence of tension and water content along the transect. As the tension goes down at around 28 m, the water content increases. The reverse takes place between 70 and 80 m along the transect.

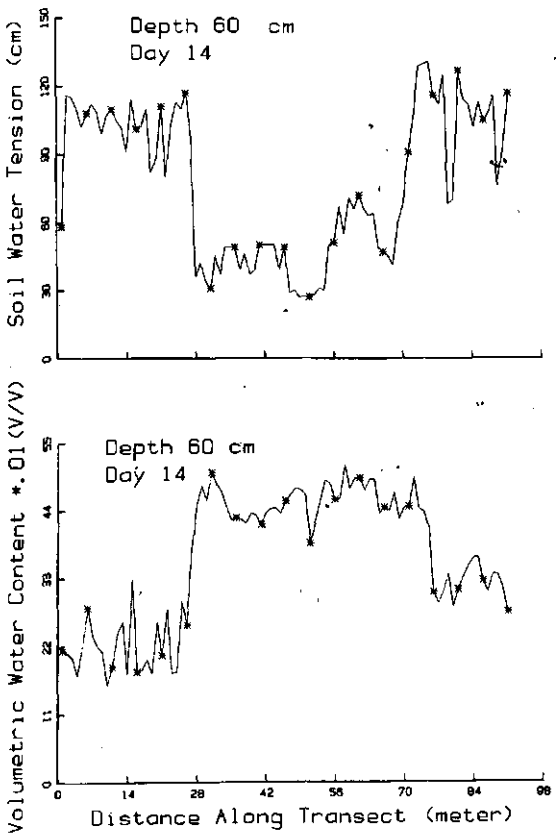


Fig. 4 Soil-water tension and water content at 60 cm as a function of distance along the transect, 14 days after flooding.

Hendrickx et al., 1985, following Webster, 1978 used the split moving window technique, to determine numerically the location

of boundaries along transects between soils with different texture or soil physical properties. Their analysis, applied to water content and texture data for the 30 cm depth along this transect, shows excellent agreement between boundaries determined on the basis of soil texture and boundaries based on soil water content. Thus, the split moving window technique is a technique that merits further evaluation for determining boundary locations along transects. However, additional work is needed to determine the influence of other parameters used in evaluating boundaries with this technique. Such parameters include vegetation, soil-water content and soil water tension at several depths, and soil chemical properties. Other techniques that could be used in analyzing data collected at regular space intervals include spatial cross correlations and/or co-kriging. These techniques provide insight into the range or spatial distance over which two variables are correlated. Such analyses are useful in predicting one variable based on observations of a second variable, taking into account the spatial dependence of each of the two variables.

Variability of field measured infiltration rates as compared to infiltration rates estimated from soil textural information

A study was conducted to compare the mean and statistical distribution of field measured infiltration rates with the mean and statistical distribution of infiltration rates estimated from soil textural information. Such information is important if one wants to utilize existing soil survey information for predicting the behavior of water in field soils. The study (Duffy et al., 1981) was conducted on a 100 ha farm at San Acacia, NM, located in the Middle Rio Grande Basin, 23 kilometers north of Socorro, NM. A detailed description of the farm layout, cropping pattern, irrigation and drainage system and water management can be found in Wierenga and Duffy, 1979. Infiltration rates were measured in the field using two methods, the ring infiltrometer method and the inverse auger hole method. With the infiltrometer method, a steel cylinder (inside diameter, 33 cm) was pressed 5 cm into the surface soil. A constant head of about 10 cm water was maintained inside the cylinder for at least 24 hours. After this period, the inflow was cut off and the rate of fall of the water table measured for a short time. The rate of fall measured over this short time was taken as the steady-state infiltration rate (K_F). The inverse auger hole method was performed for the layer just above the water table (100-150 cm), following the procedure described by Kessler and Oosterbaan (1974). The infiltrometer and inverse auger hole tests were made at 20 sites on the farm. The sites were chosen such that 2 to 4 tests were made for each soil series on the farm.

A detailed soil survey was conducted by the Soil Conservation Service. The soil was sampled at approximately 60 meter intervals and 293 bore holes were examined. A total of 11 soil mapping units and 8 soil series were identified (Duffy et al.,

1981). Based on the textural data from the 293 bore holes, infiltration rates were estimated for the surface soil and for the intermediate soil horizon (100-150 cm). Approximate relationships between the texture of a given horizon and its infiltration rate were obtained from the SCS Soil Series Interpretation forms (Form 5).

Fig. 5 presents cumulative frequency distributions for the field measured (K_F) and soil survey estimated (K_T) surface and subsurface infiltration rates. The figure shows nearly linear relationships for both the measured and estimated infiltration rates, indicating that both sets of data may be approximated by log-normal distributions. The mean ln K values were 1.0 and 0.65 for K_F and K_T, respectively, resulting in a mean measured infiltration rate of 2.73 cm/hour and a value of 1.92 cm/hour for the mean infiltration rate estimated from the soil survey

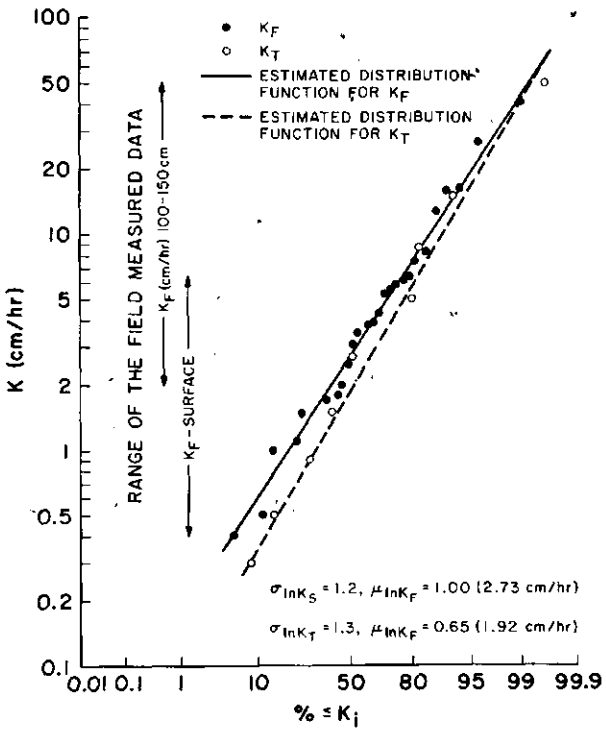


Fig. 5 Cumulative frequency distributions of the measured and texture estimated infiltration rates.

data. The standard deviations of the log transformed measured and estimated infiltration rates were 1.2 and 1.3, respectively.

Thus, it appears that for the soil in this study, the variations in infiltration rate can be estimated from soil texture data with a reasonable degree of accuracy.

In Fig. 6A the geometric mean infiltration rate of the surface soil ($\overline{K}_F$) determined for each soil series is plotted versus the geometric mean infiltration rate ($\overline{K}_T$) obtained for each soil series from the texture data. Regression of $\overline{K}_F$ on $\overline{K}_T$ resulted in:

$$\overline{K}_F = 0.57\ \overline{K}_T + 0.45 \qquad R^2 = 0.94$$

Fig. 6B shows the same for the subsoil infiltration rates (100-150 cm layer). For the subsoil, the relationship between $\overline{K}_F$ and $\overline{K}_T$ is:

$$\overline{K}_F = 0.42\ \overline{K}_T + 4.66 \qquad R^2 = 0.86$$

The data in Fig. 6 show a fair agreement between measured and estimated infiltration rates. Thus, for the soils of this study

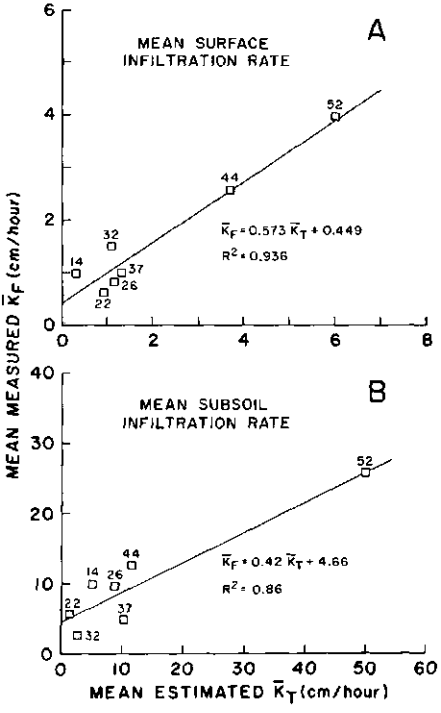


Fig. 6 Mean surface (A) and subsoil (B) infiltration rates for seven soil series (identified by numbers) on a 100 ha farm, versus mean infiltration rates estimated from soil texture data.

soil survey information may be used to determine field infiltration rates, provide a calibration is made between measured infiltration rates and those obtained from soil survey information. The variability in field infiltration rate on the other hand may directly be estimated from texture data.

Summary

The data presented in this paper have shown spatial dependence of tension and water content in irrigated soils. It appears that the degree of dependence is strongly influenced by the method of water application, time after irrigation or rain, and the presence or absence of vegetation. For example, spatial dependence was less than 1 m when water was applied to a crop through a trickle line, but was 6 m after rain or flooding. In bare soil spatial dependence, measured with tensiometers at a 1 m spacing, was up to 15 m. Spatial dependence of water content measured with a neutron meter at 1 m intervals along a transect, was 20 m or less. The variance in tension was always greater than the variance in water content, and increased as the soil dried out and the tension increased. By comparing field measured infiltration rates with infiltration rates obtained from soil survey information, it was shown that soil survey information may be used to determine the statistical properties of infiltration rate on a field scale. Thus, for the interpretation of data on spatial variability of soil properties, use should be made of soil survey information.

References

Duffy, C., P. J. Wierenga, and R. A. Kselik. 1981. Variations in infiltration rate based on soil survey information and field measurements. N. M. Agric. Exper. Station Bulletin 680. pp. 40.

Gajem, Y. M., A. W. Warrick, and D. E. Myers. 1981. Spatial dependence of physical properties of a typic torrifluvent soil. Soil Sci. Soc. Amer. J. 45:709-715.

Gumaa, G. A. 1978. Spatial variability of in situ available water. Ph.D. Dissertation. 140 pp. Univ. of Arizona, Tucson (Available as 78-24365 from Xerox University Microfilms, Ann Arbor, Michigan).

Hajrasuliha, S., N. Baniabbassi, J. Metthey, and D. R. Nielsen. 1980. Spatial variability of soil sampling for salinity studies in southwest Iran. Irrig. Sci. 1:197-208.

Hendrickx, J. M. H., P. J. Wierenga, M. S. Nash, and D. R. Nielsen. 1985. Boundary location from texture, soil moisture and infiltration data. Soil Sci. Soc. Amer. J. (submitted).

Jury, W. A. 1984. A review of published studies of spatial variability of soil water and chemical parameters. EPRI report. Palo Alto, Cal.

Kessler, J. and R. J. Oosterbaan. 1974. Determining hydraulic conductivity of soils. Drainage principles and applications.

Vol. III. Institute for Land Reclamation and Improvement. Wageningen. :254-296.

Kies, B. 1981. Solute transport in unsaturated field soil and in ground water. Ph.D. thesis. New Mexico State University, Las Cruces, NM.

Marthaler, H. P., W. Vogelsanger, F. Richard, and P. Wierenga. 1983. A pressure transducer for field tensiometers. Soil Sci. Soc. Amer. J. 47:624-627.

Matheron, G. 1965. Les variables regionalisees et leur estimation. Masson, Paris. pp. 306.

Nash, Maliha. 1984. Variation in tension, water content and drainage rates along a line transect. MS Dissertation. New Mexico State University.

Nielsen, D. R., J. W. Biggar, and K. T. Erh. 1973. Spatial variability of field-measured soil-water properties. Hilgardia 42:215-259.

Peck, A. J. 1983. Field variability of soil physical properties. Advances in Irrigation. Volume 2:189-221.

Rendu, J. M. 1978. An introduction to geostatistical methods of mineral evaluation. South African Institute of Mining and Metallurgy, Johannesburg, South Africa. 84 pp.

Russo, D. and E. Bresler. 1981. Soil hydraulic properties as stochastic processes: 1. Analysis of field spatial variability. Soil Sci. Soc. Amer. J. 45:682-687.

Saddiq, M. H., P. J. Wierenga, J. M. H. Hendrickx, and M. Y. Hussain. 1985. Spatial variability of soil water tension in an irrigated soil. Soil Science (in press).

Vieira, S. R., D. R. Nielsen, and J. W. Biggar. 1981. Spatial variability of field measured infiltration rate. Soil Sci. Soc. Amer. J. 45:1040-1048.

Warrick, A. W., D. E. Myers, and D. R. Nielsen. 1985. Geostatistical methods applied to soil science. Advances in Soil Science Vol. 1 (in press).

Webster, R. 1978. Optimally partitioning soil transects. J. of Soil Science. Vol. 29:388-402.

Webster, R. and H. E. de la Cuanalo. 1975. Soil transect correlograms of North Oxfordshire and their interpretation. J. of Soil Science 26:176-194.

Wierenga, P. J. and C. Duffy. 1979. Impacts of irrigated agriculture on water quality in the Rio Grande below Albuquerque. Engineering Experiment Station Report. New Mexico State University, Las Cruces.

Yeh, T. C. J., L. W. Gelhar, and A. L. Gutjahr. 1985. Stochastic analysis of unsaturated flow in heterogeneous soils: Part 2. Statistically anisotropic media. Water Resources Res. (in press).

Discussion

D. Elrick: Firstly, a comment on field capacity. Although the concept of field capacity is very useful agronomically, it has no physical basis in a well-drained soil; i.e., equilibrium is never reached within the time period of a couple of days.

The discussion regarding 1/3, 1/10 or 0.06 bar can be resolved somewhat by remembering that the 1/3 bar measurement is usually carried out on medium to fine textured soils that have been air-dried and sieved and then equilibrated on a pressure plate. For coarse textured soils, the tension is reduced to 0.1 bar. And, at all times, the nebuluous concept of "field capacity" must be kept in mind. I was amazed to see such good agreement between your experimentally measured K_s values and those estimated from tables given by the SCS. I expect this is because of the rather structured nature of the soil. Would you expect such good agreement in a highly structured clay loam soil for example?

P. Wierenga: The reason I brought up field capacity is that many agronomists still equate field capacity with 1/3 bar, expecting a field soil to drain to 1/3 bar in 2 or 3 days. As my data, and those of others, show this is clearly not the case. Furthermore, in determining available water, 1/3 bar is often used as the lower limit, without taking into consideration how the water content at 1/3 bar was determined. This is not correct. For most purposes I would recommend to determine field capacity in the field, using tensiometers, rather than to determine field capacity in the laboratory on core samples or sieved samples.

The agreement between the two methods for determing the hydraulic conductivity is indeed quite good. One reason is that the survey was performed by an experienced soil surveyor from the SCS. I don't know if such good agreement would be obtained in a highly structured clay loam. However, it is certainly worth investigating.

D. McComack: In the past 30 to 50 years, many decisions have been made in the limits of soil map units and their range in properties. These decisions have been made without kriging. But, using these

decisions, soil surveys have been made on more than 900 million ha in the U.S. Do you think there are feasible ways to use kriging to make better or more useful definitions of soil map units for soil surveys?

P. Wierenga: There are undoubtedly better and more useful ways to make definitions of soil map units. Whether kriging is one of these ways needs to be investigated. In my opinion kriging, and other approaches discussed at these meetings, need to be given serious consideration for use in soil survey and soil mapping.

B. Sisson: Field capacity is a useful and quantitative concept. See Wilcox (1959-1960) Canadian Journal of Soil Science.

We need some method for identifying data points that have a large influence in the estimation of stochastic parameters.

P. Wierenga: If you mean by this that we need more complete sets of field data, I agree with you.

R. Lal: Are there differences in variability between moisture content and moisture potential? For some African soils the field capacity occurs as low as 0.03 bar. The wilting point also occurs at low tension, i.e., 2 bar. The field capacity and other physical properties must be measured in the field. We find similar variability in neutron probe calibration even over short distances of 1 to 2 meters.

P. Wierenga: There appears to be a greater variation in soil water tension measured with tensiometers, than in water content measured with a neutron probe. This may partly be related to the size of soil sample measured with these two instruments. I agree with your comments on field capacity.

L. Stroosnyder: Does the fact that we have a new tool to measure soil water tension justify the additional studies recommended by you?

P. Wierenga: I have suggested that there is a lot of work to do with regard to the variability of soil water tensions in the field, and also with regard to the relation between soil water tension and yield. Up to now, few detailed studies of the variability of tension in the field have been

done. Data on field variability of tension is necessary for a host of applications. The tool used in this paper will allow us to do such studies. Studies for determining the optimum soil water tension for maximum yield are also necessary, especially with regard to trickle irrigation. The transducer tensiometer is a better tool for this.

B. Clothier: What was the trickle emitter discharge rate? Did it cause a significant surface pond of free-water to develop? If so, was some of the variation recorded by the tensiometers more a question of surface topography than of variation in soil physical properties? Maybe this would help account for the lack of spatial structure. Finally, given all this, where do you place a tensiometer to schedule irrigation by drip emitters?

P. Wierenga: We used bi-wall tubing with a low discharge rate. The tubing was buried 10 cm below the soil surface. Thus, there was no ponding on the soil surface and surface topography had little to do with the variation in tension. We placed our tensiometers such that the cup was at 30 cm below the soil surface and nearly below the drip line. For crops other than chile peppers, one might want to place them deeper. Optimal placement of tensiometers depends on the crop, soil type, tension to be maintained in the soil and other factors.

I. Murarka: Moisture distribution measurements are presumably reflecting the effects due to spatial variability of physical properties. Do you think it would be appropriate to simultaneously measure these physical properties and then use them as covariables in the analysis of moisture data?

P. Wierenga: Yes, this would be possible, but requires more work.

J. Dane: In one of your graphs you showed a large jump in tension along the transect and a corresponding jump in water content. Can you use all these data to construct semi-variograms or are you actually taking observations from 2 different populations?

P. Wierenga: You are probably right in saying that data were taken from different populations. Using the split moving window technique, one obtains well defined boundaries along the transect. One could use this technique to distinguish segments with clearly different physical properties and then construct semivariograms for each segment. Another approach would be to first detrend the data and construct one semivariogram for the entire transect.

J. Bouma: You mention that ranking of tensiometer results remains often the same upon wetting and drying. Is this correlated with differences in texture or other soil properties? This could give a relation with soil survey?.

P. Wierenga: I did not say that ranking of tensiometer results remains the same. I said that Vachaud and Warrick have observed this for water content. It would be nice if this held true for tensions but this needs to be investigated. I would expect it would be true and, if so, there should be clear correlation with soil survey data.

Soil variability and soil survey

J. Bouma, Soil Survey Institute, P.O. Box 98, 6700 AB Wageningen, the Netherlands

Introduction

Soil survey has been engaged in defining the spatial and temporal variability of soils in landscapes since the early part of the century by distinguishing different soil types and associated soil survey interpretations (e.g., Arnold, 1983). Delineated areas on the soil map are named after well defined soil series which are assumed to occupy around 80% of the delineated areas (Soil Survey Staff, 1951). This composition of mapping units on the soil map has been investigated by many researchers (e.g. Wilding, 1983). Application of (geo)statistical techniques in soil survey is and remains a crucial activity because the degree of detail in map legends as well as in soil-map interpretations should be restrained by the spatial variability of soil properties in the field. The purpose of this paper, therefore, is to: (i) analyse sources of variability and means to reduce them; (ii) compare soil delineations as made by soil surveyors to those obtained by interpolation of point data and (iii) discuss future developments. The discussion will be focused on the use of soil survey data for practical applications, emphasizing both land evaluation for actual and for potential conditions to be realized by soil and water management. Definition of land potentials for future applications appears to be particularly relevant at this time (e.g. Bouma, 1984). An attempt is made to avoid repetition of reviews presented elsewhere.

Sources of variability

Method selection

Spatial variability is determined by subjecting experimental or observational data to (geo)statistical procedures. Those procedures are discussed in statistical text-books and in other papers of this workshop. Attention will therefore be focused here on the methods and procedures by which data are obtained. This aspect appears to have been somewhat neglected so far. When discussing methods, emphasis is most frequently placed on technical aspects in terms of the type of equipment to be used and associated calculation procedures. Questions as to whether certain methods can be universally used or only in a limited number of soils, have received relatively little emphasis. The same can be said about choosing optimal sample sizes. Operational aspects of the various methods receive usually less emphasis than the technical ones.

Substantial variability can originate from using the wrong method at the wrong time, at the wrong place or by applying, for instance, a complicated technical procedure, even though only relatively untrained personnel is available. Some methods use complicated calculation procedures including substantial error even when applied professionally (e.g. Vachaud, 1982). Others yield data directly. A qualitative review of sixteen methods for measurement of the hydraulic conductivity of saturated soil (K_{sat}) and of eleven methods for measurement of K of unsaturated soil (K_{unsat}) was presented by Bouma (1983), emphasizing aspects such as: (1) time needed for preparation, execution and calculations; (2) costs of personnel and materials; (3) complexity and (4) accuracy.
Seven arbitrarily selected examples will be discussed to illustrate the relevance of method selection:

Example 1: The auger-hole method is widely and successfully used to measure K_{sat} below the water-table in sandy soils. After emptying an auger hole, the velocity by which water re-enters the hole

is measured and K_{sat} is calculated with a calculation procedure, based on a sand-flow model. Flow in bi-porous soils, containing water conducting macropores and a very slowly permeable matrix, cannot be well characterized by this method. For example, Bouma et al. (1979) reported an average K_{sat} of 5 mm day^{-1}, in a clay soil, as measured with the auger-hole method while the real value was 500 mm day^{-1}. K_{sat}, as measured by the auger-hole method, is governed by the degree of interception of water conducting macropores by the auger hole. Augering of the hole induces smearing of its walls which explains the low values observed. Smearing may not be complete, however, and thus an extra variability factor is introduced as a function of the variable number of intercepted macropores.

Example 2: The double-ring infiltrometer is widely used to measure infiltration rates. The second ring is applied to avoid lateral flow of water from the inner ring. The method works well in sandy soils (e.g. FAO, 1979). However, problems occur in bi-porous soils with cracks where water runs away laterally. As a result, very high infiltration rates are measured in the inner tube. Such rates are much higher than rates occurring in an entirely flooded field. The location of the infiltrometer rings is crucial as it determines how many cracks or other macropores are present in the surface of infiltration. Placement, as such, is therefore a major source of variability associated with the method being used and not with the physical characteristic being investigated.

Example 3: The double-tube method for measuring K_{sat} (Bouwer, 1962) works well in sandy soils, even though the amount of required labor is very high. The method is based on lateral movement of water from the outer-into the inner tube. This movement can occur in soils with cracks, but not in soils with well defined strictly vertical or horizontal macro-pores. A K_{sat} of 20 cm day^{-1} was measured with the double-tube method in a silt loam soil in Wisconsin, while sprinkling irrigation with rates of up to 120 cm day^{-1} did not produce surface ponding of water. Obviously, the

K_{sat} of 20 cm day^{-1} did not characterize the real flow system. The rapid infiltration rate in the field was possible because of worm channels extending into a sandy subsoil. These worm channels did not contribute towards flow from the outer to the inner tube.

Example 4: Sometimes questions have to be answered that cannot be solved by applying existing methods because they are unsuitable. If, for some reason, existing methods are used, questionable results are obtained. In our work, we encoutered this problem when asked to measure the vertical K_{sat} of an indurated spodic horizon. Use of sampling cylinders or infiltrometers would have resulted in fracturing of the horizon, making measurement results irrelevant. It was decided to carefully chip out an in-situ column and encase it in gypsum. Then, the steady infiltration rate in the column was measured while pressure heads were registered simultaneously in soil above and below the spodic horizon (Dekker et al., 1984). Fortunately, there was time available in this particular project to develop a new method. Often, this time is not available during contract work.

Example 5: This example does not focus on the use of a particular method, but, rather on the occurrence of the phenomenon of "bypass flow" (earlier called: "short-circuiting"). This process describes vertical movement of "free" water through an unsaturated soil matrix. As such, the process is not discussed in current soil physics text-books. It is important in many soils and it affects results of measurements and adds to the observed variability when methods are used that assume presence of homogeneous soil. Methods have been developed to measure bypass flow (e.g. Bouma et al., 1981). Recognition of this phenomenon may help to explain what appear to be erratic, highly variable measurement results at first sight.

Example 6 : Various methods for measuring soil permeability require in situ measurement of moisture contents and pressure heads.

Variability associated with applying the neutron probe was discussed by Vachaud (1982). An extra variability factor is introduced when this method is applied in bi-porous soils where neutrons intercept widely spaced water-conducting pores in an unpredictable manner. Effects of using different sizes of tensiometer-cups on measured pressure heads were reported by Bouma et al. (1982). Large cups intercepted macropores, small cups did not. Observed pressure heads differed accordingly. The suggestion was made to excavate tensiometer cups after experiments and to describe pore patterns in surrounding soil, possibly by using dyes. Thus, sub-populations of data may be distinguished.

Example 7: Different methods are often available to measure the same soil characteristic. In contrast to examples 1, 2 and 3 where some methods produced incorrect results, different methods being considered in this example produce good results. Still, data are different and it is important to define one's purpose in making measurements when selecting one of the methods. Besides, a significant source of variability is created when different methods are concurrently applied. An example is taken from the work of Dr. P. Stengel (INRA, Avignon, France) who measured soil porosity with two different techniques (Fig. 1). Data in Figure 1A were derived from in-situ measurements determining the volume of a balloon being filled inside a small soil excavation. Data in Figure 1B were derived from large, undisturbed cores of approximately one liter content. The higher values in Figure 1A were due to irregular extensions of the balloon into cavities next to the excavation, which did not occur in the rigid cylinder.

The above examples can be extended to other methods, in demonstrating that certain physical methods cannot be applied in all soils. Soils with macropores or with strongly contrasting soil horizons impose, for example, drastic boundary conditions for the flow system. In many projects researchers are forced to obtain data within a limited period of time, as specified in a contract. From a scientific point of view *lack* of data is preferable to *incorrect*

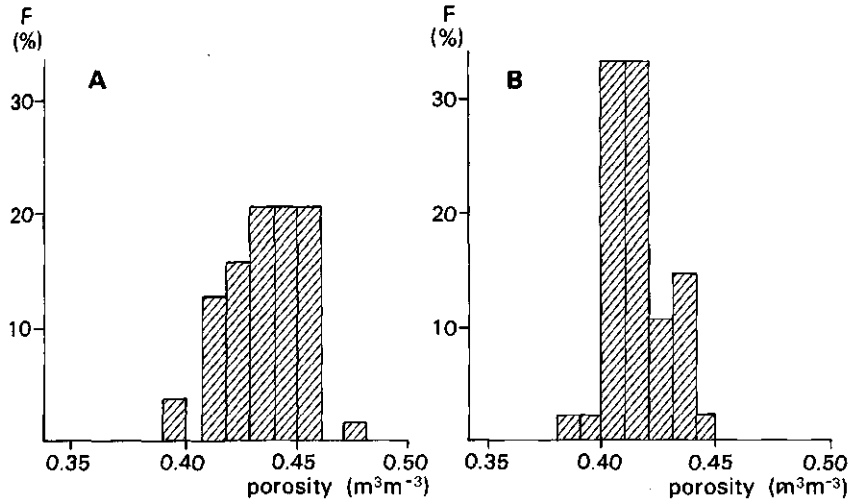


Fig. 1 Comparison of soil porosities obtained with two methods. Method A used the volume of a balloon being filled inside an excavation. Method B used large, undisturbed cores. F is the frequency of measured porosity classes (data courtesy of Dr. P. Stengel, INRA, Avignon, France).

data. Unfortunately, in the real world the opposite appears sometimes to be true.

Sample location

Sampling at regular depth intervals is often applied with good results in relatively homogeneous soils with weakly developed soil horizons. When clear soil horizons exist, however, it is preferable to sample by horizon (e.g. Peterson & Calvin, 1965). A sample containing fragments of two adjacent, and as such quite different, soil horizons, will yield data that are hard to interpret. However, it should be realized that pedological horizons as distinguished in soil survey, are not always good "carriers" of data that are relevant for the particular interpretation being pursued. Some pedological distinctions may be irrelevant in this context, while relevant aspects may not be reflected in the horizon classification. For example, when determining hydraulic conductivity

curves for subsoil soil horizons in an area of 125 ha in the Netherlands, we found three types of significantly different curves, while seven different pedological soil horizons had been distinguished. Of course, such a finding does not present a problem here because it is attractive to have fewer types of curves to work with (Wösten et al., 1985). A problem would arise when quite different types of curves would occur within one type of pedological soil horizon. Then, obviously, soil survey data on soil horizons would be inadequate to serve as a "carrier" of physical data. The fact that this particular problem did not occur in the study cited above, does not imply that it could not exist elsewhere.

Sample size

Many measurement procedures use standard sample sizes, because of fixed dimensions of sampling cylinders or of equipment being used. For example sampling cylinders with a fixed volume of 100 cm^3 have been used extensively in different laboratories. Equipment, such as the double-tube or air-permeameter, comes in standard sizes. There is good justification to vary sample size as a function of soil structure, as a means to reduce variability among replicate measurements (e.g. Bouma, 1983). Soil structure descriptions can be used to tentatively define representative elementary volumes of samples (REV's), which are the smallest sample-volumes that can represent a given soil horizon by producing a consistant population of data. To do so, the elementary units of soil structure (ELUS) have to be distinguished. These are individual sand grains in sandy soils and natural aggregates ("peds") in aggregated soils. Peds can vary in size up to several liters each in very coarse, prismatic subsoil structures. Even though emphasis in soil structure descriptions is often placed on the solid phase in terms of soil grains and peds, real emphasis should be on the nonsolid phase where transport processes take place. Of course, by describing grains and peds, information is also provided about the pores between them. In addition, pores that do not result from the packing

of grains or peds, should be considered separately. Such pores are, for example, root and worm channels with a cylindrical shape. As a general rule we have proposed that REV's should contain at least 20 ELUS but preferably more, or that any sample taken should have a representative number of channels per unit surface area. When applied literally, samples can become very large (e.g. table 1). If so, selective sampling of structural elements or soil fragments that contain representative quantities of pores and peds, should be considered. In any case, a structural description should be made. When worm channels occur in a soil, they may have such a large effect on soil hydrology that measurements of infiltration rates into individual channels may be a better procedure than measurements in a given area (e.g. Bouma et al., 1982).

A very tentative classification has been proposed for sample sizes in four broad textural classes (Table 1) (Bouma, 1983).

Table 1. Hypothetical Representative Elementary Volumes (REV's) of soil samples as a function of soil texture and structure. Selective sampling of separate soil structural features must be considered when samples become very large.

Class	Texture	Structure	Hypothetical REV (cm^3)
a	sandy	no peds	10^2
b	loamy	small peds	10^3
c	clayey	medium peds continuous macropores	10^4
d	clayey	large peds continuous macropores	10^5

Defining REV's as a function of field descriptions of soil structure needs to be further investigated. Data by Anderson & Bouma (1973) illustrate the potential of the procedure (Table 2). Unrealistically high K_{sat} values were measured in soil cores containing fewer than twenty ELUS. Values measured with a gypsum-covered column having a volume of 12 liters, averaged 70 cm day^{-1}.

The reason for the high K_{sat} values in the small cores is the high and unnatural vertical continuity of cracks between the peds in small samples.

A second example was presented by Bouma et al., 1979. They measured K_{sat} of a heavy clay soil in large, gypsum covered, samples of 16 liters. Thus, a population of data was obtained that allowed the statistically significant conclusion that K_{sat} had increased as a result of tile drainage. Use of small cores produced a highly variable population of data that did not allow such a conclusion.

Relationship between REV's and ELUS are currently investigated in a joint project between the Netherlands Soil Survey Institute and the Agronomy Department of Cornell University, USA.

Table 2. Measured hydraulic conductivity values of saturated soil (K_{sat}) in soil cores of varying height but with a diameter of 7.5 cm, containing different numbers of elementary units of structure (ELUS). Measurements were made in a silt loam soil with medium sized peds with an average volume of approximately 30 cm^3. The largest sample was a gypsum-covered column of soil with a diameter of 30 cm (derived from Anderson & Bouma, 1973).

Sample Volume (cm^3)	ELUS (no)	K_{sat} (cm day^{-1})	S (cm day^{-1})
230	8	650	350
330	11	320	320
460	15	100	80
780	26	75	30
12 000	400	70	20

The foregoing discussion is summarized in Figure 2 which emphasizes the need to use soil survey information, when selecting sites and soil profile characteristics, when selecting methods and when taking samples.

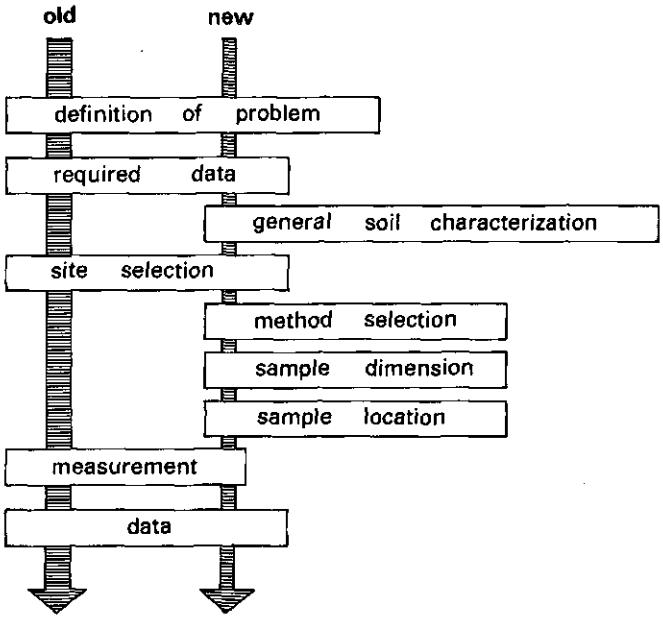


Fig. 2 Diagram illustrating site selection and measurement of data without using soil survey information (old procedure) and a new procedure that does consider this information (see text).

Soil maps as indicators of spatial variability

Introduction

So far, the discussion was focused on ways in which data are obtained. Now attention will be focused on use and interpretation of data.

Spatial variability studies in relation to soil mapping have been made or reviewed by several authors (e.g. Wilding, 1983; Nielsen et al., 1983; Webster & Burgess, 1983; Burrough, 1983). In this paper attention is focused on using soil survey for land evaluation and, more specifically, on the following points:
(1) How do predictions based on a soil-map legend compare with interpolations (e.g. kriging) using point data,

(2) What could be the future role of (geo) statistics in soil survey.

Determination of the spatial variability of pedological properties within delineated areas of the soil map is of interest but particular emphasis is now increasingly placed on the interpretation of soil maps in terms of land evaluation. In this context, characteristic land qualities are emphasized such as: moisture availability, trafficability, aeration-status etc. (FAO, 1976).

Van Kuilenburg et al. (1982) studied the spatial variability of a calculated soil moisture supply capacity by comparing estimates derived from the soil map with values obtained by applying three interpolation techniques to point data. The area studied was 4 km² and contained sandy soils. In total, 530 survey borings were made and 661 independant test borings that were used to judge soil survey interpretations and results of interpolations. A root mean squared error of 32 mm of water was obtained when estimates were based on representative soils of the various mapping units, while this error was 29 mm when kriging of point data was applied. Moisture supply capacity ranged from 50 to 240 mm; indicating a relatively high error for both methods. In this particular example, application of kriging hardly led to better results than interpretations based on the soil map. The basic conclusion was that the land quality "moisture supply capacity" was quite variable. This aspect should be reflected in the legend of the map.

Another study was made in which emphasis was placed on the distinction of major soil horizons in an area of cover-sands, overlying boulder-clay. These horizons were first distinguished by the usual pedological criteria. Next, hydraulic conductivity (K-h) and moisture retention curves (θ-h) were measured sixfold. Horizons that had statistically identical properties were grouped together, reducing the total number of different horizons that could be distinguished (Wösten et al., 1985). This procedure does not focus on a land quality, but, rather, on a land characteristic, which is a property that can directly be measured (FAO, 1976). The two physical properties, being distinguished, are, of course, impor-

tant for simulation programs for the soil water regime which is, in turn, crucial for most land qualities being distinguished for land evaluation (e.g. Bouma, 1984). As stated, horizons were defined that had identical K-h and θ-h properties. These were directly attributed to the delineated areas on the soil map, providing a means to predict these data for any location in the study area by extrapolation. Simultaneously, interpolations with the kriging technique were used, based on point data as obtained by borings during the soil survey. The comparison of both procedures was based on 60 independent borings that were made at random using a method to be described by De Gruyter in the next paper. Results, reported in detail by Bregt & Bouma (1985) showed that application of kriging resulted in less accurate predictions than the ones obtained by the interpretation of the soil map.

Obviously, more of such studies are needed to evaluate the potential of both traditional soil survey and of modern interpolation techniques. Care should be taken to not ignore the potential of available procedures by focusing completely on new interpolation techniques. In fact, integrating both procedures would appear to be most attractive. In contrast to many geological applications, differences among soils in a landscape are often associated with visible landscape features at the surface and with vegetational patterns. These features are reflected on soil maps and should be considered when defining spatial variability. Soil maps could therefore be used to define clearly different subpopulations of soils within an area to be characterized on the basis of landscape or other visible features. Use of statistical sampling techniques within each of these subpopulations could function to define their internal properties in a quantitative manner.

As is, the application of (geo)statistical techniques in soil survey is often a check of the existing map-legend, testing the delineated areas on the soil map. A more realistic use would be an application before soil mapping because then the proper degree of detail of the legend to be chosen for the particular area being

considered, can be defined on the basis of the variability being observed (see Fig. 3).

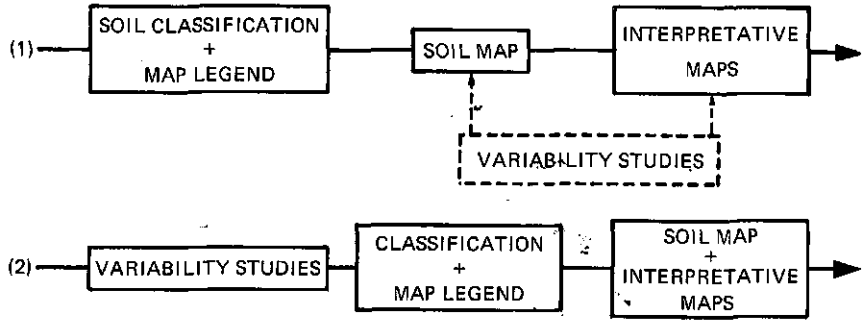


Fig. 3 Diagram illustrating some actual variability studies which focus on existing maps (procedure 1). A more promising procedure 2 uses variability studies to define legends of maps with a known variability.

Soil survey and (geo)statistics in future

The above discussions, which cover the role of soil survey information when studying the spatial variability of soil properties, relate to future work and can be summarized as follows for discussion purposes:

(1) Data on soil variability are derived from multiple measurements. Too little emphasis is being placed now on selection of methods and sampling locations, using soil survey data. The diagram in Figure 2 illustrates the proposed procedure. However, more case-studies are needed to show the usefulness of soil structure descriptions for determining optimal sample sizes and of soil maps and soil horizon descriptions for determining optimal sample locations and depths. Geostatistical studies are needed in addition, to define the optimal distance between multiple samples.

(2) Aside from studies that examine the pedological purity of mapping units in the field, attention should be paid to the variabil-

for land evaluation. Sometimes, pedological distinctions are made that are irrelevant for certain interpretations. This offers no problems, since the number of distinctions can easily be reduced to include only those that are relevant from a functional point of view. A problem occurs when a single pedological distinction has more than one interpretative meaning. Then, obviously, pedological data are less useful as "carriers" of interpretative information. So far, we know far too little about this "carrier" function of pedological data for various interpretations and its variability.

(3) Soil maps are made by soil surveyors who use their best judgment in drawing boundaries between mapping units. Those boundaries can also be obtained by using modern interpolation techniques between point data. More studies should be made comparing both procedures. This can only be done by independant test borings. A mixture of both procedures could perhaps be interesting for future work. Certain landscape features are so clearly evident that (geo) statistical techniques are not needed to distinguish them. Hence, these features can be used to define subpopulations in terms of areas which can be characterized by considering point data within these areas.

(4) Sometimes, soil survey appears to be particularly focused on map legends and soil classification schemes, which are, of course, only intended to be a means towards a purpose and not a purpose in itself. The key function of variability studies in future could be the definition of optimal map legends, as illustrated in Figure 3. Completion of country-wide surveys reduces the need for continued use of a unified legend. Legends should be composed for any particular area and possible application. Availability of soil information systems which allow flexible handling of basic data, is necessary for this type of approach.

References

Anderson, J.L. & J. Bouma, 1973. Relationships between hydraulic conductivity and morphometric data of an argillic horizon. Soil Sci. Soc. Amer. Proc. 37: 408-413.

Arnold, R.W., 1983. Concepts of soils and pedology. In: Pedogenesis and Soil Taxonomy. L.P. Wilding, N.E. Schmeck and G.F. Hall (eds.): Developments in Soil Science IIA. Elsevier, Amsterdam: 1-21.

Bouma, J., 1983. Use of soil survey data to select measurement techniques for hydraulic conductivity. Agric. Water Managem. 6 (2/3): 177-190.

Bouma, J., 1984. Estimating moisture related land qualities for land evaluation. In: Land Use Planning Techniques and Policies. SSSA Special Publication no. 12. Am. Soc. of Agronomy, Madison, Wisc., USA: 61-76.

Bouma, J., L.W. Dekker & J.C.F.M. Haans, 1979. Drainability of some Dutch clay soils: a case study of soil survey interpretation. Geoderma 22 (3): 193-203.

Bouma, J., L.W. Dekker & C.J. Muilwijk, 1981. A field method for measuring short-circuiting in clay soils. J. Hydrol. 52 (3/4): 347-354.

Bouma, J., C.F.M. Belmans & L.W. Dekker, 1982. Water infiltration and redistribution in a silt loam subsoil with vertical worm channels. Soil Sci. Soc. Am. J. 46 (5): 917-921.

Bouwer, H., 1962. Field determination of hydraulic conductivity above a water-table with the double-tube method. Soil Sci. Soc. Am. Proc. 26: 330-335.

Bregt, A. & J. Bouma, 1985. Spatial patterns of soil physical characteristics as derived from soil survey and from kriging of point data. Soil Sci. Soc. Am. J. (in press).

Burrough, P.A., 1983. Problems of superimposed effects in the statistical study of the spatial variation of soil. Agr. Water Managem. 6: 123-145.

Dekker, L.W., J.H.M. Wösten & J. Bouma, 1984. Characterizing the soil moisture regime of a Typic Haplohumod. Geoderma 34: 37-42.

FAO, 1976. A framework for land evaluation. FAO Soils Bulletin 32. Rome, Italy.

FAO, 1979. Soil survey investigations for irrigation. FAO Soils Bulletin 42. Rome, Italy.

Nielsen, D.R., P.M. Tillotson & S.R. Vieira, 1983. Analyzing field measured soil-water properties. Agr. Water Managem. 6: 93-111.

Petersen, R.G. & L.D. Calvin, 1965. Sampling. Chapter 5 in: Methods of Soil Analysis. Part 1. C.A. Black (ed.) Am. Soc. of Agronomy. Agronomy Series no. 9: 54-73.

Soil Survey Staff, 1951. Soil Survey Manual. Us. Dept. Agriculture Handbook no. 18. U.S. Dept. Agr., Washington.

Vachaud, G., 1982. Soil physics research and water management. Whither Soil Research. Proc. ISSS Congress, India: 32-59.

Van Kuilenburg, J., J.J. de Gruijter, B.A. Marsman & J. Bouma, 1982. Accuracy of spatial interpolation between point data on soil moisture supply capacity, compared with estimates from mapping units. Geoderma 27 (4): 311-325.

Webster, R. & T.M. Burgess, 1983. Spatial variation in soil and the role of kriging. Agr. Water Managem. 6: 111-123.

Wilding, L.P. & L.R. Drees, 1983. Spatial variability and pedology. In: Pedogenesis and Soil Taxonomy. L.P. Wilding, N.E. Schmeck and G.F. Hall (eds.): Developments in Soil Science IIA. Elsevier, Amsterdam: 83-113.

Wösten, J.H.M., J. Bouma & G.H. Stoffelsen, 1985. The use of soil survey data for regional soil water simulation models. Soil Sci. Soc. Am. J. (in press).

Discussion

M. Nash: What is the size of the delineated areas on the soil map once the soil variability has been determined for different soil properties by geostatistical techniques?

J. Bouma: From a purely cartographic point of view, the minimum size of delineated areas is a function of map scale: very small areas cannot be observed. More important is the observation that different soil properties (e.g. % clay; pH; % org. matter) may have different ranges. So which one is to be used? It is suggested to use the spatial characteristics of the land quality that is considered to be most relevant for the study concerned. In my paper it was the land quality: moisture supply capacity, which was characterized by different land characteristics.

K. Flach: The minimum size of the mapping unit depends on the purpose of the soil survey. Map scales must be adjusted accordingly, and they should never determine minimum size.

L. Stroosnyder: The quantification of soil surveys and their interpretations is indeed necessary and quite useful, but we should not forget that to do so requires many data that are usually not available in most developing countries. More qualitative surveys are still very valuable in that context.

D. Nofziger: What operational problems are involved when measuring modeling parameters that have vastly different representative elementary volumes (REV's)?

J. Bouma: Studies on REV's using morphological soil-structure data have so far centered on measurement of the hydraulic conductivity. It is conceivable but not likely that other physical parameters would have different REV's. We don't know yet. Of course, a sample can never be too big. So it may be advisable to aim for uniform large sample sizes for a given soil horizon, even though this could represent a certain "overkill" for some parameters.

B. Clothier: The difference between the 5 mm/hr auger-hole K_s and the K_s of the gypsum-coated column of 50 cm/hr is physically important. One is dominated

by the matrix properties and the other by macropores. Consequently the conductivity appropriate to pre- or non-ponding infiltration is 5 mm/hr. Free water infiltration will be typified by 50 cm/hr. Saturated (or near-saturated) conductivity is not a unitary property, but needs to be measured appropriately for the purpose required.

Further, the difference between the two values may be more important than their absolute values. The former (matrix) property is soil-texture dominated. The latter (macropore) is controlled by management.

J. Bouma: I do not completely agree with your assessment. The K_{sat} of 5 mm/day is found because water-conducting macropores are partly closed by puddling of the walls of the borehole. Water movement still occurs along the macropores, as the matrix has a very low conductivity. The auger-hole method is used to assess drainability of soils. The measured value would lead to the (incorrect) conclusion that tile-drainage would not be feasible. However, it is widely and successfully applied, based on the real K_{sat} of 50 cm/day. When considering pre- or non-ponding conditions we need hydraulic conductivities of unsaturated soil, which can be measured with the crust test.

C. Wang: We used the auger-hole method to measure K_{sat} of clay soils after scratching the walls of the borehole. Thus, the puddling problem is overcome. We do have serious problems in soils with very fine sandy textures, where the borehole caves in.

J. Bouma: I wonder about your reference level for K_{sat}. Usually, the only way to remove the adverse effects of puddling is by drying and cracking. Also, augering the hole under dry conditions and coming back later for the measurement, may work. I stress that it is not my intention to discuss the merits of the method as such, but to use the example to illustrate the effect of method selection on variability observed. My paper gives several other examples.

H. ten Berge: Especially in soil physics and hydrology the problems of spatial variability and ERV are introduced by wanting to measure site-specific

properties like K_{sat}. I believe that the connection between soil physics on the one hand and regional soil surveys, management practices etc. on the other hand should be established through (inverse) modelling. In most surveys, hydraulic soil properties are determined to be applied in deterministic models, not to develop these models.

J. Bouma: We apply indeed our measurements in deterministic models. Basic data have to be representative so as to allow generation of realistic output data from the models. We certainly don't want to build "black-box" models. A representative K_{sat} value for a soil, obtained by a correct method in an adequate sampling volume, is, of course, considered to be input data for the model.

C. Topp: There is a need to better define soil morphology data in relation to the determination of soil physical characteristics. A problem when using soil morphology data is its qualitative character.

J. Bouma: Standard soil structure descriptions, as made in the field, are indeed qualitative in nature. Quantification has particularly been developed in micromorphology using thin sections. Staining tests and drawing of ped faces and macropores on transparent sheets, will allow more quantitative assessments of macrostructure. More work is needed here.

C. Topp: A major challenge for this gathering is how to bridge the gap in communication between those, such as pedologists, who of necessity operate in a qualitative domain and those, such as physicists, who require or desire more quantitative information. J. Bouma has suggested transfer functions as one mechanism. The calibration or definition of these transfer functions is very labor intensive and I ask are there other mechanisms which require less labor? I suggest that one other mechanism is for physicists to operate with pedologists to calibrate and quantify the pedologists well-developed observation capabilities.

J. Bouma: I doubt whether transfer functions always require so much labor. Specific measurements must always be made. In your studies, for

example, you used the air-permeameter to measure K_{sat}. The idea of the transfer functions is to correlate them in some way with data obtained in soil survey. For example, correlations with texture are attractive in this context. But such simple correlations often don't work. In my paper I used an example in which pedological soil horizons and soil series were used as "carriers" of physical information. I believe you have done the same in your studies. This seems to be a good example of using observations by pedologists to estimate physical data. We need much more work in this area.

Transect sampling for reliable information on mapping units

J.J. de Gruijter, Institute TNO for Mathematics, Information Processing and Statistics, P.O. Box 100, 6700 AC Wageningen,
B.A. Marsman, Netherlands Soil Survey Institute, P.O. Box 98, 6700 AB Wageningen, The Netherlands.

Introduction

In soil survey there is a need for using efficient sampling methods to gather reliable information on soil conditions in mapping units. The reasons for this are discussed by e.g. Wilding (this issue) and Bouma (this issue). Wilding & Drees (1983) give a general review of sampling methods in the context of spatial variability and soil survey. We started a search for a sampling method which meets the following conditions.
1. The method must lead to unbiased estimates of the average of soil properties within mapping units.
2. The efficiency should be as high as possible, i.e. with fixed costs the accuracy of the results is maximized, or with fixed accuracy the costs are minimized. Time spent on fieldwork forms a major component of costs involved.
3. It should be possible to quantify the accuracy of the results in an objective way.
4. Field work and statistical calculations should be as simple as possible, for practical operational reasons.

Only samples of medium size (say n between 10 and 100) can be afforded in practice. With that size it would hardly be possible to properly verify the assumptions underlying regionalized variable theory. Estimation of variograms would often be problematic too. So sampling and statistical inference based on this theory would generally violate the third condition mentioned above. We confine ourselves therefore to classical sampling theory (see e.g. Cochran, 1977). Units are then selected by a procedure with some random component, so that the observations are mutually independent and accuracy can be quantified without resort to unproven assumptions. For the selection of an efficient sampling design it is imperative to use the surveyors knowledge of the spatial variation and logistics of field work in the area to be sampled. Input of the field surveyor is crucial.

Random sampling by point transects seems promising, partly because of its operational advantages. Transect sampling has been applied in soil science rather frequently during the last decennia. See e.g. Powell & Springer (1965), Steers & Hajek (1979), Wang (1982) and Bigler & Liudahl (1984).
It is not always possible to judge the methods used for sampling

and statistical analysis because no sufficient details are given. Where this is possible, however, it must in some instances be concluded that the published results are biased to an unknown degree. The main reasons for this are that starting points of transects were chosen at a fixed distance from the boundary of delineations, that observation points initially falling in 'non-soil' area were shifted, and that in calculating averages the values were not weighted according to the sampling design used.
We developed a new sampling method by which unbiased estimates are obtained. In this article we give some reasons why we expext this method to be efficient in terms of costs and accuracy. In a future article we hope to quantify its efficiency and to compare this with alternatives.
It should be noted that we concentrated on easily accessible terrains, like most rural landscapes in The Netherlands.

A method of sampling by random transects

General description of the method

Conceptually the soil universe of the mapping unit is divided into a very large number of small square cells, forming a finite population of discrete elements. Observation points are located in the centres of these cells.

As a form of two-stage cluster sampling, two mutually perpendicular transects with equidistant observation points are randomly selected from each of a number of randomly selected delineations of a mapping unit.
In the first stage the delineations act as sampling units. They are selected with replacement and with probabilities proportional to size, i.e. number of cells or area.
In the second stage the sampling units are clusters of cells in a linear configuation. Any of these clusters can be conceived as constructed by taking cells from a given row or column, going through a delineation while constantly skipping a fixed number of cells. (The skipped cells belong to similar clusters which may be selected from the same row or column). So each cluster corresponds to a possible transect with equidistant observation points. From each of the delineations selected in the first stage two mutually perpendicular transects are selected at random.

Generally these transects will cross and could then have an observation point in common. However the transects, being sampling units, have to be disjoint. Therefore we introduced an extra restriction on the way the transects are formed: the cells of a given delineation are divided into two strata, like the black and white fields of a checkboard. Now all transects along rows are composed of cells from one stratum only, while the other stratum produces the transects along columns. From each of these strata one transect is selected at random with probabilities proportional to size, viz. number of observation points. If a delineation has been selected more than once in the first stage, one pair of transects has to be selected for each time that this delineation is selected.

Fig. 1. Location of transects in the delineations of two mapping units on the soil map of Lievelde, scale 1 : 10 000.

A technical difficulty is that delineations may contain 'non-soil' area, e.g. buildings, farmyards, roads and ditches. In general such areas do not belong to the population to be investigated. Some of them may be recognizable on the map, others only in the field. Shifting of points falling in these areas according to some rule is not a good solution as this leads to a systematical oversampling of their neighbourhoods. The selection procedure described below proved to be practical while strictly satisfying the above conditions on selection probabilities.
The procedure is illustrated by Figure 1, which shows the location of transects in two mapping units of a 1 : 10 000 soil map. Figure 2 shows a detail of this map, with transects and observation points in two delineations.

Preparation of sampling frames

A rectangle is drawn around each of the delineations, with an orientation that may be chosen at random or on purpose, and with dimensions as small as possible. The lower and left sides of the rectangles function as X and Y axes respectively, with a length-unit of, say, 1 mm.
Thus each rectangle forms a local coordinate system, to be used as a sampling frame for the enclosed delineation. Overlapping of frames presents no problem. The frames are numbered and listed with their cumulative areas. (Note that the areas of the delineations need not be measured.)

First stage: selection of delineations

Step 1: Read from a table with random numbers a number which does not exceed the total area of the frames. Compare this number with the listed cumulative areas. Select the frame with the smallest cumulative area that is still larger than the random number.

Step 2: Read from a table with random numbers a pair of X and Y coordinates within the frame selected by Step 1. Both coordinates must be either even or odd. If the point in question falls in the delineation and not in 'non-soil' area recognizable on the map, select this delineation and mark the point on the map. Otherwise, drop this frame and repeat Step 1.

Repeat Step 1 and 2 until sufficient selections have been made. Recall that any delineation may be selected more than once. The number of selections will strongly influence the accuracy of results as well as sampling costs.

Second stage: selection of transects

Step 3: Use each of the random points resulting from Step 2 as the starting point in a transect parallel to the X-ax. Mark the other observation points of the transects, going in both directions from the starting point to the boundary of the delineation. The distance between adjacent points must be an even number of millime-

ters and must be constant within delineations. (Recall the checkboard device: the number must be even to stay on the same color, i.e. in cells of the same structure.) Skip any point which falls in 'non-soil' area recognizable on the map.

Step 4: For each transect resulting from Step 3, select at random another transect from the same delineation so that they form a mutual perdendicular pair. To this end, read a pair of X and Y coordinates from a table with random numbers, now under the restriction that one is even and the other odd. If the point falls in the delineation, and not in 'non-soil' area recognizable on the map, accept it as a starting point and mark it on the map. Otherwise, try another pair of random coordinates and continue if necessary, until a starting point is obtained. Mark the other observation points of the transects as in Step 3, but now parallel to the Y ax.

Field check on starting points

In order to maintain the specified selection probabilities, realizing that some 'non-soil' areas may be only recognizable in the field, special attention should be given to the starting points during fieldwork.
At the beginning of the fieldwork in a given delineation it should be checked whether the random point by which it was selected (also the starting point for the transect parallel to the X ax) lies in 'non-soil' area. If so, the delineation is dropped as yet, including its pair of transects. It is substituted by a reserve from an extra sequence of pairs selected beforehand by Step 1 to 4. If not so, the delineation and its first transect is actually included in the sample and observations along the transect are made.
Then the second transect, parallel to the Y ax, is checked. If its starting point appears to lie in 'non-soil' area, it is dropped and substituted by a reserve in the same delineation. Such reserves are to be selected beforehand by Step 4. If other than starting points lie in 'non-soil' area, these are just skipped.

Fitting to local circumstances

Several possibilities can be used, single or in combination, to fit our sampling method to local circumstances, i.e. the spatial variation expected and the objects of sampling. The major possibilities are briefly discussed below. Using them has no consequences for the way the sample data are to be analysed statistically, except for stratification (see: Statistical analysis of sample data).

Lay-out of transects

The directions of the transects are determined by the orientation of the sampling frames around the delineations. This may be chosen at random or on purpose, and is allowed to vary between delineations. These choices may affect the accuracy of the results, but not their unbiasedness. (Recall that the orientation of the frames

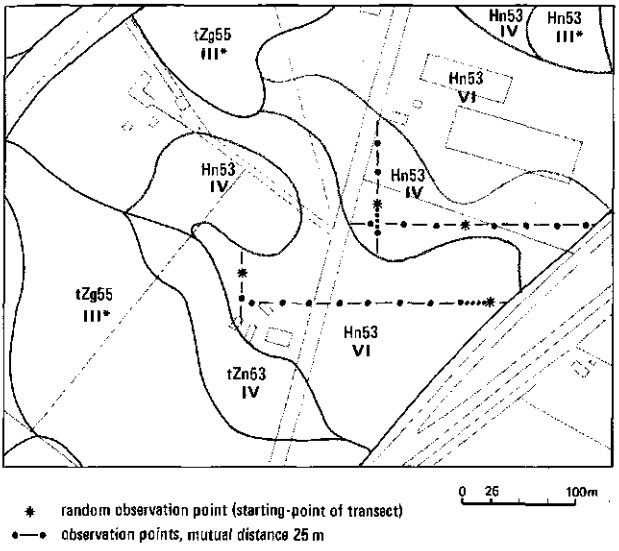


Fig. 2. Detail of soil map Lievelde with transects in two delineations.

only determines how the delineations are divided into cells and how these cells are divided into two strata).
It would be most efficient in terms of costs and accuracy to sample the delineations by transects along a trend if that existed. If the direction of a trend were known before sampling, this could be done by orientating the sample frame accordingly and selecting a pair of parallel instead of perpendicular transects. This can also be done by the above procedure, if one leaves out the restrictions related to the checkboard device.
However, the directions of possible trends common to the soil properties in question will usually not be known beforehand. In that case the directions of the transects cannot be optimized with respect to accuracy. They may still be chosen on purpose.
For instance, specific directions may have operational advantages in the field, and minimizing frame areas will speed up Step 2 and 4 of the selection procedure. The directions may also be chosen in such a way that genetic hypotheses related to trends can be tested.

Similar to the directions of transects, the distances between adjacent points must be constant within but may vary between delineations. (So far we kept them constant in our own applications for simplicity). Smaller distances lead to more observation points per transect, which increases accuracy as well as sample costs. A compromise must be found between both effects. This will depend on the specific autocorrelation structure: the closer the correlation

between observations in a transect, the less information is gained. We are preparing a publication dealing with optimization of this choice (De Gruijter and Marsman, in prep.). Our experience up till now is limited to 1 : 10 000 soil maps of coversand areas. We have the impression that a distance of 20 to 30 m is a reasonable choice for such maps.

Sampling large delineations by transects with relatively close observation points could lead to only few transects, with many points in each. This would be inefficient if variations between delineations are not very small compared to those within. To avoid this, the transects can be delimited by splitting large delineations into two or more sections. This is to be done between Step 1 and 2 of the selection procedure. The entire delineations still function then as primary sampling units, but expansion of transects from their starting points is limited to the sections in which they fall.

Stratification in the first stage

To increase the efficiency of the sampling design one may stratify in the first stage. To that end, the set of all delineations is first divided into a number of strata, e.g. according to size or location. The more homogeneous strata are formed, the more gain in efficiency is achieved.
Then the selection procedure is applied to each of the strata separately. To enable a straight-forward assessment of the accuracy, at least two delineations should be selected from each stratum. The results on the individual strata may also be interesting by themselves.

Other types of populations

Our sampling method may also be used in cases where the population to be sampled is a single area rather than a mapping unit, e.g. a parcel or some mapped area as a whole. The area should then be divided into sections to be used as primary sampling units. Rectangular or square sections have, of course, the advantage that their sides can directly be used as sampling frames. See Bregt and Bouma (in prep.) for an application in which the purity of a map as a whole is estimated.

Why this method?

Two-stage and cluster sampling are the two major features of our method. In sampling technique in general both devices are applied to reduce the costs of visiting sample units. This advantage will be achieved in the present context too. Experience showed that locating a single point in the field in a sufficiently accurate way is one of the major components in sampling costs. Transect sampling reduces this drastically because once the starting point has been located, the other points follow easily by pacing, at least if they are not too far apart. As long as the points are not too

close, we expect the mentioned advantage to outweigh the negative effects on the accuracy due to autocorrelation between observations at points near by.

Two-stage sampling makes it possible to concentrate the sampling effort on a limited number of delineations. This reduces the time needed for travelling between transects, and it opens the possibility to produce results on individual delineations with sufficient accuracy to be useful. We expect that two mutually perpendicular transects will make this accuracy fairly independent from the direction of a possible unknown trend.
Using map delineations as sampling units in a two-stage design seems natural; see e.g. Ragg & Henderson (1980) and Wang (1982) for earlier applications.

Selection with probabilities proportional to size, as applied in both stages, has two advantages. Firstly, the results will usually be more accurate than by selection with equal probabilities (Cochran, 1977). Secondly, the resulting samples are 'self-weighting' in the sense that no weight-factors are needed in the statistical formulas. As shown in the next paragraph, the main calculations are therefore simple.

Statistical analysis of sample data

The statistical analysis of data obtained via the present sampling method is in essence given by Cochran (1977; par. 11.9). We use another notation to simplify the typography.

The formulas below are applicable to quantitative as well as qualitative soil properties. The latter have to be represented by 'indicator' variables with 0 and 1 indicating respectively absence and presence of an attribute. For instance, if purity is to be estimated, 1 is recorded if the observed profile belongs to the taxon predicted by the map and 0 otherwise. The fraction of profiles in the right classes is then obtained by calculating the mean of the corresponding indicator variable.

First stage without stratification

If the set of delineations has not been stratified, an unbiased estimator of the population mean $M(y)$ of a property y is simply:

$$m(y) = \sum \bar{y}_i / n, \qquad (1)$$

where n denotes the number of pairs of transects in the sample and $\bar{y}_i$ denotes the mean of the $i$the pair, calculated as the arithmatic mean of the two transect means. (Note that the cells of each delineation were divided into two strata of practically equal size and that the transect means are unbiased estimators of the respective stratum means.)

An unbiased estimator of the variance of $m(y)$ is:

$$v(m(y)) = \sum\{\bar{y}_i - m(y)\}^2/\{n(n-1)\} \quad (2)$$

Its estimated standard error, $s(m(y))$, is calculated as the square root of $v(m(y))$. Confidence limits for the population mean can be calculated according to:

$$m(y) \pm ts(m(y)), \quad (3)$$

with t read from a table of Student's distribution with n-1 degrees of freedom. If the distribution of the $\bar{y}_i$ is not too skewed, and n is not too small, this will give a reasonable approximation.

An unbiased estimator of the variance among cells is (cf. Cochran, 1977, par. 5A.11):

$$v(y) = m(y^2) - m(y)^2 + v(m(y)), \quad (4)$$

where $m(y^2)$ is an unbiased estimator of the population mean of the squared values per cell. This term is to be calculated in the same way as $m(y)$, after squaring the values of the individual observation points.

It is indicated above how statistical estimates can be made about soil conditions in a mapping unit as a whole. However, estimates are often also required for some part of the population, referred to as domain in the following. For instance, separate estimates of properties may be required for different taxa within a mapping unit. Other domains just arise from the fact that some properties are only defined for a subset of the population. For instance, a property like 'humus content of the B horizon' is only defined for profiles having a B. Such properties have to be considered within their domains of definition.

As a preliminary to estimation within domains, an auxiliary variable x is introduced for each domain, with value 1 if the profile belongs to the domain, and 0 otherwise. Furthermore, a variable y' is introduced with the same values as y within the domain, but 0 elsewhere. The mean of y within the domain, $M(y|x)$, is then estimated by the ratio (Cochran, 1977, par. 11.12):

$$m(y|x) = m(y')/m(x), \quad (5)$$

where $m(y')$ and $m(x)$ denote the estimated means of y' and x respectively, calculated in the same way as $m(y)$ in Eq. 1.

The variance of $m(y|x)$ is estimated by:

$$v(m(y|x)) = \sum\{\bar{y}'_i - m(y|x)\bar{x}_i\}^2/\{m(x)^2 n(n-1)\} \quad (6)$$

An estimate of the standard error of $m(y|x)$ can be calculated by taking the square root of this variance.

The variance among cells within the domain can be estimated with the same procedure as in Eq. 4:

$$v(y|x) = m(y^2|x) - m(y|x)^2 + v(m(y|x)), \quad (7)$$

where

$$m(y^2|x) = m(y'^2)/m(x), \quad (8)$$

and $m(y'^2)$ denotes the estimated mean of the squared values of y' per cell, calculated in the same way as $m(y)$ in Eq. 1.

First stage with stratification

In this case an unbiased estimate of the mean is obtained as a weighted average of estimated stratum means:

$$m(y) = \sum W_h m(y_h), \quad (9)$$

where W_h denotes the weight of the hth stratum, i.e. the area of the delineations in this stratum divided by the total area of the mapping unit. The mean of the hth stratum is estimated as before:

$$m(y_h) = \sum \bar{y}_{hi}/n_h, \quad (10)$$

where $\bar{y}_{hi}$ denotes the mean of the ith pair of transects in the hth stratum (again calculated as the arithmatic mean of both transect means), and n_h denotes the number of pairs in this stratum.

The variance of $m(y)$ is estimated similarly:

$$v(m(y)) = \sum W_h^2 v(m(y_h)), \quad (11)$$

where the $v(m(y_h))$ are calculated according to:

$$v(m(y_h)) = \sum\{\bar{y}_{hi} - m(y_h)\}^2/\{n_h(n_h-1)\} \quad (12)$$

An unbiased estimate of the variance among cells can be calculated with Eq. 4. (The estimator $m(y^2)$ in this equation should of course be calculated in the same way as $m(y)$ in Eq. 9).

For estimations within domains the same auxiliary variable x and transformation of y to y' is used as before. First the means of x and y' are estimated, in the same way as $m(y)$ in Eq. 9. Then the mean within the domain is estimated with Eq. 5.

The variance of $m(y|x)$ can be estimated by:

$$v(m(y|x)) = \sum W_h^2 v(m(y_h|x)), \quad (13)$$

where

$$v(m(y_h|x)) =$$

$$[\sum\{\bar{y}_{hi} - m(y|x)\bar{x}_{hi}\}^2 - n_h\{m(y_h) - m(y|x)m(x_h)\}^2] \;/\; \{m(x)^2 n_h(n_h-1)\} \quad (14)$$

The variance among cells within the domain can be estimated with Eq. 7 en 8. (The estimator $m(y^2)$ in Eq. 8 should of course be calculated in the same way as $m(y)$ in Eq. 9).

Applications

Since we developed the sampling method it has been applied in four projects of the Netherlands Soil Survey Institute, to investigate:
- soil conditions in two related mapping units on the 1 : 10 000 soil map of Lievelde (province of Gelderland);
- soil conditions in three single mapping units and two groups of physically similar mapping units on the 1 : 10 000 soil map of Sleen (province of Drenthe);
- soil conditions in a parcel, to assist one of our experts in quantifying possible damage from reclamation activities;
- purity of two soil maps of the Hupselse Beek area (Gelderland), produced by different methods (Bregt & Bouma, in prep.).

An application to major mapping units on a sheet of the 1 : 50 000 soil map of The Netherlands is being planned.

In the following some results from the Lievelde project are presented as an example. One of the mapping units studied in this project is defined as 'veldpodzolen' (De Bakker & Schelling, 1966) or Typic Haplaquods in slightly loamy medium fine sand, with a mean highest water-table between 40 and 80 cm, and a mean lowest level deeper than 120 cm (Van der Sluijs & De Gruijter, in press). Seven delineations of this unit were selected: three from a stratum with small delineations (i.e. $< \pm 2$ ha), and four from the larger ones. From each of these delineations a pair of transects was selected, with one transect in E-W direction and the other N-S. Using a lag of 25 m this resulted in 61 observation points. By previous division of the largest delineations, the transects were delimited to a maximum length of 300 m. See Figure 1 and 2 for the location of transects and observation points.
At each observation point a profile description was made by augering to 2 m depth. Field estimates of humus content, loam content and median of sand were calibrated against laboratory analyses to remove systematic errors.

Results for some important properties are presented in Table 1, with the estimate of the mean, its 90% confidence limits, and the standard deviation of the properties (i.e. square root of the estimated variance among cells).

Only a part of the observed profiles had a B horizon or a heavy subsoil as defined in this project. So the properties of these layers had to be estimated within their domains, i.e. the part of the mapping unit where they occur. It appears from Table 1 that the areas of these domains were estimated only approximately, but this is a usual feature of samples of this magnitude.
The estimates of the means of the properties themselves are sufficiently accurate. It is clear from the standard deviation of available moisture in the root zone, depth to heavy subsoil, and mean highest and lowest water-table, that there is rather a large variation of soil physical conditions in this mapping unit. In view of this internal variation it is questionable whether units like this should be treated as distinct entities in soil physical simulation studies. These and other results are discussed in detail by Marsman & De Gruijter (in prep.).

Sampling efficiency

With an average of 0.10 day per point, 6.2 days of fieldwork were needed for the 61 points of the Lievelde project. This includes travelling, finding the location of observation points, augering

Table 1. Mean with approximate 90% confidence limits and standard deviation of some soil properties in a mapping unit of Typic Haplaquods (Hn53-VI) on the 1 : 10 000 soil map of Lievelde.

Soil property	Mean and 90% conf. limits	Standard deviation
A1 horizon:		
thickness (cm	26.2 ± 3.4	7.2
humus content (%)	5.7 ± 0.9	1.4
loam content (%)	14.5 ± 2.3	3.9
median of sand (um)	157.5 ± 2.2	6.5
B horizon:		
domain area (%)	65 ± 19	
thickness (cm)	22.3 ± 1.9	11.8
humus content (%)	2.3 ± 0.8	1.3
loam content (%)	10.0 ± 1.5	3.4
median of sand (um)	156.0 ± 2.9	7.5
Heavy subsoil:		
domain area (%)	56 ± 35	
upper boundary (cm)	160.6 ± 13.8	25.5
loam content	20.3 ± 0.4	1.4
median of sand (um)	122.4 ± 0.9	2.0
Root zone:		
thickness (cm)	41.0 ± 3.7	7.4
available moisture (mm)	75.1 ± 13.4	23.4
Mean highest water-table (cm)	50.8 ± 5.6	14.2
Mean lowest water-table (cm)	134.7 ± 15.5	24.4

and description. We regard both this and the accuracy of the results as acceptable.
This does not mean that our sampling design is more efficient than any other. It probably isn't. Neither the number of small and large delineations, nor the distance between points have been optimized.

Comparing different sampling designs in terms of efficiency is incomplete as long as only their relative precisions are considered. Also costs have to be taken into account when making a rational choice of a sampling design. Average costs per observation point may vary with a factor two or more between methods, due to differences in time needed for travelling and finding the observation points. An early example of balancing both factors is given by Wilding et al. (1965). De Gruijter & Marsman (in prep.) hope to give a quantitative analysis of the efficiency of the present and related sampling designs.

References

Bakker, H. de & J. Schelling, 1966. Systeem van bodemclassificatie voor Nederland. De hogere niveaus. Pudoc, Wageningen.

Bigler, R.J. & K.J. Liudahl, 1984. Estimating map unit composition. Soil Survey Horizons 25: 21-25.

Bouma, J. (this issue). Soil variability and soil survey. Proceedings SSSA-ISSS Workshop on Spatial Variability, Las Vegas, 1984.

Bregt, A.K. & J. Bouma (in prep.). Spatial patterns of soil physical characteristics as derived from soil survey and from kriging of point data.

Cochran, W.G., 1977. Sampling techniques. Wiley, New York.

Gruijter, J.J. de & B.A. Marsman (in prep.). Two-stage designs for spatial point sampling: efficiency and optimum allocation.

Marsman, B.A. & J.J. de Gruijter (in prep.). Soil conditions in two related mapping units of Typic Haplaquods as studied with random transects.

Powell, J.C. & M.E. Springer, 1965. Composition and precision of classification of several mapping units of the Appling, Cecil, and Lloyd series in Walton County, Georgia. Soil Sci. Soc. Am. J. 29: 454-458.

Ragg, J.M. & R. Henderson, 1980. A reappraisal of soil mapping in an area of southern Scotland: I. The reliability of four soil mapping units and the morphological variability of their taxa. J. Soil Sci. 31: 559-572.

Sluijs, P. van der & J.J. de Gruijter (in press). Water-table classes: a method used to indicate seasonal fluctuation and duration of water-tables on dutch soil maps. Agr. Water Management.

Steers, C.A. & B.F. Hajek, 1979. Determination of map unit composition by a random selection of transects. Soil Sci. Soc. Am. J. 43: 156-160.

Wang, C., 1982. Variability of soil properties in relation to size of map unit delineation. Can. J. Soil Sci. 62: 657-662.

Wilding, L.P. (this issue). Spatial variability: its documentation, accomodation and implication to soil surveys. Proceedings SSSA-ISSS Workshop on Spatial Variability, Las Vegas, 1984.

Wilding, L.P. & L.R. Drees, 1983. Spatial variability and pedology. In: Pedogenesis and Soil Taxonomy. L.P. Wilding, N.E. Schmeck and G.F. Hall (eds.): Developments in Soil Science IIA. Elsevier, Amsterdam: 83-113.

Wilding, L.P., R.B. Jones & G.M. Schafer, 1965. Variation of soil morphological properties within Miami, Celina and Crosby mapping units in West-Central Ohio. Soil Sci. Soc. Am. J. 29: 711-717.

Discussion

J. Culling: My concern relates to the problem of soil characteristics that are highly variable in both space and time. Your water-table data are an excellent example. It seems to me that an adequate analysis requires both spatial and temporal considerations. In this context, climatic factors need to be involved.

J. de Gruijter: Your point is very well taken. We should be careful to not overemphasize spatial variability, because temporal variability is often more pronounced. We find that simulation models are particularly useful to express temporal varability, because data gathering at one site during a period of several years is often prohibitive in terms of cost.

E. Runge: We are in an interdisciplinary phase of soil science. You don't primarily use your own discipline but communicate it to others. We must recognize the inherent conflict of "knowledge base" versus "problem base" research. We must be tolerant of the difference and recognize each others strengths, so as to maximize knowledge being generated. We want to get a year smarter as well as a year older. We must continue this good start of a dialogue between Commissions S1 and S5.

P. Greminger: Do you think there is a possibility to build a transfer function to get an idea about the impact of acid rainfall on the pH of the soil profile as a function of time as well as of space?

J. de Gruijter: The impact of acid rainfall on soil pH as a function of time is governed by many interrelated factors, such as type of vegetation, parent material, climatic conditions and soil profile development. Attempts are being made at several research institutes to develop complex deterministic simulation models to describe and possibly predict the rate of acidification processes. The term transfer function would, in my opinion, not really apply to such models. I would suggest to use there term only to relate relatively simple land characteristics (e.g., texture, etc.) to more complex land characteristics (e.g., CEC, P-adsorption) as mentioned in my paper. Variability in space adds yet another aspect to

the above. However, well defined transfer functions allow a better evaluation of the spatial (and temporal) variability aspect because of the larger amount of available data when using relatively simple land characteristics.

Spatial variability: its documentation, accommodation and implication to soil surveys

L. P. Wilding, Soil & Crop Sciences Department, Texas A&M University, College Station, TX 77843

Introduction

Spatial variability of soil is not an academic question. It is a real landscape attribute; our unwillingness or inability to identify it in no way decreases its magnitude or existence. It is no stranger to the pedologists for it is the very essence of their profession. As scientists we must document the magnitude and form of soil variability; accommodate its existence in models of soils; and transmit accurately the expected pattern and implication of spatial changes to users of soil resources. Soils are not material specific; many soil properties are not single valued, many are transient, and many are not randomly distributed but rather systematically time and spacially dependent. The dilemma is that soils are not isotropic media but rather they are strongly anisotropic laterally and vertically (Wilding and Drees, 1983).

Soil Surveys - Science with a Little Art

Spatial variability within landscape bodies is a continuum. The purpose of soil surveys is to partition this continuum into natural or artificial classes that have greater homogeneity for selected soil properties than the continuum as a whole. Soils with similar properties and environments are expected to behave similarily. Limiting the range of soil variables permits more accurate predictions of expected responses to alternative soil management imputs and land use. By this means, soil surveys represent a powerful vehicle for technology transfer.

The methodology in conducting soil surveys, the accuracy of the soil map, and the specificity of its application are often misunderstood. Hence, the following discussion outlines the scientific basis, data aquisition, and kinds of soil variability observed in making soil surveys in the U. S. (West et al., 1984). The scientific basis of a soil survey is that soils and their location on the landscape are predictable to an experienced soil scientist who has a knowledge of the geology, vegetation, climate, and landform patterns of the area. The soil scientist is not able to observe or sample the soil at every point on the landscape. Only enough observations are made during mapping to

determine soil/landscape relationships and to confirm predictions of soil models established from these relationships. Thus, map unit delineations are derived, to a great extent, from inferences gained from a small sampling of the landscape. This is possible because soil properties change systematically with landscape position. Visible changes in slope, vegetation, surface color, and drainage pattern enable a soil scientist to locally extrapolate soil/landscape relationships previously established. Thus, a soil map unit is a landscape (cartographic) unit that reflects the dominant soil conditions of a landscape element or segment (Fig. 1).

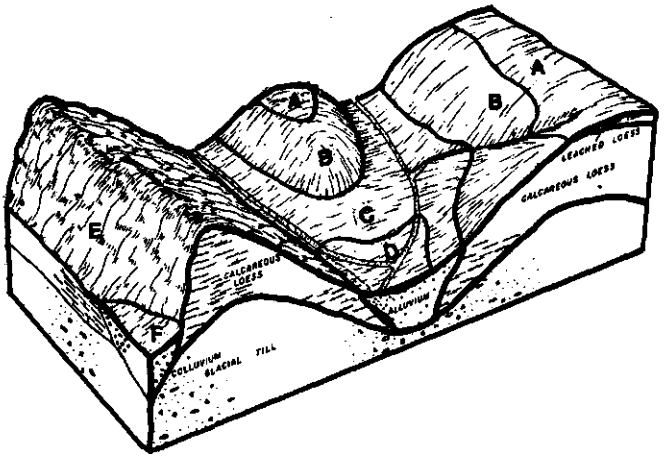


Figure 1. Schematic illustration of soils (A through F) occurring on different landscape segments (stability) and derived from different parent materials (Revised from Fig. 2.5 Grossman, 1983).

Most map units are named for the dominant soil series occurring in the delineated landscape body. Unfortunately, one of the common properties of soils as natural landscape bodies is variability. In addition to the systematic variation observed above, variability also occurs in an undefined (random) manner. Soil scientists are aware of this random variability and make every attempt to design map units and to map soils that restrict property variability to limits that permit meaningful interpretations of soil use, management, and behavior. However, soil properties are not homogenous within mapped areas. They often have ranges that exceed the limits imposed by the definition of the soil series naming the map unit. Thus, map units may consist of a number of soils either similar or dissimilar to the named series.

The goal of the soil scientist is to design the map units and delineate them in a manner that limits the number of inclusions. However, inclusions are almost always present both as similar

soils with properties just outside the series range and as areas of dissimilar soils too small to delineate at the scale of mapping. Serious errors, from a user's stand point, may occur when a soil map is used to greater detail than the accuracy of mapping warrants. Detailed inspection is needed if minor soil variation is important to the intended use. The map unit descriptions should alert the user to expected variations.

In summary, making a soil survey is a modeling exercise involving both the scientific method and an element of art. A soil model is developed by correlative studies with landform conditions; hypotheses are formulated, tested by ground truthing and verified; the model is revised and new hypotheses are formulated for testing when the model fails. The soil/landscape portrait thus evolved is an artwork of the soil scientist much like a painter who transfers his mental image of a subject to the canvas; the greater the familiarity the painter has with the subject, the more accurately expressed the mental construct will be.

Spatial Variability - A Continuum

Spatial variability in soils is a continuum from megascopic to microscopic levels of resolution. It is dependent on the property of interest, area or volume observed or intergrated, and methods of determination (Fig. 2). An inverse relationship exists between the field of view and level of resolution (Fig. 2). To generalize information gained at high levels of resoltuion will require a telescoping series of observations from visual to subvisual levels. Further, a soil survey of the experimental area with appropriate placement of soils into a classification system, such as Soil Taxonomy, will foster technology transfer and generalization of the research results at all levels of resolution.

Pedologists have long been cognizant of long-range systematic changes, especially those correlative with surficial features, in soils, but only of recent have they focused on the magnitude of short-range spatial changes. Spatial variability in soils rarely increases linearly with distance; more commonly it is an expoential function and strongly dependent on the pedogenesis of soils under given environments (Fig. 3). Spatial variability which generates the greatest problems in soil survey work are those changes in soils that occur subsurfically with no expression at the surface. This may be a consequence of parent material stratification, differential pedogenic processes, biological activity, or temporal changes in soil properties. Generally soil surveys made in areas of of low relief and/or of soils developed from fluviatile or tectonic parent materials yield map delineations of least reliability and accuracy.

Systematic variability occurs in soils at microscopic and submicroscopic levels. Particle orientation, zonation of organic

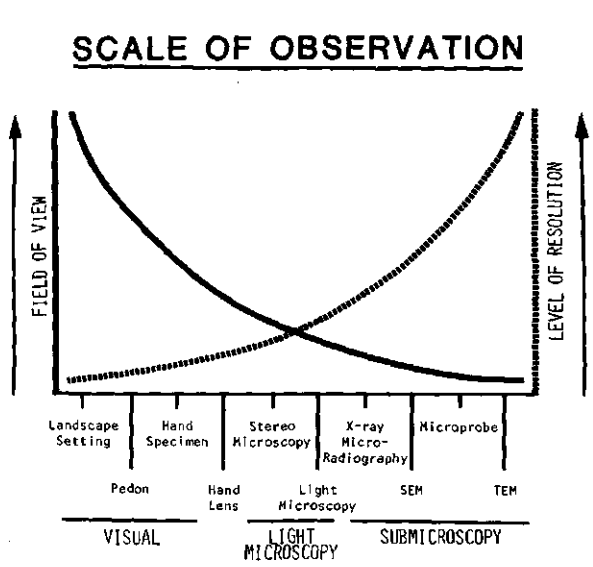


Figure 2. Schematic illustration of the relationship between increasing levels of resolution and the area of the field under view (Fig. 1, Wilding, L. P. and K. W. Flach. 1985).

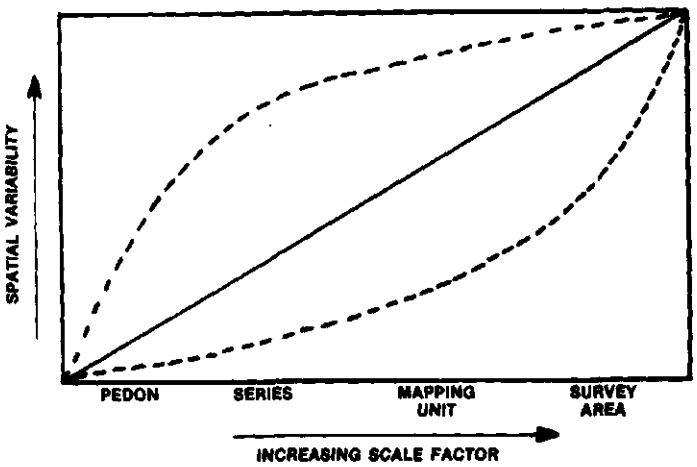


Figure 3. Schematic spatial variability reflecting increasing scale factor (Fig. 4, Wilding and Drees, 1978).

and inorganic components and pore size distribution patterns are frequently not random. Skeletal grains, ped structural surfaces and walls of other voids commonly serve as loci for concentrations of colloids and solutes and/or the reorientation of clay particles. Frequently, the sample volumes analyzed are sufficiently large to integrate the systematic variability resulting from loci specific reactions and fluxes. Soil surveys should aid the scientist in determining those soil properties which have systematic organization at given levels of resolution versus those soil conditions that are more likely random. They should also serve to indicate the scale at which soil volumes are likely to integrate both systematic and random variability.

Statistical Considerations and Sampling Schemes

A pedologist should have a much better concept of soil property relationships, covariance and soil distribution patterns than a statistican without soils experience. With such prior knowledge care must be exercised not to knowingly confound systematic and random variability in a sampling scheme. For example, if soils systematically change as a function of relief, then the most efficient, non-biased sampling scheme is one that tranverses normal to drainage systems. Likewise, sampling a soil, either laterally or vertically, should not be at random. No useful purpose is served in compositing soil materials from two separate populations of known morphological, physical, chemical and biological difference. Random sampling is suitable only when soil differences are not evident.

Classical statistics has not advanced our knowledge of the causal factors responsible for soil property variation within a landscape (Wilding and Drees, 1983). Often the soil factor is masked by removing it as blocking or replication error. Soils often exhibit non-orthogonal variability spatially--variance is not isotropic laterally or vertically.

Soil observations are not necessarily spatially independent and frequency functions are usually not normal but skewed log normal or gamma distributions. How do we overcome all of these constraints that are assumptions basic to most classical statistics?

The question of sampling scheme, statistical analyses to be employed and observational interval continue to plague most pedologists. These answers really depend on objectives of work, nature of classes being sampled and precision of the results expected at a given confidence level.

Sampling Schemes

Sampling schemes briefly considered include: random, transect, grid, and systematic radial transects or some combination of these. On a given delineation, the first three schemes are

illustrated in Fig. 4. and the latter in Fig. 5

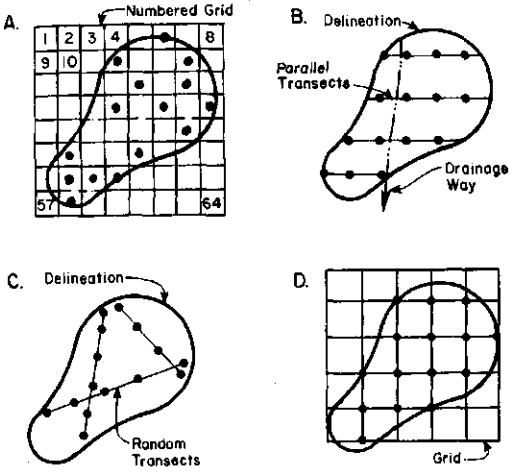


Figure 4. Sampling schemes to determine composition of mapping unit delineations. Observations located at random (A), along parallel point transects (B), along random point transects (C), and at intersection of grid lines (D), Same number of observations in each scheme (Fig. 4.6, Wilding and Drees, 1983).

The **random sampling scheme** is unbiased and statistically sound. It is commonly preferred among statisticians but may cluster data spatially unless a large number of observations are taken. It may also confound systematic and random error. The **point-transect sampling scheme** always involves the question of random vs. systematic orientation. Stratified random point transects may be aligned normal to an anticipated soil gradient (Wang, 1982). Point transects may take advantage of trenches, powerlines, highways, etc., that may not be aligned with a soil bias. They are useful in remote areas where vegetation or topography obscure cultural features. They may be oriented parallel to each other to form a grid if desired. The **grid sampling scheme** provides equally spaced observations. It is better adapted to geomorphic-pedogenic studies. With elevation control it can be used to simulate 3-D computer-generated surface nets of limiting or diagnostic soil properties. It is the best design for geostatistics and semi-variance analyses. For the **systematic radial transect scheme**, a series of radial transects at 60^{o} of each other are laid out from a central point with multiplicity of observations increasing with distance from the center (Fig. 5). This results in telescoping series of observations with increasing area that may be coupled with the grid scheme to yield short-range to long-range variability across a large sampling area.

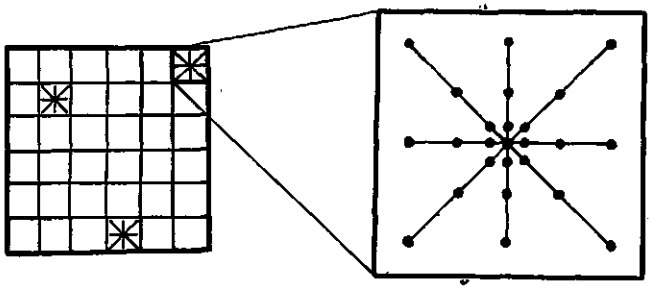


Figure 5. Grid and systematic radial sampling scheme used to sample undisturbed and disturbed soils (Bearden, 1984).

Number of Observations or Transects

The number of observations or transects needed to characterize soil properties in a certain sampling unit is determined by the population variance of the property, the confidence level chosen, and the probable error tolerance about the mean that is acceptable (Fig. 6). While the variability of a property in a landscape may be fixed, one way to decrease the population variance is to stratify the sampling units so that the total variation can be partioned as much as possible to differences among strata (Steel and Torrie, 1960). Wang (1982) discuss means to accomplish this with sampling of natural soil units.

The confidence level is an expression of probability that a statement is correct and is reflected by the "t" value selected. The higher the t-value the more samples or transects will be needed. While a confidence level of 99% or 95% is common in many fields, in soils a confidence level of 70 to 80% is probably more realistic in terms of time and money imputs that are practical to a sampling scheme.

The confidence interval or deviation allowed from the mean (probable mean error) is an estimated range of the population mean that is likely to occur within a given soil unit sampled. If the variance of the property observed is high, (CV's above 25 to 35%) our precision of estimate of the true mean at a given confidence level may need to be decreased to accommodate a practical number of samples (Fig. 6).

Frequently we must accept either a lower confidence level (probability) or higher confidence interval (probable mean error) to maintain a sampling scheme within reason. Under some circumstances, perhaps many, limits within 20% or even 50% of the mean with a probability level of 80% or lower may permit sufficiently accurate mean estimates. This would depend on the property in question, the magnitude of the mean, the critical confidence limits for interpretations, and the risk one is willing to take in making an error in judgement.

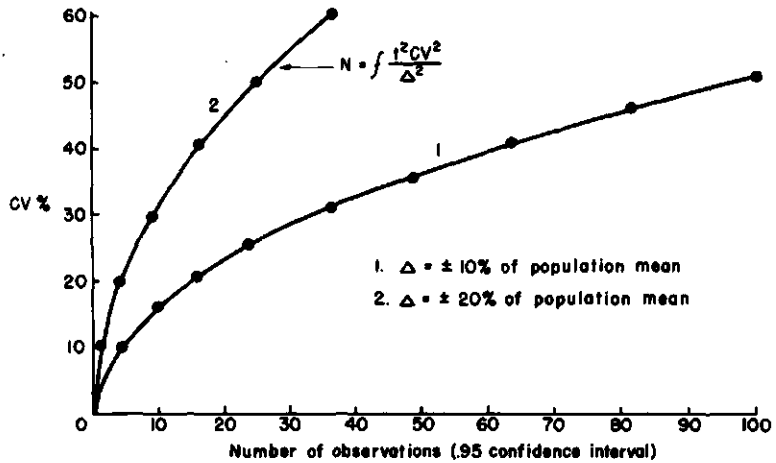


Figure 6. Coefficient of variability (CV) versus number of observations necessary to estimate the population mean within specified limits (Fig. 4.4, Wilding and Drees, 1983).

Table 1 gives relative ranking of soil property variability in a landscape. It is based on the work of this author and extensive compilation of data in the literature. This table is intended to be utilized as a guideline when on-site data are not available; it is not to be interpreted rigidly. Quite often of greater importance than the absolute CV's or probable mean errors of a given property is the class range of a property or set of properties which can accommodate similar behavior or response interpretations. Under such circumstances the CV's might be quite large but because most of the values fall within an acceptable class range, the spatially property variability would not adversely affect interpretations. This is the important consideration.

Observational Interval

No definitive statement can be made about the observational interval because it is a function of soil conditions at a given site and scale of mapping. It can be determined arbitrarily by a predetermined ground distance on a photobase map or statistically by semi-variance or autocorrelation methods. The latter have been used relatively little by soil surveyors (see later discussion). More commonly the intervals are based on knowledge of soil complexity gathered during the mapping phase, on the character of the parent material variability, and on a realistic number of samples that can be taken within an available

Table 1. Relative ranking of variability of soil properties that occur in landscape units of a few hectares or less in size. These values represent lateral variability for equivalent horizons or depths as appropriate for given properties (Modified from Table 4.2, Wilding and Drees, 1983).

Variability of Property	Number of profiles needed*	Property
Least (CV's <15%)	< 10	Soil color (hue and value) Soil pH Thickness of A-horizon Total silt content Plasticity limit
Moderate (CV's 15 to 35%)	>10 to 25	Total sand content Total clay content Cation exchange capacity Base saturation Soil structure (grade and class) Liquid limit Depth in minimum pH Calcium carbonate equivalent
Most (CV's >35%)	< 25	B2 horizon and solum thickness Soil color (chroma) Depth to mottling Depth of leaching (carbonates) Exchangeable hydrogen, calcium, magnesium and potassium Fine clay content Organic matter content Plasticity index Soluble salt content Hydraulic conductivity Water content

* Employing 95% confidence interval and limit of accuracy 10% of mean.

time frame. Wang (1982) suggests observational intervals be: (1) either 1/10 or 1/20 length of transects with no fewer than 10 to 20 observations per transect; or (ii) 1/2 length of shortest transect. He prefers (i) because it gives every transect the same number of degrees of freedom and makes statistical computation easier. If one is attempting to quantify the soil composition in a landscape unit, ground distances between observations of 20 to 200 m are common, while for pedon considerations, distances of 0.25 to 0.5 m are in order (Wilding and Drees, 1983).

Generalized Aspects of Spatial Variability

Following are a few generalities that should be kept in mind in considering spatial variability in soil surveys.

- Reliability in accurately predicting many soil properties decreases with depth. Fewer observations are made at greater depths (ie >2m) in the soil than near the soil surface; thus there is less ground truth control at depth.
- Spatial variability in soils is closely allied with the nature of the parent material from which soils are formed. Parent materials from least to most variable include: loess, till, fluviatile deposits, tectonic rocks, and drastically disturbed soil materials.
- Static soil properties are less variable than dynamic ones (i.e. OM, texture, mineralogy, solum depth, soil color vs hydraulic conductivity, soil moisture content, salt content, microorganisms, exchangeable cations and redox conditions. Soil Taxonomy is strongly biased towards static properties in subsoil horizons.
- Properties which can be closely calibrated to a standard (or quantified in the field) are less variable than those which are qualitative (i.e. texture, color, pH vs structure, consistence, porosity, root abundance, etc.).
- Properties which are differentiating or accessory to differentiating properties used in mapping and classifying soils will be less variable than those of an accidental nature. Not all properties are considered in making a soil surveys; only those that are considered most important to the objectives of the survey or classification scheme.

Geostatistics and Soil Survey

The use of geostatistics in establishing map unit concepts and in the mapping process is yet mostly untested (Wilding and Drees, 1983). Few pedologists are familiar with this statistical approach. Frequently the mathematical and statistical foundation is beyond their comprehension -- they are afraid to use a tool they do not understand either in terms of principles or limitations. One of the likely applications of geostatistics to soil survey is to determine the sampling interval that assures spatial independence, but the question remains how to generalize

this information. It may also play an important role in developing soil property/landform correlations and to aid in designing map unit concepts. For high intensity soil surveys (i.e. experimental research plots) it should be a valuable tool to kring the area so mean and variance estimates could be predicted for soil properties at all plot loci.

Questions being raised about this tool for soil survey application are as follows:

- How can geostatistics be used without collecting an inordinate number of samples? Many observations are needed to determine how few could have been collected to obtain the same information base.
- How can geostatistics be used as an extrapolative tool? Soil surveys are mostly an extrapolation of knowledge gained from one landform and extended to the next.
- How might it be used for pedon and microscale variability?
- If one finds that the observations are spatially dependent over the area investigated, how can classical statistics be applied to the data set?

Verification and Accommodation of Spatial Variability

Because of the limited number of observations collected in the conduct of soil surveys, there is a required verification phase of the mapping units and the soil characteristics represented by these delineations. These two phases will be discussed under the topics mapping verification and pedon (sampling unit) verification.

Mapping Verification

Transects - Perhaps the most common means to verify mapping unit composition (spatial variability of soils within mapping units) is to employ stratified random point transects (Steers and Hajek, 1979). This procedure has been employed as a means to sharpen the correlation process at the end of the survey. It has also been used as a routine mapping tool during the survey in areas of low relief, high vegetative density or where there is little surficial predictive capability of the landscape (Wang, 1982). Results from this work have been incorporated into soil surveys as probability statements of the map accuracy in terms of given soil conditions and/or interpretations inferred from observed soil properties. Arnold's (1979) graphical bionomial confidence limit method has been used for this purpose (Fig. 7). Results with elevation control have also been plotted as two-dimensional crossections illustrating distributions of soils as a function of topography, parent materials, limiting horizons, etc. (Fig. 8 and 9).

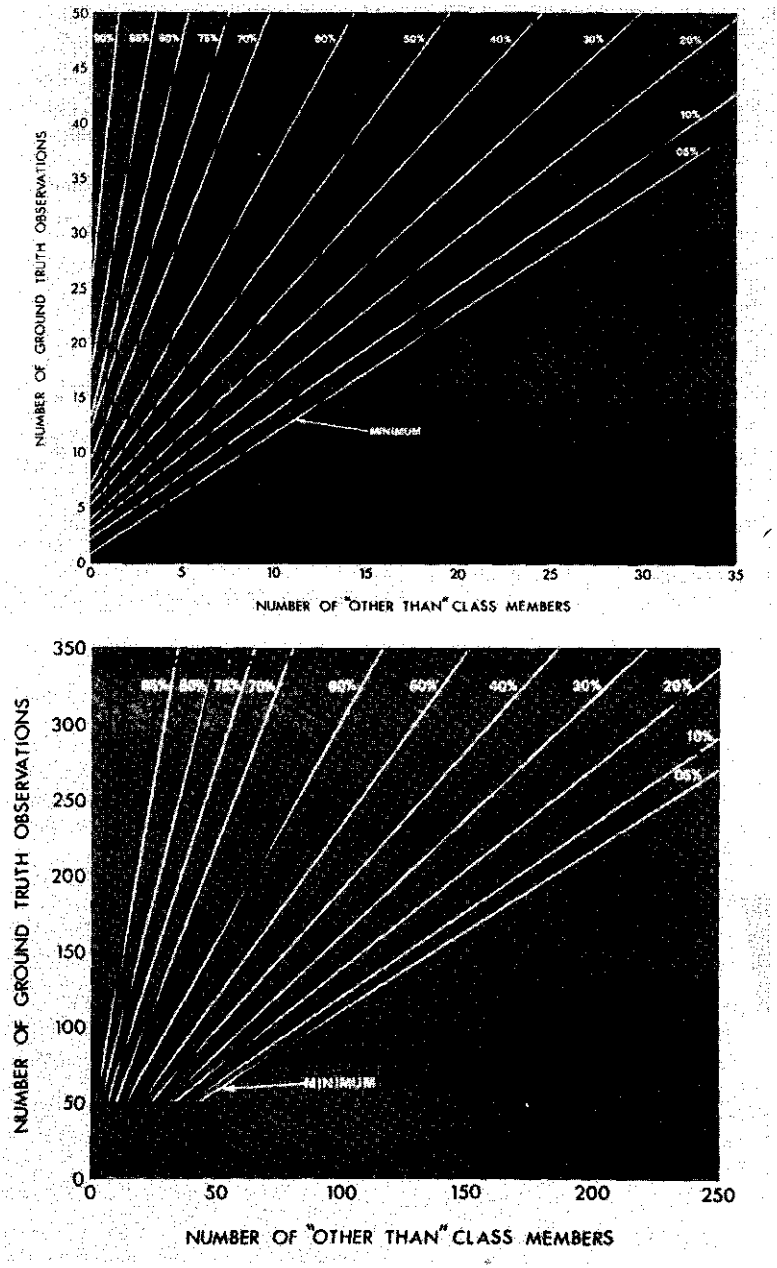


Figure 7. Arnold's (1979) Binomial Confidence Limit Graphs: (A) 0 to 50 and (B) 50 to 350 samples (Taken from Wang, 1982).

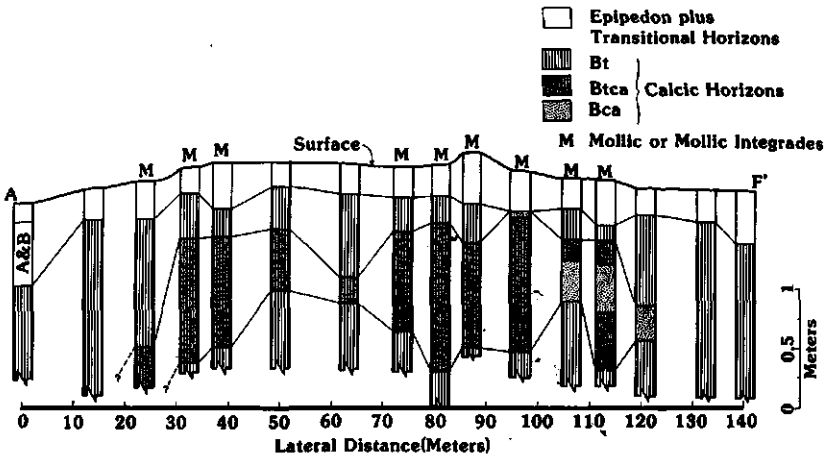


Figure 8. Cross section along a topographic microhigh of the Addicks soil in Harris County, Texas (Fig. 5, Sobecki and Wilding, 1982).

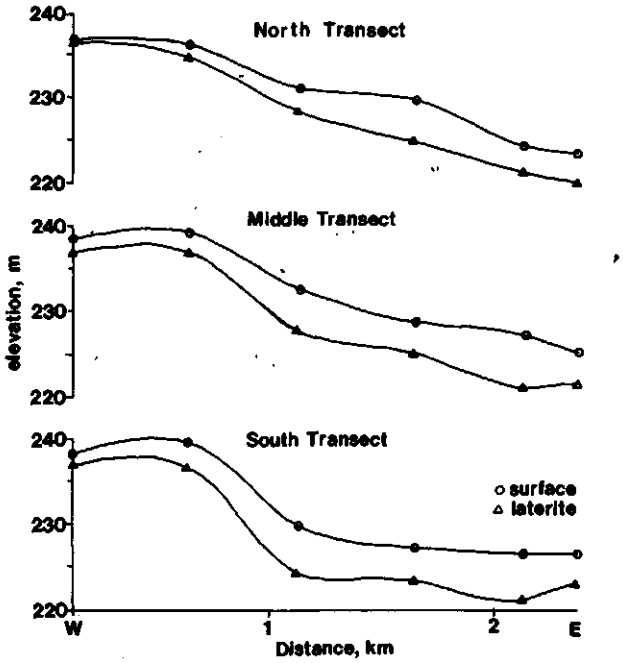


Figure 9. Elevation of soil surface and underlying laterite along three transects in a soil survey area of Niger, Africa (Taken from Fig. 2, West, et al., 1984).

Grids - The grid sampling scheme has been utilized with elevation control to simulate 3-D surface nets of surface relief, A/Bt contact, and differential thickness of parent materials (Fig. 10) in the "Sand Sheet" land resource region of southern Texas and the distribution of soils on a landscape in the Coast Prairie land resource area of Texas. (Fig. 11).

Such a display permits one to readily visualize the configuration of horizons governing solute and water movement over short distances and, thus, probable genesis of such soils. It also permits one to more accurately place soils on the landform. However the cost and time required for such sampling efforts may limit their application to research emphasis rather than routine mapping methods.

Grid sampling schemes have also been used to accurately determine the composition of soils within previously mapped areas of natural or disturbed soils. This approach has been used in attempts to determine the accuracy of older soil surveys when compared to the same areas which have been remapped under modern soil survey standards (Rehage, et al., 1982). Comparisons were based on interpretations of the grid points for alternative land uses in view of former versus current map unit delineations. If much of the variability is at short-range and many of the inclusions in the mapping unit have similar restrictive interpretations as the dominant soil, remapping of older soil surveys may not greatly improve the mapping and interpretation accuracy even though the size of the delineations may be smaller (Rehage et al., 1982). Similar approaches have been used to compare the mapping variability of surface mine disturbed lands with adjacent undisturbed soils (Bearden , 1984).

Ground Penetrating Radar - Ground-penetrating radar (GPR) systems, specifically designed for soil reconnaissance, can provide continuous profiles charting the depth and extent of diagnostic subsurface horizons. Unlike most other radar systems, GPR transmits repetitive, short duration pulses of electromagnetic energy into the soil, rather than into the air. When these pulses strike an interface separating layers of contrasting electrical properties a portion of the signal is reflected back to the radar. The depth to the interface is determined by the time delay of the echo. By towing the radar across the surface, a continuous profile of subsurface conditions can be developed (Personal Communication, Mr. Jim Doolittle, SCS Soil Scientist, Florida).

Present GPR systems do not work equally well in all soils. The radar signal is dissipated and the probing depth restricted by high contents of moisture, clay, or salts in solution. Though the effectiveness of the GPR is extremely site specific, probing depths of 1 to 2 feet in clays and as great as 40 feet in sands can be expected.

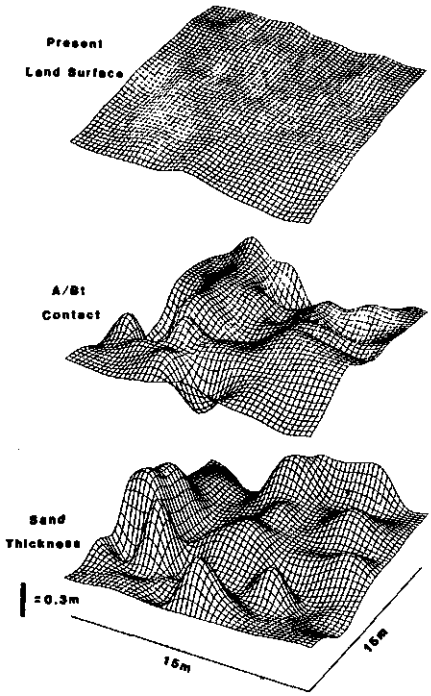


Figure 10. 3-dimensional surface plots of soils in the Sand Sheet region of southern Texas (unpublished data by author).

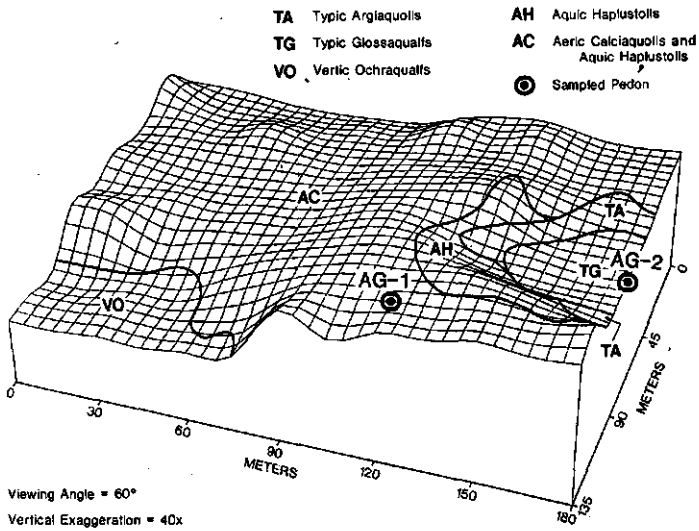


Figure 11. Soil-topography relationships in the Coast Prairie region of Galveston County, Texas (Fig. 6, Sobecki and Wilding,1982).

During the past three years, USDA's Soil Conservation Service has tested GPR on a wide variety of soils in diverse geographic locations, principally to assess the variability within map units delineated by traditional survey procedures. The GPR has also been used successfully to measure subsidence in the Florida Everglades, and sedimentation in lakes; to characterize geomorphic contacts and features; to document the fractures, voids, and solution cavities in rock; to locate buried artifacts, bodies, utility lines, and cables; and to determine the presence and distribution of buried waste materials (Figs. 12 and 13).

Pedon Verification

Except under cyclic conditions a sampling unit for Soil Taxonomy is equivalent to a $1m^2$ area over the depth of the soil. Even under cyclic conditions, such as gilgai in Vertisols, there is good basis for restricting the sampling unit to an area of about $1m^2$ which corresponds to microhigh and microlow topographic relief forms (Wilding, 1982). Within such a sampling volume, soils can express considerable lateral variability especially in more labile properties such as water content, salt content, exchangeable cations, clay content, organic matter, free iron oxides, etc. After one has established the magnitude of lateral variability for given soil properties by horizon, one can use this information to determine the number of samples that need to be collected to estimate the mean of most soil properties of interest within given tolerance limits, at a specified probability level. Figures 14 to 17 give examples of such variability for selected properties within pedons sampled by the authors (Wilding and Drees, 1983).

The problem is that few data of this nature are available because of the time and expense of such sampling. How does one attain horizon pedon variability of this nature for major soils within a survey area? How far can one generalize this information base? With only 3 or 4 horizons sampled for lateral variability within a pedon, how can this information be used to determine significant differences among horizons not sampled laterally in the same pedon especially if different variances exist for different properties at several depths.

Under the above constraints, a composite sampling scheme is proposed for sampling pedons. By employing a sampling scheme where the same subhorizon is multiple-sampled in a lateral vector and then composited to represent a single horizon sample, the mean conditions can be estimated with greater accuracy. Figure 18 illustrates several schemes which have been employed for this purpose. **It is important that only morphologically similar subsamples are composited.** Multiple subsampling four times instead of once will decrease the probable sampling error in half. Using determined subsample error terms for pedon properties, Fig. 19 illustrates absolute class boundary probable errors for several Soil Taxonomy class limits.

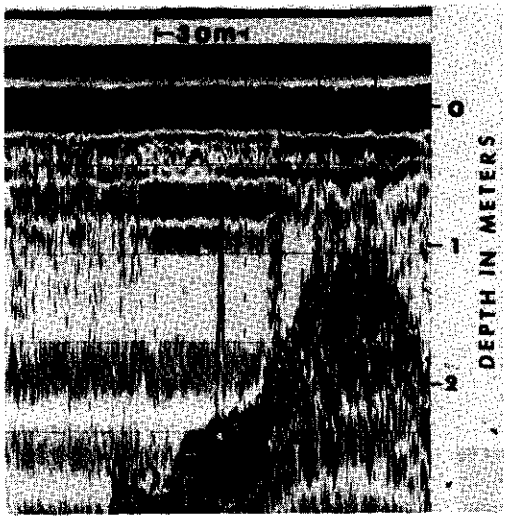


Figure 12. GPR profile of Immokalee fine sand mapping unit (sandy, siliceous, hyperthermic Arenic Haplaquods). Multiple horizons are often clearly expressed on graphic profiles. For example, the GPR profiled a water table, multiple spodic horizons (Bh), and an argillic horizon (Bt) (courtesy of Mr. Jim Doolittle, USDA-SCS Soil Scientist, Florida).

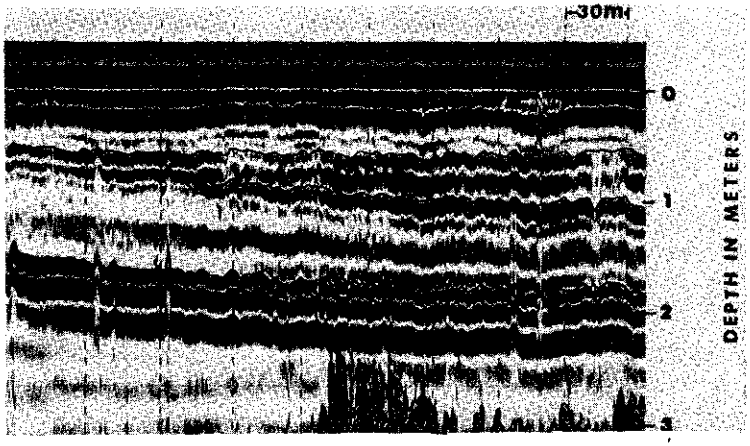


Figure 13. GPR profile of an area of Candler (hyperthermic, uncoated Typic Quartzipssamment) and Apopka (loamy, siliceous, hyperthermic Grossarenic Paleudults) fine sands, 5 to 12% mapping unit complex in Florida. The deep sands of the Candler without an argillic horizon within 2 m is shown on the left with the Apopka soil and its argillic horizon (Bt) on the right (courtesy of Mr. Jim Doolittle, USDA-SCS Soil Scientist, Florida).

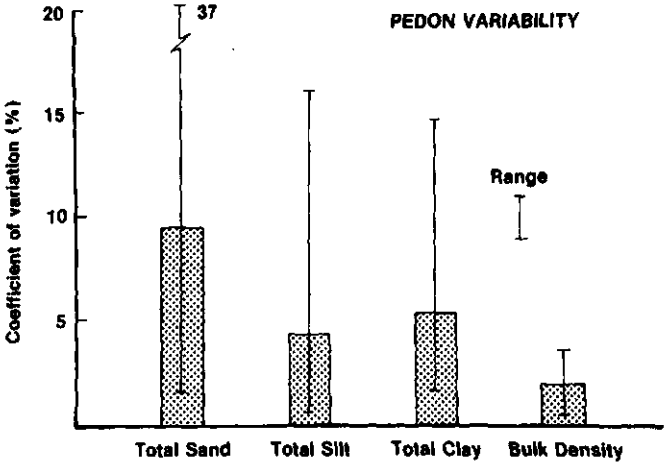


Figure 14. Variability of total sand, silt, and clay contents and bulk density due to combined sampling and analytical errors for five morphologically uniform pedons of about 1 m^2 area (from fig. 5, Smeck and Wilding, 1980).

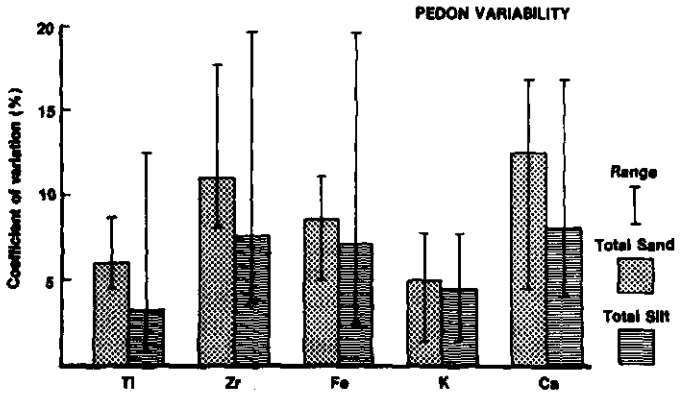


Figure 15. Variability of elemental Ti, Zr, Fe, and K and Ca in sand and silt fractions due to combined sampling and analytical errors for five morphologically uniform pedons of amout 1 m^2 (from fig. 6, Smeck and Wilding, 1980).

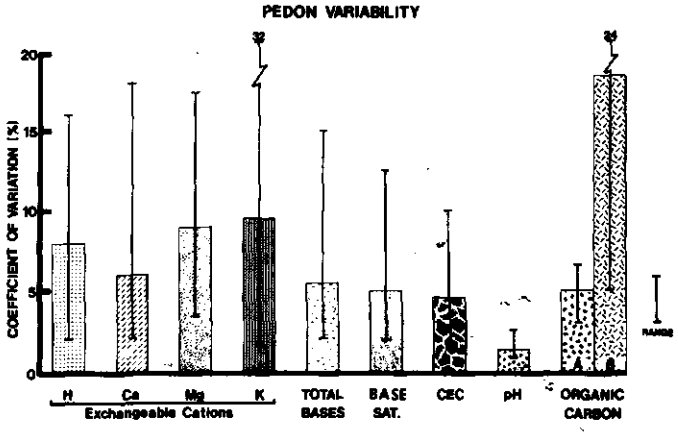


Figure 16. Observed lateral variability among selected chemical properties within horizons of similar morphology for a pedon sampling unit of about 1 m^2 area (reprinted from Fig. 5, Wilding and Drees, 1978).

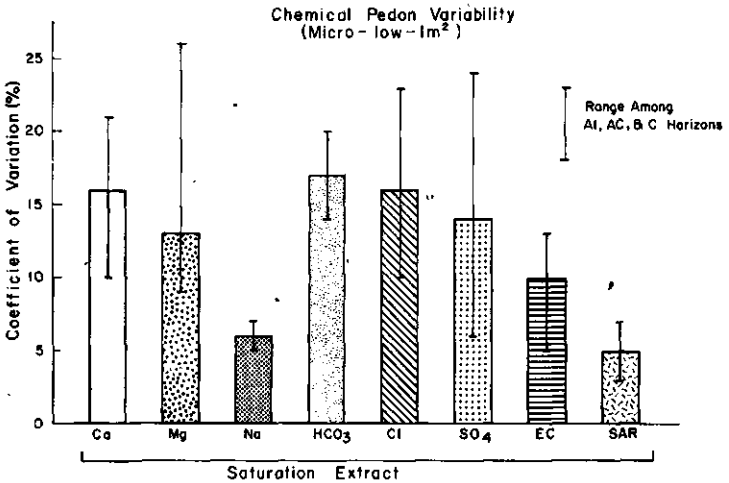


Figure 17. Observed lateral variability of saturation extract parameters within horizons of similar morphology for a pedon sampling unit in Vertisols (about 1 m^2 area of micro-low gilgai element, from fig. 4.14, Wilding and Drees, 1983).

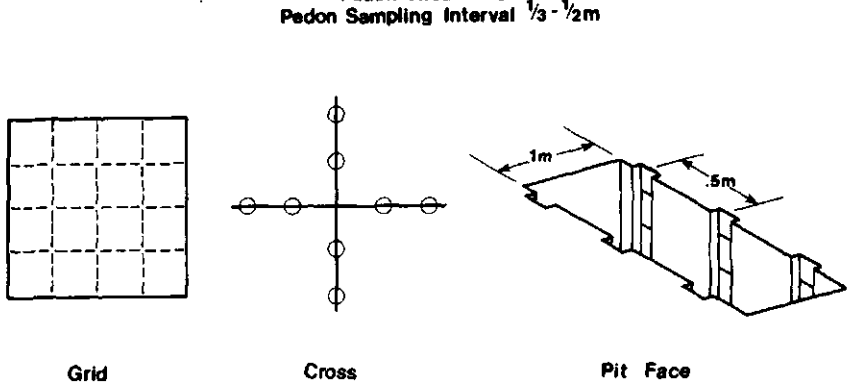


Figure 18. Multiple sub-sampling schemes for pedon sampling units, (from Fig. 6 Wilding and Drees, 1978).

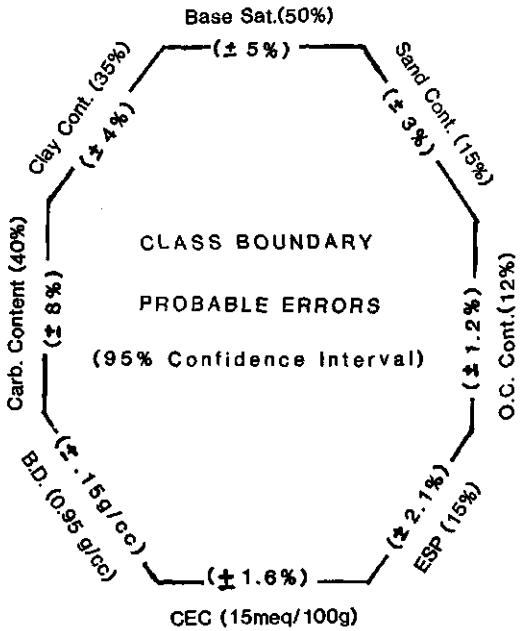


Figure 19. Absolute class boundary errors for several differentiae used at various categorical levels in Soil Taxonomy.

Soil Taxonomy class limits.

For example, the base saturation boundary probable error is 5% at 50% base saturation, sand content is 3% at the 15% limit, ESP is 2.1% at the 15% limit, etc. By using lateral subsamples these values could be halfed or brought very close to the laboratory errors of determination. Application of pedon data to Soil Taxonomy, the Canadian Soil Classification System or any other classification systems with rigid class boundary limits must consider such pedon and laboratory errors in classifying the pedon. Pedons having properties that fall within the boundary error ranges can be placed equally well into either adjoining classes.

Another means of pedon verification is to collect satellite samples about a pedon sampled for complete characterization (Mausbach, et al., 1980 and West et al., 1984). Satellite samples ranging from 2 to 5 in number might include only 2 or 3 horizons of specific interest and only those determinations of properties critical to the classification or interpretation of the soil. This provides a cost and time effective means to evalute the spatial variability of soils considered to be within the limits of a given series.

SUMMARY

Pedologists have long been aware of spatial variability and in the future we can expect to see further advances to unscramble the complexities of soil-landscape relationships. Geomorphology, multivariant analyses, and greater awareness of short-range changes in soils will gain greater emphasis. Specifically, the knowledge of soil variability is necessary to:

- Establish significant vertical and lateral differences among soil properties
- To obtain central tendency and variance statistics
- To design more accurate sampling schemes
- To establish soil property-landform relationships
- To design soil survey mapping unit legends with greater accuracy
- To determine composition of soils representing of mapping units in soil surveys
- To determine validity of class definitions and class differentiae
- To apply pedon data to soil classification systems
- To quantify pedogenic processes
- To verify soil properties and spatial variability for more accurate interpretations of soil behavior and management

A knowledge of spatial variability gained from soil surveys should provide the following kinds of information pertinent to understanding spatial variability in landscape setting:

- Terrain configuration and slope gradient
- Origin and spatial distribution of parent materials
- *Soil drainage characteristics*
- Physical and chemical restrictive layers
- Soil thickness and leaching potential
- Distribution of structural units and cracking patterns
- Small-scale ped and pore-wall physical and chemical zonation of materials
- Sampling volumes necessary to achieve integraded sampling units
- Nature of physical, chemical and biochemical boundary gradients

LITERATURE CITED

Arnold, R. W. 1979. Strategies for field resource inventories. Agronomy Mimeo 70-20. Dept. of Agronomy, Cornell Univ., Ithaca, New York. 14853.

Bearden, E. D. 1984. A comparison of undisturbed and surface mined soils in Freestone County, Texas. M. S. Thesis. Texas A&M University.

Mausbach, M. J., B. R. Brasher, & W. D. Nettleton. 1980. Variability of measured properties in morphologically matched pedons. Soil Sci. Soc. Am. J. 44:358-363.

Grossman, R. B. 1983. Entisols. In L. P. Wilding, N. E. Smeck, and G. F. Hall (eds.). Pedogenesis and soil taxonomy II. The soil orders. Developments in soil science 11B. Elsevier Publishing Co., Wageningen. pp. 55-90.

Rehage, J. A., W. G. Chervenka, & L. P. Wilding. 1982. Soil variability of claypan mapping over two generations of soil survey. Agronomy Absts. p. 238.

Smeck, N. E. & L. P. Wilding. 1980. Quantitative evaluation of pedon formation in calcareous glacial deposits in Ohio. U. S. A. Geoderma 24:1-16.

Sobecki, T. M. & L. P. Wilding. Calcic horizon distribution and soil classification in selected soils of the Texas Coast Prairie. Soil Sci. Soc. Am. J. 46:1223-1227.

Steel, R. G. D. & T. J. H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Co. Inc.

Steers, C. A. & B. F. Hajek. 1979. Determination of map unit composition by a random selection of transects. Soil Sci. Soc. Am. J. 43:156-160.

Wang, C. 1982. Application of transect method to soil survey problems. LRRI Contrib. Num. 82-02. Research Branch, Agricultural Canada, Ottawa, Ontario. Land Resource Research Institute. pp. 1-34.

West, L. T., L. P. Wilding, J. K. Landeck, & F. G. Calhoun. 1984. Soil Survey of the ICRISAT Sahelian Center, Niger, West Africa. Department of Soil and Crop Sciences/TropSoils. Texas A&M University pp. 66.

Wilding, L. P., & L. R. Drees. 1978. Spatial variability: A pedologists view point. ASA Special Publications No. 34 (Diversity of soils in the Tropics). Chapter 1, pp. 1-12.

Wilding, L. P. 1982. Genesis of Vertisols. Proceedings of the 5th International Soil Classification Workshop, Sudan (in press).

Wilding, L. P. & L. R. Drees. 1983. Spatial variability and pedology. In L. P. Wilding, N. E. Smeck, and G. F. Hall (eds.). Pedogenesis and Soil Taxonomy I. Concepts and interactions. Developments in soil science 11 A. Elsevier Publishing Co., Wageningen pp. 83-116.

Wilding, L. P. & K. Flach. 1985. Micropedology and Soil Taxonomy. In L. A. Douglas (ed). Micromorphology and Soil Taxonomy. ASA Special Publication (in press).

Discussion

R. Hammer: Since soil chemical and physical properties covary with differing rates and levels according to the soil forming factors and their interactions, why is more attention not given to the use of multivariate statistics (factor analysis and discriminate analysis) in quantifying and analyzing soil variability?

L. Wilding: Most of the multivariant statistical approaches to soils has been applied in attempts at numerical classification of soils by simple ordination methods of affinity or similarity among soil classes or by using cluster analysis to reduce the number of selected properties for classification to a few. Some of the disadvantages or problems in this approach is that: (1) a single measure of similarity involves enormous loss of information; (2) selection, measurement, and coding of multiple characters are highly subjective; and (3) many different kinds of characters must enter into taxonomic classification. Some of these are not sufficiently quantified to be written into a computer program.

It has also been observed that cluster analysis is effective with small numbers of distinct soil groups or with small areas but in large areas with a number of different soils, clusters either do not exist or are extremely diffuse.

Although this tool is in its infancy in pedological studies, the current availability of high speed computers and increased focus on quantification of soil parameters makes this an area of potential future study. The possibility of coupling multivariant analysis with kriging may overcome some of the concerns that kriging does not permit delineating soil areas with multiple property differentiate. This whole area is open to further development.

J. Bouma: Spatial variability observations are commonly made in soils at the initiation of a soil survey in the U.S. This is necessary to design mapping units appropriate for the scale of spatial variability found in the survey area. For example, if small-scale variation occurs among soils with contrasting properties such that individual delineations of each soil cannot be

delineated at the scale of mapping, then either complex or undifferentiated mapping units would be designed for this situation. Considerations of spatial variability must precede the field mapping phases of the soil survey program to adequately determine which soil conditions dominate in a given landscape and what the scale of repetitive patterns may be. The transect samples collected during the survey or at the end of it are provided to quantify the relative percentages of soil conditions within a given map unit.

The purpose of this comment was to be sure we did not leave the impression that spatial variability in soil surveys was examined only at the end of the survey.

L. Wilding: I agree with your comment and appreciate your making this point.

M. Collins: The number of observations per mapping unit, I feel, very rarely exceeds 50 and never 350! Did I understand Arnold's figure on the graphical solution to binomial confidence limits correctly?

L. Wilding: The figure is just an example of this approach to rapid determination of confidence intervals. In the paper by Wang cited in the text, he has illustrated such graphs for observations ranging from 0 to 50 and 50 to 350. For any given mapping unit delineation, the graph for 0 to 50 observations could be used but in sampling several delineations of the same mapping unit the graph for 50 to 350 observations may be appropriate.

P. Greminger: If we have some information about spatial variability, on what basis do we decide which sampling scheme to choose?

L. Wilding: Perhaps the major decision regarding the sampling scheme to be utilized revolves about the kinds of questions to be answered, the objectives of the work and the nature of the classes being sampled. I personally prefer transect or grid sampling schemes with equal point intervals and elevation control, if possible. These schemes work well in geostatistical analysis and determining soil composition within mapping units. It might be preferable to couple short-

range and long-range sampling to observe spatial variability of differing scales such as in the grid-systematical radial transect scheme illustrated. Finally, the sampling scheme will eventually depend upon the labor and financial inputs that can be allocated to accomplish the given objectives. The text of this paper covers the various pros and cons of sampling schemes.

A. Warrick: My own experience is that water content distributions tend to be somewhat less variable than you indicated. I believe it should be grouped in a lower variability class, say a CV of 10-15%.

The concept of bulking or composite sampling can be handled nicely by block kriging. This was done for determining the variance of the mean for a finite area by Webster and associates in the J. of Soil Sci. This is believe relevant particularly to "additive" property concentrations and for textural classes.

L. Wilding: I appreciate your comment on water content. Values given in this report were based primarily on fine-textured soils with significant cracking character. I suspect in these soils water content is spatially more variable than in coarser-textured soils.

I am aware of Webster's work on block kriging and appreciate your mentioning this point. However, in routine soil survey operations in the U.S., we do not sample sufficient pedons of a given soil for laboratory characterization to permit such kriging methods to be employed. Numbers for kriging application seem to be the major constraint.

J. Hendrickx: Could the CV you presented not be more an expression of sample procedures than of variability of soil properties under consideration?

L. Wilding: I don't think so. These values were taken both from our own studies as well as those from the literature where sample sizes for given parameters were quite variable. It is well recognized, however, that sample size is very important in defining spatial variability. For example, the CV for a given property may be 25-50% across a landscape unit of 5 to 10 ha or more, 5-10% for the area of sampling unit ($1m^2$)

and <5% for the sampling errors associated with laboratory determination. In that sense CV's are covariable with size of the units sampled. This is even more critical in sampling a disturbed soil material (mixed overburden from mining activities) where the CV's will be directly correlated with sample volume. However, for the data presented as relative soil property variability, the CV's are for natural soil bodies of a given mapping unit representing areas of several ha. The observations were for equivalent soil horizons and of sufficient volume or area to express a sampling unit of about $1m^2$.

K. Flach: Pedologists have been occupied with the systematic depth variation in soils or soil horizons; lateral variation can be equally systematic and can be used to characterize kinds of soils. Lateral variation needs to be studied more and included in the characterization and classification of soils.

L. Wilding: Both lateral and vertical variation in soils are important to soil use, management, classification and characterization. Pedologists have long focused on both aspects as a basis to partition soils into landscape and taxonomic classes. I assume, however, that your comment is for soils with similar horizonation but markedly different rates of lateral change. This would influence the areal extent and pattern of a given soil and its border neighbors. Similar soils but with different landscape associates might be classified differently. Patterns of soils and their complexity are important for soil management and use. Draining a poorly-drained soil, for example, often depends on whether it is in a large, nearly level, contiguous unit or as many small, isolated tracts scattered among better-drained sloping analogues. There is good justification for greater efforts to study the lateral variability in soils and potential impacts on use, management and classification.

P. Greminger: Now knowing something about the spatial dependence of the soil water characteristics and some other soil properties, how do you apply this new information to your sampling for the soil survey? And what other information is still needed for an efficient soil survey?

L. Wilding: Firstly, we still are at our infancy in understanding spatial dependence of soil water and other characteristics in soil landscapes. In fact, there is very little information on the spatial dependency of many of the soil properties used as differentiae for soil surveys and soil classification. Given that information, however, the ways it would be used to enhance the quality of a soil survey are given in the summary section of this paper. In short, it would permit us to more accurately convey the mode and magnitude of spatial variability to users of soil resources. It would also foster development of soil surveys with greater quantification of vertical and lateral spatial change.

One area where considerable future effort is needed is the development of class ranges (limits) for given soil properties which are most meaningful for interpretations of soil response and behavior. Properties may be highly variable, but if absolute ranges of spatial difference do not exceed critical class limits, then our interpretations do not suffer from spatial variability Another area for great need is to develop transfer functions for soil properties, such as those discussed in Dr. Bouma's paper, to integrate qualitative and quantitative property evaluations.

The analysis of water quality variations in stream-aquifer systems: nonpoint sources

C. J. Duffy, Utah Water Research Laboratory, UMC 82, Utah State University, Logan, Utah 84322, USA.

Introduction

The persistent problem of nonpoint source pollution, although it received considerable attention by the research community during the 1970's, still ranks as a major factor in the environmental degradation of groundwater and streams in the U.S. Historically, the analysis of diffuse sources of groundwater contamination, such as infiltrating urban runoff, septic tanks, fertilizers, pesticides, and salinity in arid regions, has been greatly complicated by two factors. There has been a lack of conceptual understanding of the subsurface flow and solute transport processes involved, as well as a serious lack of measured data for nonpoint source contaminants. In recent years significant advances in our understanding of transport processes and modeling of soil and groundwater systems have taken place, while industry and government have continued (to some degree) to address the problem of long term monitoring. With regard to transport processes and models, scientific advances have largely been the result of controlled tracer tests in the laboratory and the field, where the input or source has been regulated (i.e. a pulse or step). Regulated inputs facilitate the application of simplified forms of the convective-dispersive transport model. The analysis of the evolution of environmental residuals (i.e. isotopes, pollution and geochemical observations) in natural systems has received much less attention. This may be due in part to the large degree of spatial and temporal variability associated with nonpoint source pollution, and the attendant difficulty in "modeling" these processes. Since tracer tests for nonpoint sources of contamination are not generally feasible, the direct analysis of environmental chemicals seems to be the only alternative for studying these systems.

In the present paper the goal is to review and illustrate a simplified linear reservoir approach to nonpoint source contamination in a stream-aquifer setting. The study will include a heuristic argument as to the physical basis for the lumped model, and a demonstration of the effect of spatial and temporal variablility of a nonpoint source on subsurface out flow concentration. An attempt is made to demonstrate that simplified input/output models are useful for describing the dynamic behavior of nonpoint source contamination of streams or agricultural drains, and that solute

dispersion and spatial variability have little impact on the concentration of groundwater discharge to streams or drains.

The linear reservoir

A practical approach to nonpoint sources of groundwater contamination makes use of the "complete mixing" assumption of the linear reservoir model (Duffy, 1982; Raats, 1981; and Gelhar and Wilson, 1974). This assumption of complete mixing requires that the input to the reservoir be adequately described by its spatial average over the surface of the reservoir, and that the output be characterized by the reservoir average or well mixed concentration in the groundwater system.

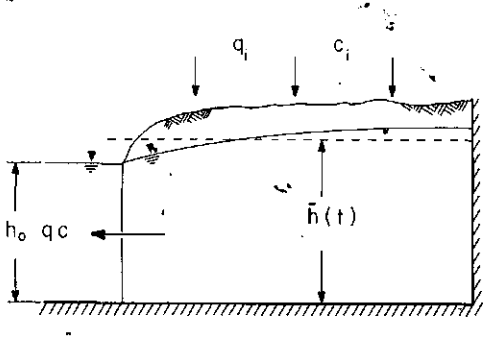


Fig. 1 A linear reservoir representation for a nonpoint source contaminant c_i

Gelhar and Wilson (1974) have shown that a mass balance for a solute of concentration c(t) in the stream aquifer setting of Fig. 1 is given by:

$$dM/dt = m_i - m_o \pm m_s \qquad (1)$$

where

$M = n\bar{h}c$ = average mass of solute c per unit aquifer area
$m_i = q_i c_i$ = input mass flux of concentration c_i per unit area
$m_o = qc$ = outflow mass flux of concentration c per unit area
m_s = other sources or sinks per unit surface area
n = porosity
$\bar{h}$ = spatial average hydraulic head

For the case of steady-state groundwater flow ($q = q_i$) (1) is given by

$$dc/dt + c/T_c = c_i/T_c \pm r' \qquad (2)$$

where r' indicates sources or sinks $[m/L^3/T]$, and $T_c = n\bar{h}/q$ is the average residence time of the system, which can also be written

$$T_c = \frac{V}{Q} = \frac{\text{fluid volume of the reservoir}}{\text{flow through the reservoir}} .$$

In simple terms T_c describes the average time to replace one pore volume of the reservoir by deep percolation or recharge. A general solution to the linear reservoir is given by the convolution integral

$$c(t) = c_I e^{-t/T_c} + \int_o^t h(t - \tau)\, c_i(\tau) d\tau, \; h(t) = (e^{-t/T_c})/T_c \qquad (3)$$

where $h(t)$ is the impulse response function, c_I is the reservoir concentration at $t = o$, and $r' = o$. From (3) one can construct other solutions for $c(t)$ by assigning the form of $c_i(t)$. For example the solution for a step input ($c_i = c_o$) is given by

$$c(t) = c_o(1-e^{-t/T_c}), \; c_I = 0 \qquad (4)$$

For situations of nonpoint source contamination under uncontrolled or natural state conditions $c_i(t)$ is likely an arbitrary function of time. In this case a discrete version of equation (3) is useful

$$c(i) = \sum_{k=1}^{i} \frac{\exp}{T_c} (- (i - k + 1)\, \Delta T/T_c) \cdot c_i(k) \cdot \Delta T \qquad (5)$$

where ΔT is the time interval and k is the time index ($k = 1, 2, 3 \ldots i$).

Physical justification for well-mixed systems

A number of authors have applied reservoir theory to hydrologic problems of flow and solute transport. Among them are Kraijenhoff van de Leur (1958) who examined the problem of nonsteady flow to drains, and Eldor and Dagan (1972) who related the exponential outflow behavior of a uniform contaminant source to convective transport effects. This study also demonstrated that the effect of dispersion would be relatively small at the point of outflow (at the stream or drain) where solutes from all stream tubes are mixed. Gelhar and Wilson (1974), and McClin (1981) have also demonstrated that an exponential solute outflow is characteristic of convective transport for a uniformly distributed, nonpoint source. Raats (1981) provides a discussion of "apparently well-mixed" systems based on travel time distributions in a convective flow field.

The concept of an "apparently well mixed" system can be demonstrated with a simple numerical example. Consider the introduction of a steady, uniformly distributed source of contaminant c_o in the homogeneous and isotropic aquifer illustrated in the inset diagram of Fig. 2. The flow is assumed to be steady and governed

by the Laplace equation (Kirkham and Powers, 1972). The governing equation for two dimensional solute transport of a conservative species in a vertical plane is given by

$$\frac{\partial c}{\partial t} + u_x \frac{\partial c}{\partial x} + u_y \frac{\partial c}{\partial y} = \frac{\partial}{\partial x}\left(D_{xx} \frac{\partial c}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_{yy} \frac{\partial c}{\partial y}\right), \qquad (6)$$

$$D_{xx} = \frac{(\alpha_x - \alpha_y)\, u_x^2}{|u|} + \alpha_y |u| \quad , \quad D_{yy} = \frac{(\alpha_x - \alpha_y)\, u_y^2}{|u|} + \alpha_y |u|$$

where α_x and α_y are the longitudinal and transverse dispersivities, and $u_x(x,y)$ and $u_y(x,y)$ are the horizontal and vertical components of the velocity field determined from the solution to the Laplace equation (Kirkham and Powers, 1972, p. 118). The position of the free surface was estimated by specifying the recharge rate (q_i = .25 cm/day) and the hydraulic conductivity (K = 150 cm/day). Equation (6) was solved using a two-dimensional finite element code called Seftran (Huyakorn et al., 1984), with $c_o q_i$ uniformly specified along the free surface, and no-flux boundaries specified elsewhere.

Figure 2 illustrates the normalized outflow concentration c/c_o at the stream versus dimensionless time t/T_c, for various values of the longitudinal dispersivity α_x. The longitudinal to transverse dispersivity ratio α_x/α_y was taken to be 10 in all cases, while the residence time was calculated to be $T_c = nh_o/q$ = 3.28 years. Comparison of the linear reservoir model with the convection-dispersion results suggests that dispersion has very little impact on the aquifer outflow concentration, and that the exponential outflow behavior of the linear reservoir model provides a reasonably good analogue for describing outflow from uniformly distributed nonpoint sources.

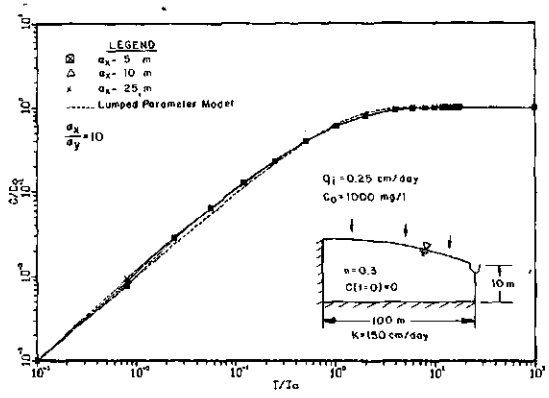


Fig. 2 A comparison of the outflow concentration for the convection dispersion model and the linear reservoir

Spatial and temporal variability of nonpoint sources

It is reasonable to conclude from the previous numerical experiment that spatially and temporally uniform, nonpoint sources in stream-aquifer systems demonstrate a behavior not unlike a "well mixed" linear reservoir. However, as was pointed out in the introduction, inputs to natural systems are subject to highly nonuniform conditions in both space and time. In this section we will relax the uniformity condition and examine the performance of a groundwater reservoir subject to spatial and temporal variations in the source strength. Ideally this analysis should allow c_o to be treated as a concurrent space-time stochastic process. However, the preliminary analysis to be presented here deals with spatial and temporal variations separately. Future work will explore this coupling.

Spatial variability

We first assume that spatial variations in the nonpoint source $c_o(x,t)$ are described as a stationary stochastic process, and thus can be characterized by its mean, variance and correlation structure (correlation scale). We will also make the physically unrealistic assumption that c_o is uniform in time. The approach is to generate various stochastic realizations of $c_o(x)$, input this data to the convection dispersion model described earlier (eq. 6) and examine the outflow concentration from this model . The theoretical spatial correlation structure was assumed to be exponential

$$\rho(k) = \exp(-k/\lambda) \qquad (7)$$

where λ is the spatial correlation scale, k is lag distance, and $\rho(k)$ is the autocorrelation function. Using a lag-one autoregressive algorithm (Salas et. al., 1980), several space series were generated, and these are presented in Fig. 3 for λ = 1 and 10 meters and a coefficient of variation for c_o of, σ/μ = 0.1 and 0.5. Figure 4 illustrates the normalized outflow concentration $c(x,t)/\bar{c}_o$ versus t/T_c for the convection-dispersion model, and the linear reservoir. The most striking characteristic is the very regular behavior of the outflow, demonstrating the large degree of averaging which occurs for areally distributed sources, since none of the spatial fluctuations evident in the field data are observed in the output. It is clear that there are significant differences between outflow from the convection-dispersion model and the well-mixed linear reservoir during early time ($t/T_c < 0.3$). However at latter time ($t/T_c > 0.3$) outflow concentrations are again apparently "well mixed." During this early time period, notice that the simulation with the largest variance and the largest correlation scale also shows the greatest difference with the linear reservoir result, while the simulation with the large variance and the small correlation scale provides the result closest to the linear reservoir. Most of this early time difference is likely due to variations in the simulated

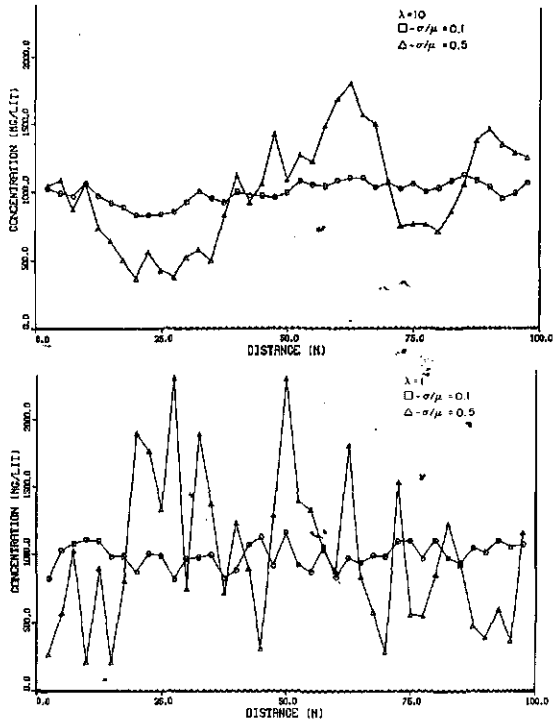


Fig. 3 Realizations of a spatially variable nonpoint source $c_o(x)$ generated from an exponential autocovariance ($\bar{c}_o$ = 1000 mg/l)

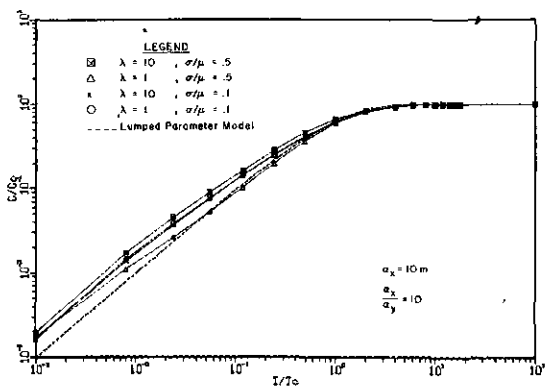


Fig. 4 Normalized output concentration $c/\bar{c}_o$ versus t/T_c from the numerical solution to the 2-d convection-dispersion model (Eq. 6) for the spatially variable input processes $c_o(x)$ of Fig. 3. The dashed line is the simple linear reservoir result Eq. 4.

source strength very near to the drain. It seems likely that most of this difference would disappear if the source strength was realistically simulated as a two-dimensional field, and the groundwater system treated as a three-dimensional one.

Although this preliminary effort remains somewhat inconclusive, it does suggest that complex spatial variations of the input process for nonpoint source contamination do not dramatically affect the concentration of groundwater entering the stream. However, additional research is necessary to better understand what are the limiting conditions where the "well-mixed" assumption may apply to nonpoint source problems subject to spatial and temporal variability.

Temporal variability

If we accept the concept that spatial averaging produces an apparently "well-mixed" system, then the interpretation of temporal variations of nonpoint sources can be examined within the theoretical framework of linear filter theory for stationary stochastic processes (Koopmans, 1974). A linear filter, as illustrated in Fig. 5, simply converts a stationary input time series into stationary output, while the (groundwater) system performs as a filter. The filter can be defined either in the time or frequency domain. For stochastic processes this is most conveniently done in the frequency domain. The approach makes use

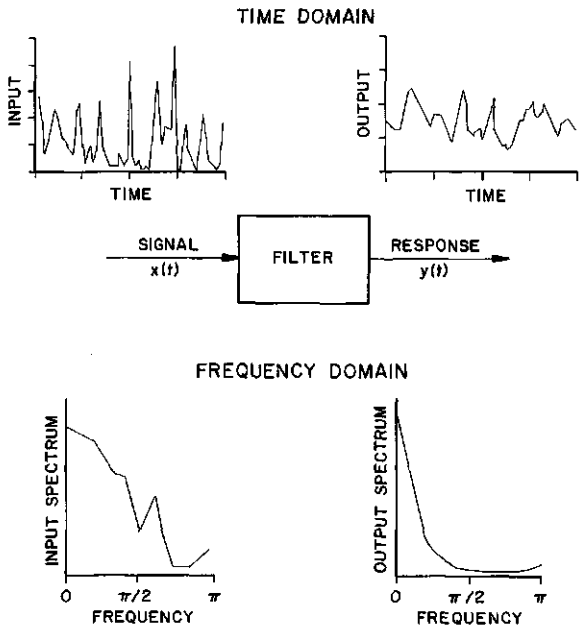


Fig. 5 A linear filter in the time and frequency domain

of the frequency response function from linear systems theory, and the mathematical technique of spectral analysis of time series. Jenkins and Watts (1968) provides a good summary of the theory as well as several examples.

Treating the linear reservoir model (eq. 2) as a stochastic differential equation, Duffy (1982) provides a solution in two parts. The first part, or transfer function, is a measure of amplitude attenuation (actually amplitude squared) between input and output versus frequency, and is presented in terms of the output/input spectral density ratio $\phi_{c_o c_o}/\phi_{c_i c_i}$. The spectral density $\phi(\omega)$, or simply the spectrum, describes the variance distribution of a time series as a function of frequency ω. The second part of the frequency domain solution to the linear reservoir is the phase spectrum $\Theta(\omega)$, which is a measure of the lag time between input and output for each frequency in the record. Figure 6 summarizes the theoretical transfer function and phase spectrum versus dimensionless frequency for the linear reservoir model, along with the output/input variance ratio versus T_c/λ. In this case λ is the temporal input correlation scale for c_i, assuming an exponential autocorrelation in time, similar to eq. 7, with lag k now having units of time. Details of the theory can be found in Duffy (1982).

The frequency domain solution of the linear reservoir model of Fig. 6 can be readily applied to field problems such as parameter estimation (T_c) where the input/output spectra are known, or for characterizing the input time series itself, given T_c and the output time series.

Qualitatively the filtering characteristics of a linear reservoir are quite simple. The transfer function suggests that high frequency variations in the input are attenuated over a broad range ($1 \leq \omega T_c \leq 30$). The negative phase spectrum increases with frequency up to a maximum of $\pi/2$. The output/input variance ratio suggests that contaminant fluctuations in the input will be attenuated to near zero when the residence time is large and/or the correlation scale is small ($\lambda \rightarrow 0$). This suggests that uncorrelated input time series are more effectively filtered than correlated inputs. Duffy (1982) has demonstrated that the frequency domain approach is useful for characterizing the response and estimating parameters of nonpoint source salinity in an irrigated agricultural setting.

Multiple reservoirs

A potentially useful extension of reservoir theory to nonpoint source problems is the area of multiple reservoir systems. Eriksson (1971) provides a general theory for multiple reservoir modeling of composite natural systems, while Duffy et al. (1984) treat an agricultural drainage system as a multiple reservoir system. Figure 7 illustrates two simple arrangements for the linear reservoir, serial and parallel. Also shown is the impulse

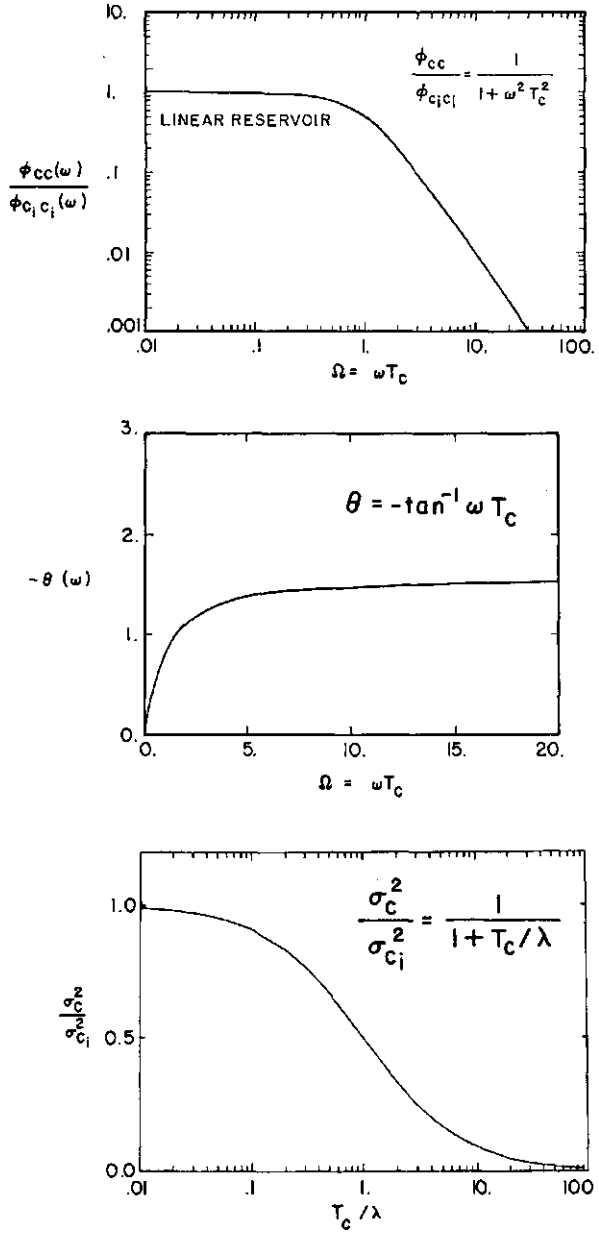


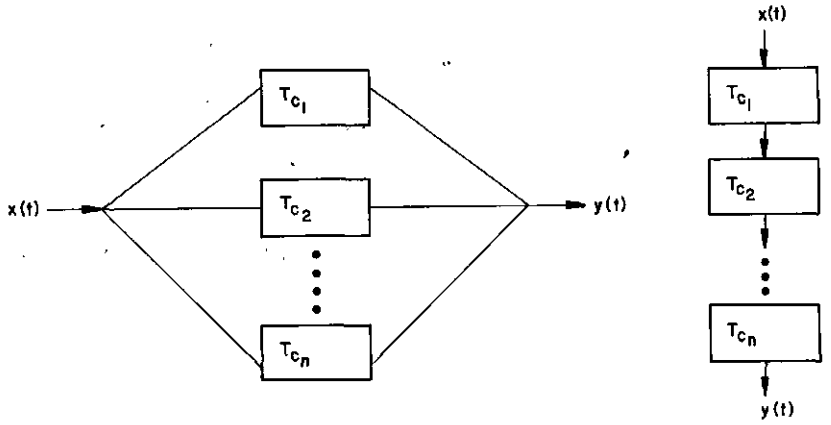
Fig. 6 From top to bottom, the transfer function, phase spectrum and output/input variance ratio for the linear reservoir model.

response function for each system of reservoirs, where x(t) is taken to be an impulse input.

As an example of the multiple reservoir approach, van der Molen (1973) has examined the process of leaching saline soils with fresh water using serial reservoirs. In this case each reservoir corresponds to a different soil layer, with outflow from the overlying layer producing inflow to the next. Comparison of the serial reservoir solution of van der Molen (1973) for leaching with the convection dispersion model indicated a close agreement between outflow concentrations for as few as four reservoirs. For a situation of solute displacement in a layered aquifer with different hydraulic conductivities, Simonett (1981) has used a combined serial-parallel arrangement of groundwater reservoirs to describe chloride transport to drains. In general the flexibility of using systems of reservoirs in series and/or parallel would seem to provide a useful tool for analysis of nonpoint source systems.

Summary

The basic goal of this paper has been to emphasize the utility of simplified models of nonpoint source contamination, and their relation to problems of spatial and temporal variability of the contaminant source. There has been no attempt here to incorporate



$$h(t) = \sum_{i=1}^{n} \frac{1}{T_{c_i}} e^{-t/T_{c_i}},\ i = 1,n\ ,\ h(t) = (T_c\Gamma(n))^{-1} e^{-t/T_c} (t/T_c)^{n-1}$$

Fig. 7 a) Parallel linear reservoirs with unequal residence times T_{c_i}, and the corresponding impulse response function. b) serial reservoirs and the impulse response function for constant T_c (Nash, 1957), $\Gamma(n)$ = gamma function

the effects of spatial and temporal variations in the hydraulic properties of this system (hydraulic head, hydraulic conductivity and porosity), or in incorporating the important effects of the unsaturated zone. and thus the results are incomplete. However, it does seem reasonable to conclude that simplified lumped parameter models, given an appropriate conceptual framework can be very useful for studying the integrated, input-output effects of stream-aquifer or agricultural drainage systems. In addition it was found that magnitude of dispersion, and stationary spatial variations of the nonpoint source, have relatively little impact on the outflow concentration to streams or drains.

References

Duffy, C.J., L.W. Gelhar and P.J. Wierenga, 1984. Stochastic models in agricultural watersheds. J. Hydrology 69: 145-162.

Duffy, C.J., 1982. Stochastic modeling of spatial and temporal water quality variations in groundwater. PhD Dissertation, New Mexico Institute of Mining and Technology, Socorro, New Mexico.

Eldor, M. and D. Dagan, 1972. Solutions of hydrodynamic dispersion in porous media. Water Res. Res. 8: 1316.

Eriksson, E., 1971. Annual Review of Ecological Systems, 2, 67-84.

Gelhar, L.W. and J.L. Wilson, 1974. Ground-water quality modeling. Ground Water 12(6): 399-408.

Huyakorn, P., et al., 1984. Seftran: A simple and efficient flow and transport code. Tech. Report prepared for Holcomb Research Inst. IGWMC, by GEOTRANS, Inc.

Jenkins, G.M. and D.G. Watts, 1968. Spectral analysis and its applications. Holden Day, San Francisco, California.

Kraijenhoff van der Leur, D.A., 1958. A study of nonsteady groundwater flow with special reference to a reservoir coefficient. Ingenieur 70B: 87-94.

Kirkham, D. and W.L. Powers, 1972. Advanced soil physics. Wiley Interscience, New York.

Koopmans, L.H., 1974. The spectral analysis of time series. Academic Press, New York.

McLin, S.G.. 1981. Validity of the generalized lumped parameter hydrosalinity model in predicting irrigation return flow. PhD Disseration, New Mexico Institute of Mining and Technology, Socorro, New Mexico.

Nash, J.E., 1957. The form of the instantaneous unit hydrograph. Int. Assoc. Sic. Hydrology, Pub. 45(3): 114-121.

Raats, P.A.C., 1981. Residence times of water and solutes within and below the root zone. Agric. Water Management 4(1,2,3): 63-82.

Salas, J.D., J.W. Delleur, V. Yevjevich and W.L. Lane, 1980. Applied modeling of hydrologic time series. Water Resources Publications, Colorado State University, Fort Collins, Colorado.

Simonett, M. J., 1981. Irrigation return flow modeling at San Acacia, New Mexico. Independent Study Paper, New Mexico Inst. of Mining and Tech., Socorro, New Mexico.

van der Molen, W.H., 1973. Salt balance and leaching requirement. In: Theories of field drainage and watershed runoff, Chapt. 9, Vol. II, International Institute for Land Reclamation and Improvement, Wageningen, Netherlands.

Acknowledgments

The research on which this paper is based was financed in part by the United States Department of the Interior as authorized by the Water Research and Development Act (P.L. 95-467) (Contract No. 374702).

Discussion

B. Shumway: You have nicely modeled input-output relations in time in terms of a lumped parameter system which is strictly "causal", i.e., the ouptut is strictly dependent on the past values of the input. In space, there is no clearly defined past. How would you handle this case?

C. Duffy: It is true that this study concentrates on input-output behavior, however, we are modeling the system as a two-dimensional vertical flow and transport problem, with recharge rate and concentration a function of space ($q_i(x)$ and $C_o(x)$).

If you follow a particle along a given streamline from its origin on the water table surface to its point of outflow at the stream, then each particle does have a clearly defined history with respect to space. That is, particles that originate at a point on the water table some distance from the stream have a longer travel time than those that originate close to the stream. The point of the spatial correlation structure, has a relatively minor effect on the outflow concentration, or in other words, the outflow depends primarily on the spatial mean input concentration C_o. This is due to the mixing effect of all streamlines at the point of outflow. In this case spatial variability of nonpoint sources has only a secondary effect on groundwater outflow at the stream due to the natural mixing processes in the aquifer.

It is true, however, that the ARI model which I use to generate the spatially variable input process is not really appropriate, since it only allows correlation in one direction (i.e. a causal system only depends on (X-1) values of the spatial process). An improved generation technique would use both (X + 1) and X - 1) terms to generate the input process.

A. Warrick: Could some of the differences attributed to correlation length and variances be also dependent on the particular value of the source chosen near the outlet? These might dominate the early breakthough and disappear at longer times.

C. Duffy: I believe that this is exactly the case. That is, early breakthrough of the source of contamination near the stream is the cause of the deviations from the lumped parameter result. The problem with this experiment is that we only consider the input process, or nonpoint source, to be one dimensional $C_o(x)$. Effectively the problem we are solving amounts to a single slice in a much larger watershed or drainage situation. If we simulated the more realistic three-dimensional groundwater system with a two-dimensional, a really distributed source, I expect that these early time deviations would disappear by averaging and the "well mixed" outflow assumption would still be apparent.

Measurement and interpretation of spatially variable leaching processes

R. J. Wagenet, Department of Agronomy, Ithaca NY 14853, USA

The variability of leaching processes in the unsaturated zone has received much attention in the last decade of soil science research. Measurements of the movement of water and solutes under field conditions have demonstrated that accurate estimation of fluxes must consider spatial variations in the soil physical, chemical and biological processes that affect water and solute distribution. A number of statistical and mathematical approaches have demonstrated with relatively few field measurements that variation in hydraulic conductivity, water flux, pore water velocity, and dispersion produce substantial variation in water and solute distribution in the soil profile. In the course of producing these sometimes elegant descriptions, there has been little development of similarly advanced sampling technology or the investigation of variability as it relates to management of water and chemical applications. These issues impact both the amount of spatial variability that we estimate exists in the landscape and our subsequent interpretation of it. Before designing further field investigations and developing new mathematical and statistical interpretive techniques, it is important that we momentarily pause and consider several rather fundamental questions which have strong implications with respect to measurement and interpretation of spatially and temporally variable leaching processes. These questions are:

1. What is the influence of the physical size of a measurement or of a particular measurement technique upon the degree of estimated variability?
2. How much variability in water and solute movement is due to inherent (intrinsic) soil variability (arising from soil forming factors) and how much is due to imposed (extrinsic) variability (non-uniform or geometrically patterned applications of water and solute or the influence of vegetation patterns)?
3. Is there a temporal stability to the spatial structure of variability?
4. What is the influence upon spatially variable water and solute fluxes of spatial and temporal variations in the physical, chemical and biological processes that affect

leaching (e.g. adsorption/desorption, precipitation/dissolution, crop uptake, degradation, mineralization)?

5. If models are to be used to describe the displacement of water and solutes in spatially variable systems, what is the appropriate model form to be used for research versus management purposes?

These questions have been only infrequently addressed in the soil science literature by those scientists interested in leaching processes, with the result that little consensus exists concerning appropriate measurement, interpretation or simulation procedures. A perusal of the soils literature for the last several years suggests that issues related to data analysis and interpretation are receiving much attention through application of autoregressive, geostatistical, spectral and related techniques. Similarly, the development of modeling approaches that consider stochastic versus deterministic relationships is also an attractive topic for many scientists. Yet, the fundamental first step of making reliable measurements that are representative of the natural system has not always been appreciated. We must be careful in our enthusiasm to study spatially variable soils that we do not generate substantial data sets of relatively worthless information because of lack of appreciation of basic principles of measurement and interpretation. This issue seems particularly relevant to scientists interested in the variability of leaching processes. The objective of this paper is to review several studies that have indicated the types of problems encountered in field sampling and data interpretation.

Sample size and representative elementary volumes

One of the first questions to be answered in any field study of variability is the measurement method to be used. Often the choice of a method is based upon previous experience of the investigator and the cost and logistics of collecting an adequate number of samples. There is usually very little concern for the implications posed by use of a particular technique with respect to the degree of spatial variability that will be measured. For example, if the technique involves use of a calibrated device, such as a neutron moisture meter to measure soil water content, the statistics of the calibration curve can (and often are) considered when water content at a particular location is calculated. However, if the estimation of water content is by gravimetric sampling, there is usually little consideration of whether a 2-inch or 4-inch bucket auger used at 50 locations in the field will provide sample populations of similar mean and variance for soil water content. The relationship between the volume or size of a sample and the variability of a number of samples is widely appreciated (Fig. 1) with methods developed to treat it. Despite several examples in the soil science literature that indicate its importance, sample size has seldom been considered as a fundamental issue to be resolved before initiating a sampling program.

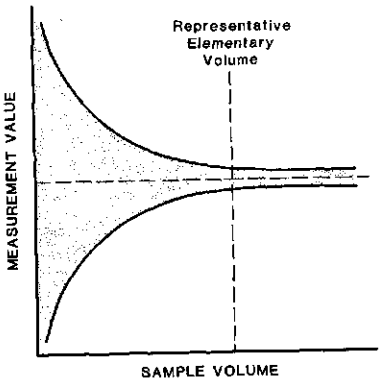


Fig. 1 Relationship between sample volume and possible measurement values

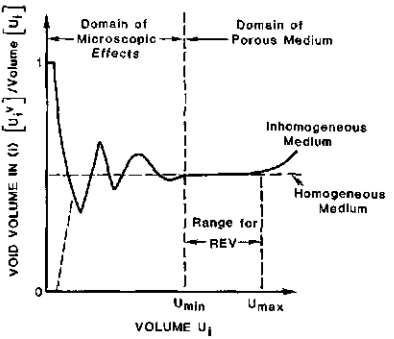


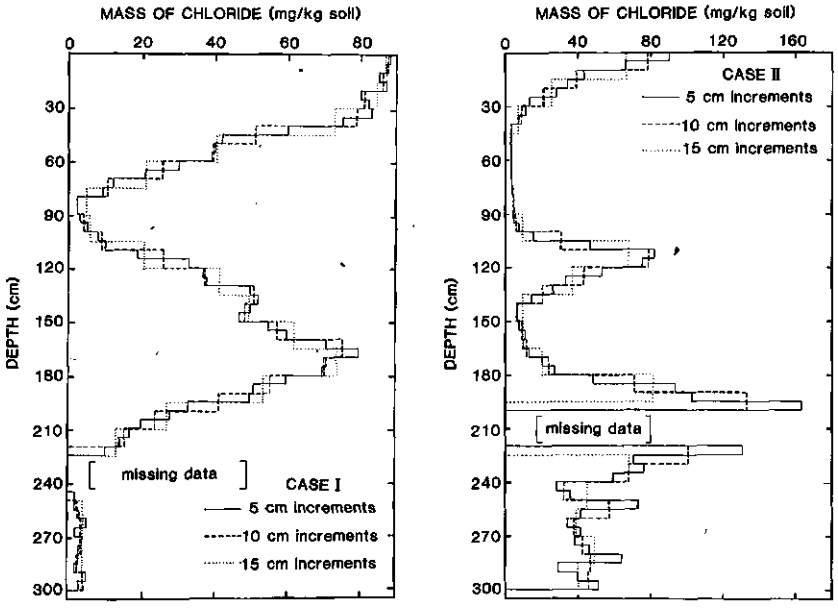
Fig. 2 The representative elementary volume defined by continuum theory

In soil leaching processes, as in other disciplines that consider flow phenomena, a continuum approach is commonly employed (Bear, 1979). This conceptualizes the flow system as heterogeneous at the microscopic level, with solid space and pore space distributed throughout the domain occupied by the soil. This microscopic system is difficult to describe mechanistically, and we therefore do not consider processes at the microscopic level of pore-pore interactions. Rather, we consider a representative elementary volume (REV) of the soil (or porous media), which can be defined on a macroscopic level, and which is continuous through space (Fig. 2). The property of interest within the system is defined to be some average value of its microscopic variation within the REV. When we discuss spatial variability two issues must be considered. First, what is the correct value of the REV for the soil and property of interest, and second, what is the spatial relationship of successive REVs in a spatial domain? Geostatistical analysis treats the latter issue through techniques described elsewhere in this symposium. The former issue of a correct REV quickly becomes an issue of sample size related to our assessment of spatial variability. It is useful to review a few studies that illustrate the need for determining the REV and that are drawn from cases that have implications for soil leaching processes.

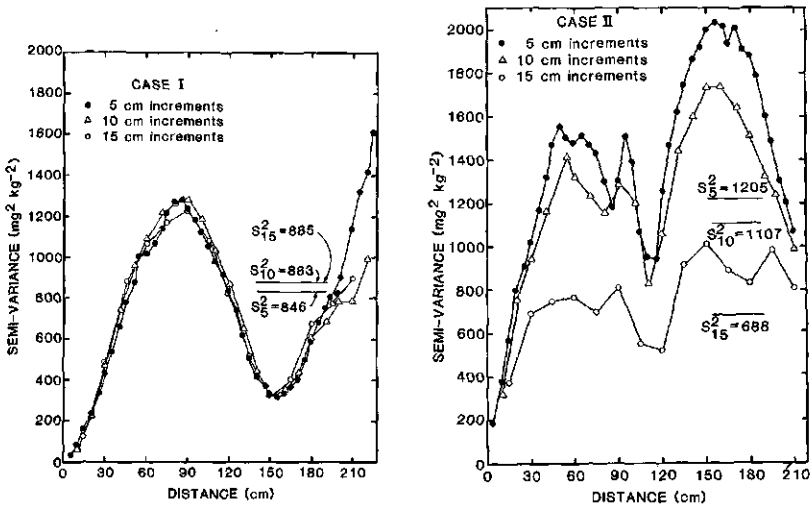
Case 1. The volume-variance relationship in core samples.

A continuous core sample separated into finite depth increments is often used as a means of measuring solute or water distribu-

tion with depth in the soil profile. Data generated in a core with enough depth increments can be treated by several techniques (e.g. Gelhar et al. 1983) as a step in analyzing the spatial variability of water and solute movement. However, a fundamental assumption in geostatistics is that the regionalized variable (in this case solute content) is measured at a point in the system which has an infinitely small volume or geometric support associated with it. In a core, as the length of a depth increment increases, this assumption is less valid, with the result that the analysis of spatial relationships is no longer comparing two points separated by a distance but two average values over a core length separated by the distance between their midpoints. The size of the core increments influences the estimated degree of spatial variability, as demonstrated in Figures 3-6. Consider two hypothetical profiles (Fig. 3 and 4) of measured chloride (mg/kg soil) resulting from application of salty water, as would be determined by taking samples at increments of 5, 10, or 15 cm. Case 1 presents a smoothly varying



Figs 3 and 4 Hypothetical distributions of chloride in smoothly varying (Case I) and irregularly varying (Case II) cases.



Figs 5 and 6 Semi-variograms of chloride content developed for different sampling depths from data of Cases I (Fig. 3) and II (Fig. 4).

profile, and Case II a more irregular chloride distribution. Calculated semi-variograms for Case 1 (Fig. 5) demonstrate very little difference in the estimate of spatial structure whether the 5, 10 or 15 cm core size is used. However, the data of Case II lead to different shaped semi-variograms as the core size changes (Fig. 6), due in great measure to the irregular chloride distribution. If functions were fitted to each Case II semi-variogram they would be different, as presumably would be the results of any next step taken. Interpretation of the spatial variation of leaching processes from such data, whether by correlation, spectral or other analysis can be expected to suffer from such considerations until we better determine, for example, the REV for chloride leaching in different soils. As further stochastic or mixed stochastic-deterministic descriptions of leaching processes are derived, the use of core data to test the models can be expected to increase, and although there are statistical methods available to treat the volume-variance relationship (such as regularization), no soil leaching studies have so far considered this issue.

Case 2. Soil moisture sampling

The physical size of the sample collected has been shown to influence the degree of measured variability of field soil water content. As an example, Hawley, et al. (1982) collected eight different sized cores from ten locations in a $2m^2$ area. Core length was a standard 100 mm, with core diameters ranging from

Table 1. Variability of water content measured with different core sizes on three dates (Hawley et al. 1982).

Core Size diameter (cm)	Measured Soil Water Content (%)					
	Mean	Variance	Mean	Variance	Mean	Variance
	Aug. 14		Sept. 18		Sept. 26	
10.2	9.755	0.408	10.539	0.068	15.346	0.434
7.6	9.804	0.387	10.650	0.296	14.747	0.464
5.1	9.699	0.251	10.781	0.180	14.936	0.094
3.8	9.709	0.489	10.834	0.113	14.916	0.355
2.5	9.029	0.391	11.024	0.298	14.411	0.917*
1.9	9.067	0.803	11.227	0.266	14.700	0.924*
1.3	8.964	0.935	11.750	1.937*	14.830	0.949*
0.95	8.494	3.482*	13.025	1.262*	15.013	1.905*

*indicates variances significantly higher ($\alpha = 0.01$) than other variances on that date.

10.2 cm (4 in) to 0.95 cm (0.375 in). This produced sample volumes ranging from 824 cm^3 to 7 cm^3. Sample collection was accomplished on three different dates with the results dependent upon the date (Table 1). The variability of the soil water content was found to be a function of the time elapsed between the sampling date and the last previous rainfall. The wettest sampling date, September 26, followed two days in which the total rainfall was more than 15 mm; the driest, August 14, had been preceded by two days without rain. The four smallest core samples all exhibited variances greater than the large samples on the wettest date, but on the driest date only the variance of the 0.95 cm diameter core was significantly greater. The intermediate date, September 18, was also intermediate in wetness, with only the two smallest cores more variable than the other sizes. These results demonstrate not only that smaller sample volumes are more variable than large ones, but also that there is a relationship between mean soil moisture and the minimum desirable sample volume. The results also indicate that the local soil moisture variability decreases as the average soil water content decreases, which is itself a function of the time since the previous rainfall or irrigation event. This implies that the REV for water content will have to be large enough to accommodate the effects of transient soil water conditions and will probably lead to different REVs for different soil types.

Case 3. Variability of soil solute content

A similar study to Case 2 investigated the relationship between measured chloride distribution with depth and sample volume (Hassan et al. 1983). Two core samplers were used, a 7.9 cm diameter auger and a 2.1 cm diameter sampling tube. A series of soil samples were collected from plots to which had been applied a known quantity of chloride. These samples were then dried, extracted and analyzed for chloride. There was higher total recovery as well as higher variation among values from the smaller sampler. The coefficient of variation (CU) for the total recovery was 8% for the large sampler and 14% for the small sampler. The implications of these results for leaching studies is more obvious when the estimated distribution of chloride with depth is considered (Fig. 7), as values generated from the small samples indicated the chloride content to be more variable in the profile than did values from the large samples. The conclusion from such a study is that if measured concentrations are to be used to develop estimates of the variability of solute fluxes in a field or are to be used to develop or test models, it is obviously important to take as large a sample as

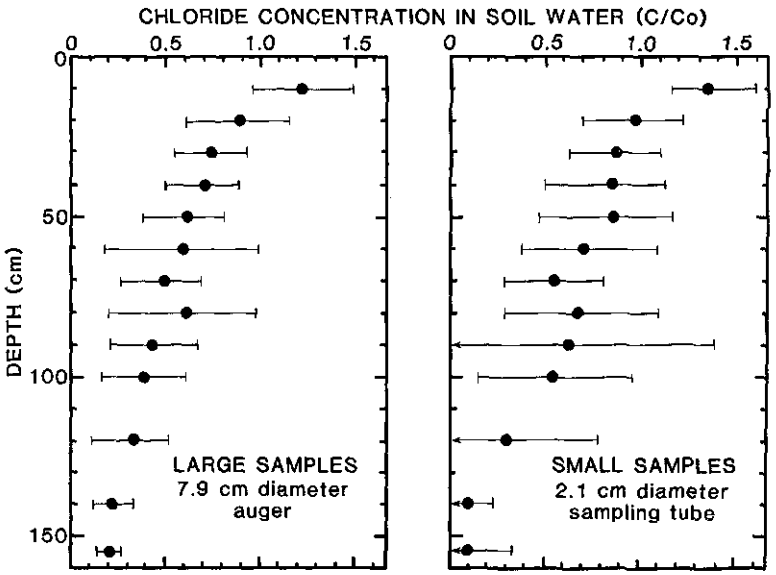


Fig. 7 Chloride concentration measured in the field with two different sized samples (from Hassan et al. 1983).

possible to insure that a REV is being sampled. Additionally, Hassan et al. (1983) noticed a possible effect from contamination of deeper samples by surface material when using the larger auger, and exclusion of soil from the small sampling tube due to friction and compaction. These very practical concerns obviously influence results and should be identified as components of sampling programs when spatial variability is the issue to be treated.

A final point before leaving the issue of measuring soil solute variability concerns the use of porous ceramic cups as extraction devices to measure soil solution concentrations. The volume of soil sampled is a function of the negative pressure applied to the cup, and the soil water content. The sample collected reflects the contributions of those pores that will empty in response to the applied suction. It is almost certain that in many soils this sample will not constitute a REV, yet it is a methodology that is widely used. A very practical point, often the determining factor, is that the porous ceramic cup is the only available means to repeatedly and directly sample the same point in the profile. However, it appears that ceramic extractors do not always provide reliable estimates of the true soil solution as they do not sample all water in a localized volume, and in many cases do not collect sample from a REV of soil. Their use should be carefully considered in future studies of solute spatial variability.

Case 4. Variability of hydraulic conductivity measured on extracted core samples

It is a common practice to collect cores in the field in as undisturbed a condition as possible, transport them to a laboratory and there subject them to a series of standard tests to determine hydraulic conductivity (K). The influence that the size of the core has upon the estimate of K is often neglected, but the effect can be substantial. Anderson and Bouma (1973) collected ten undisturbed cores of 7.5 cm diameter for each of four core lengths: 5, 7.5, 10 and 17 cm. The study focused on an argillic horizon of a silt loam soil with abundant root channels and subangular blocky peds. The value of the saturated hydraulic conductivity, K_{sat}, was measured using standard laboratory techniques for core samples. Mean values of K_{sat} were found to decrease as the core was lengthened (Fig. 8). The variance of the ten samples also decreased when K_{sat} was measured in longer cores. These results were interpreted by considering the relationships between soil-water flow, soil structure, and a finite-sized core sample. It was concluded that larger pores tend to become discontinuous throughout longer cores (much as they would in the field), which would result in a considerable drop in hydraulic conductivity. Shorter cores would more often be characterized by the presence of continuous pores through the length of the core, thereby producing generally higher K_{sat} values. The longer cores produced more

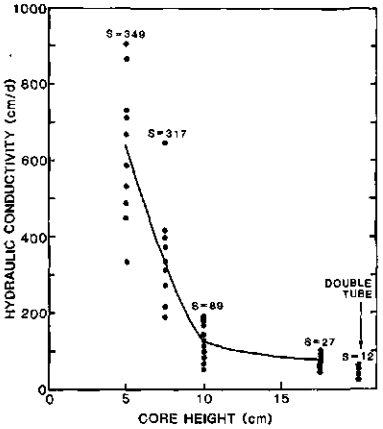


Fig. 8 Saturated hydraulic conductivity measured in different size undisturbed core samples (from Anderson and Bouma, 1973).

consistent K_{sat} values for the same reason, as attenuation of a large pore at any point in the core above the outflow would produce a similar reduction in K_{sat} in replicate cores. The values of the 17 cm high core were closest to those measured in-situ by a double tube method although the mean of the former was still greater than the mean double tube value by about a factor of two (Fig. 8). It is obvious that use of a core of insufficient size can provide inaccurate estimation of the spatial variability of K_{sat}, even if many replicates of the core are used. Not only does the core size influence the mean and variance of a sample population, the relationship of the values determined on any core to a "real", in-situ determined K_{sat} is not clear.

Case 5. Variability of hydraulic conductivity measured in-situ

Recognition of the influence of soil structure upon the value of K_{sat} measurements from different sized cores leads to an additional question. What influence does the physical size of an in-situ sampled area have upon estimation of K_{sat}? That is, if soil structure has an influence in the vertical dimension then does it also play a role in the horizontal through the number of peds included in an in-situ, attached soil column? This question has recently been addressed in a field study (Lauren, 1984) that measured K_{sat} in-situ using five different sized soil columns carved from a textural B horizon but with the bottom of each column left attached to the soil profile. The procedure used on the three largest-sized columns (160x75 cm, 120x50 cm,

50x50 cm) was adapted from the column method of Bouma (1977), and consisted of encasing in gypsum the four exposed vertical sides of the excavated column leaving the top surface exposed for infiltration of water. Each column was 20 cm in height. Steady state infiltration of water into the column was used as the measure of K_{sat}. The two smaller-sized columns were formed by inserting into the 50x50 cm square two circular sections of pipe, one 20 cm in diameter and the other 7.6 cm diameter. K_{sat} was also determined on these columns, in the attached condition, by an infiltration measurement. The experiment was organized such that each set of five measurements was made at approximately the same x-y spatial coordinate location. This was achieved by locating, for the three largest column sizes, each successively smaller measurement within the confines of the previous column. The only deviation from this design was that the two circular columns were taken adjacent to each other from within the 50x50 cm square. The experiment was repeated at 40 locations separated by 10 m on a 400 m transect. The results (Table 2) demonstrate that the sampled populations of K_{sat} from all column sizes were strongly log-normal. The mean, mode, median and variance of these populations depend greatly upon the size area sampled. Excluding the largest sample size, an increase in surface area of the column produced a general decrease in the mean and variance of measured K_{sat}. This trend was not consistent at the largest size for unexplained reasons. However, the coefficients of variation (CV) for the three largest sizes were essentially identical. The CVs for the two smaller sizes were also identical although greater in value than for the larger samples. This behavior of smaller sampling areas can be explained by considering the probability of any small

Table 2. Saturated hydraulic conductivity measured in-situ on five different sized volumes (Lauren, 1984)

Column Size (cm)	Column Area (cm^2)	Saturated Hydraulic Conductivty					
		Mode	Median	Mean	Variance	CV	Number
		----	(cm/day)	-----	(cm^2/day^2)	(%)	of Obs.
160 x 75	12000	8.6	17.6	25.0	647	101	40
120 x 50	6000	5.5	10.9	15.4	233	99	39
50 x 50	2500	5.4	11.3	16.3	290	104	40
20 (dia.)	314	6.9	20.3	34.9	2391	140	37
7.6 (dia.)	45	4.9	16.0	28.9	1897	150	38

sample containing within it a representative cross-sectional area of soil, complete with the cracks, channels or preferred pathways that determine K_{sat}. The smaller the sample, the greater the probability that the flow regime will occasionally be dominated by a single crack or channel. This will result in both a larger mean and variance of the sampled population, as the occasionally extreme value of K_{sat} will increase both these statistical moments. It remains to quantify the relationship between these values and soil structural units, such as the ped size, degree of cracking or presence of large pores.

Geostatistical analysis of the five transects measured in this study produced different estimates of spatial structure depending upon the column size used (Fig. 9). The semi-

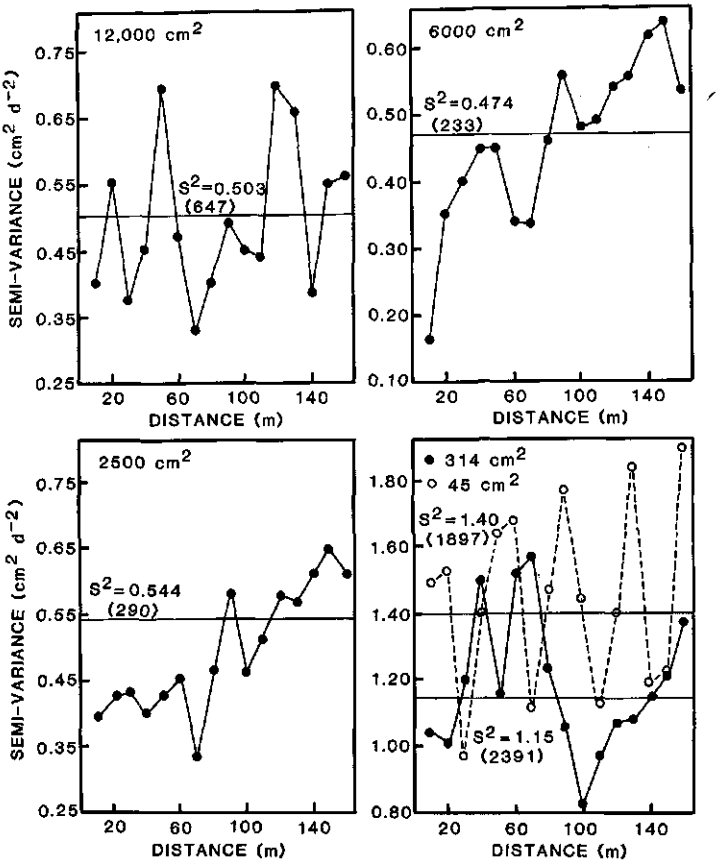


Fig. 9 Semi-variograms of saturated hydraulic conductivity measured in-situ over different sized areas (from Lauren, 1985).

variogram of K_{sat} values from the largest and two smallest columns showed no spatial structure, while a spherical or linear model could perhaps be fitted to the semi-variograms calculated using the data from the two intermediate column sizes. This is not an unexpected result, particularly if the results from the small columns are misleading due to the size area sampled not being representative of the average cross-sectional properties of the soil (i.e. the REV for K_{sat} is greater than 314 cm^2 by 20 cm deep for this soil). There is also, apparently, an effect upon estimated spatial structure produced by using too large a sample size (12,000 cm^2) indicating perhaps an error in experimental methodology when trying to apply the column method to so large a block. The further analysis and interpretation of this data is in progress.

These results have obvious implications for in situ studies of hydraulic conductivity. The same degree of variability is not measured by the column method in different soils, as a CV of only 10% was measured by Baker and Bouma (1976) using 25 cm diameter columns in silt loam soils. The proper size column to be used (the REV) for measurement of K_{sat} is obviously influenced by soil structure. Criteria have been presented (Bouma 1983) based upon soil structure for separating soils into classes of REV appropriate for measurement of K_{sat}. These hypothesized classes are relatively untested, but represent a logical approach to deciding how large a sample constitutes a REV in different soils. It is clear that until we resolve the relationship between soil structure, sampling size and the degree of estimated spatial variability, the extrapolation of results from one field study to another situation will be a tenuous proposition. A number of studies (e.g. King and Franzmeier 1981; Nowland, 1981; McKeague et al., 1982) have had mixed success in predicting K_{sat} from soil pedological measurements, perhaps due to lack of information on the REV for K_{sat} in different soils. However, when appropriate measurement techniques are used, and spatial structure is measured in hydraulic properties, calculation of integral scales and estimation of field sizes and sample numbers are possible (Russo and Bresler 1982). These techniques offer promise for extrapolating the results of a study at one site to other geographic locations. However, without reliable measurements based upon knowledge of the REV, use of such equations will be only an academic exercise. Increased attention must be paid to the relationships between soil structure, hydraulic properties and the REV, particularly if a long-range objective is to relate measured spatial variability to common and widely used soil data bases, such as the soil survey.

Temporal stability of spatial variability

A basic premise of studies that measure spatial variability of leaching processes is that the spatial structure of variability (if any) is preserved over time. Not much research atten-

tion has been paid this issue, although intuitively it seems a reasonable assumption. As spatial variability in water and solute movement is dictated by variations in the geometric arrangement of particles in the soil matrix, it can be presumed that in most soils these arrangements are relatively stable over time. This is visually apparent from observation of wet and dry spots in the field, indicating localized positions of relatively rapid or slow water movement. The lack of research attention paid to temporal stability of spatial variability is understandable, particularly with respect to soil parameters such as hydraulic conductivity that are used to predict water and solute movement. Most techniques used to measure these parameters are tedious and labor-intensive. This often makes multiple observations on one field at one point in time a logistical nightmare, let alone repeated sets of multiple observations, as would be needed to study temporal variability.

Vachaud et al. (1984) have recently presented a straightforward method of assessing, and then using for further sampling programs, the temporal stability of the structure of spatial variability. In their study, a neutron moisture meter was used to measure soil water content every 10 cm to a depth of 100 cm at 17 sites in a 2000 m^2 grass field. Data were collected 24 times over a 2.5 year period at irregular intervals. Soil water storage in the first meter was calculated for each access tube at each sampling date. Geostatistical analysis of the data showed that water storage values were spatially independent, but classical statistics showed that at each sampling time the water storage values were normally distributed. As an example, Fig. 10 (Vachaud et al. 1984) presents the cumulative probability density function for the two extreme situations: the wettest

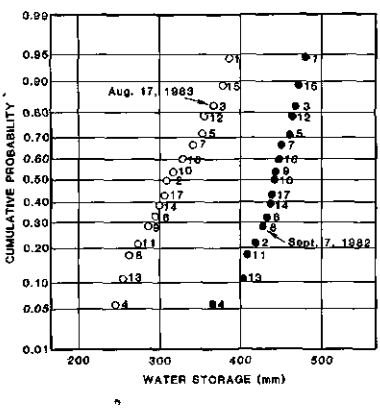


Fig. 10 Temporal relationships of spatial variability of soil water storage (from Vachaud et al., 1984).

(Sept. 7, 1982) and the driest (Aug. 17, 1983). The important characteristic of these plots is the relative position in the cumulative probability of each individual access tube, here denoted by the number located to the right of the plotted point. Note that although the total amount of water in the profile at each access tube changed over time, the mean soil water storage for the field could have been determined by sampling only at position 10, and the 95% probability limits on the water storage determined using locations 1 and 4. This is potentially quite a useful result, as it indicates that positions in the field can be identified with one soil sampling at time zero and used thereafter for repeated sampling to estimate, e.g. the mean and variance of water storage. This infers that simple gravimetric sampling and analysis at the first stage of a field study could identify statistically important field locations at which could then be located more detailed and intensively monitored sampling devices. Only a few such sites, intensively monitored, could then be used as measures of the limits of field variability.

Extrinsic and Intrinsic Variability

The spatial variability of soil properties has long been recognized by pedologists. The relationships between parent material, topography, elevation, vegetation, time and the resultant soil's physical-chemical properties has led to appreciation of the variable nature of soils in the landscape. This variation arising from natural processes can be termed intrinsic soil variability, as it is a component of the essential nature of the soil and has arisen at least partially from within the soil body by weathering and leaching phenomenon. This intrinsic variability is manifested in such forms as the gradation of particle sizes in alluvial outwash soils, the development of distinguishable differences with depth (the soil profile), and the presence of macropores, structural cracks, and root channels. Such variability should be distinguished from the imposed variability of water and solute within the soil profile that results from the distinct application geometries used in many soil-water-chemical management programs. The imposed, or extrinsic, variability can result from patterned irrigation application (furrow or sprinkler) and patterned fertilizer or pesticide applications in bands or rows behind sprayers. This patterned variability will often persist over time, depending greatly upon the dynamics of the flow regime, which is itself variable according to the intrinsic variability. These issues should be recognized, particularly as soil sampling programs are designed that will be incorporated into other studies of chemical redistribution within the soil profile and in which extrinsic variability is to be imposed upon the intrinsic. For example, Hornsby et al. (1983), in a field study in which aldicarb was applied in bands, investigated the proper location of soil sampling efforts on successive dates to monitor the variability of aldicarb leaching. They found that the pattern of leaching variability in their very sandy soil was almost

completely determined by the uniformity and separation distance between band applications. That is, the extrinsic variability defined the limits of subsequent solute leaching variability. Whether this result would be transferable to other soil types is not resolved. It is clear, though, that these two types of variability must be separately considered, particularly as results from a soil series in one location are extrapolated to the same series in another location that is not subject to the same management (different extrinsic variability).

Such issues have been recognized in studies of the spatial variability of pesticide applications, and the results demonstrate the degree of effect that can be imposed. Taylor et al. (1971) collected a total of 108 cores in a 0.13 ha field plot soon after its treatment with dieldrin. A 50-fold variation in dieldrin was measured (overall CV of 80%), with much of the observed variability attributed to application using a boom sprayer followed by a discing operation. Similar extrinsic variability has been reported in a number of other studies (Fryer and Kirkland 1970; Harris et al. 1949: Wauchope et al. 1977), including one (Walker and Brown 1983) which demonstrated that careful application methods could reduce measured CVs compared to those observed under more conventional procedures.

Leaching variability in non-cropped and cropped situations

The studies of Nielsen et al. (1973) and Biggar and Nielsen (1976) that demonstrated the spatial variability of the steady-state infiltration rate, K(θ) relationship, apparent diffusion coefficient (D), and pore water velocity (v) stimulated much subsequent theoretical and experimental activity (e.g. Van De Pol et al. 1977; Carvallo et al. 1976; Luxmoore 1982; Hornsby et al. 1983; Vauclin et al. 1983). These studies have confirmed that water flow and solute transport are quite spatially variable. Generally, it can be stated that field measured sample populations of rate parameters related to leaching, such as infiltration rate, K(θ), D and v are characterized by skewed distributions, often log-normal. Sample populations of capacity parameters, such as bulk density, saturated water content and cation exchange capacity have been shown to be less skewed in their distribution, and often are normally distributed about their mean. The raw data from these and similar studies has been used in a number of conceptual exercises to (1) simulate water flux during steady drainage (Warrick et al. 1977), (2) assess the relative impact of D and V on variability of solute leaching (Amoozegar-Ford et al. 1982), (3) develop stochastic models that represent the relationships between sample populations of spatially variable transport properties and the calculated water and solute movement in the field (Dagan and Bresler 1979; Bresler and Dagan 1979, 1978; Jury 1982), and (4) develop field sampling programs that consider the spatial distribution of sampling points as a necessary first step in measuring leaching variability (Warrick and Nielsen 1980).

All these studies had as a basic objective the determination of the intrinsic variation in soil properties related to water and solute movement. However, the implications of these studies may differ depending upon whether the soil is cropped or uncropped, and in the cropped case, may depend upon the leaching fraction as well. Intuitively, it seems obvious that root zone processes of water and solute extraction have an influence upon the variability of water and chemical fluxes within that region of the profile. The old adage that "a plant root takes water from where it is most available" recognizes not only variability of water in the soil, but also the interaction between the variability and water extraction. A field study that attempted to simulate water content and solute concentration in both cropped (corn) and non-cropped soil profiles (Dudley et al. 1981) provided indication that plant root extraction homogenizes leaching processes within that region. Tensiometers and neutron probe data were used to develop in-situ soil hydraulic properties, and soil solution extractors were used to sample the soil profile within and below the root zone in several treatments that differed in water and salt management. A numerical model of water and salt movement (including description of chemical reactions) provided much more accurate simulation of total salt in cropped than in non-cropped cases. A single measured, spatially averaged relationship between K-θ-matric potential was used in simulating all situations. It was hypothesized that this relationship produced simulations that more closely matched observed data in the cropped case due to a uniformity of water and solute fluxes within the root zone produced by plant extraction process. The removal of water more readily from pores where the matric potential was closest to zero, which are also pores that would be physically larger and more conducive to rapid water and solute movement (the extremes of leaching variability), reduced the influence these large pores could have upon the leaching process. The use of an average K-θ-matric potential relationship to describe a cropped area characterized by such processes was approximately correct, and reasonable simulations were thereby produced for cropped cases. Solute concentrations in the non-cropped profiles, closely resembling the experimental design used previously by e.g. Biggar and Nielsen (1976) or Van De Pol et al. (1977), were not well-described by the predictions of the numerical model. This is consistent with the previous work that indicates K-θ-matric potential relationships to be extremely variable soil properties. It is to be expected that variable solute concentrations produced by such spatially variable water flow conditions could not be predicted using a deterministic numerical model. It should also be recognized that the leaching fraction used in this study was rather low (less than 10%), and therefore the influence of the root zone in integrating and homogenizing water and solute fluxes was probably maximized. It can be expected that as the leaching fraction increases the influence of the root zone will decrease, and the variability of water and solute leaching will approach that measured in non-cropped cases. This

will be particularly true if increased leaching fractions are produced by intermittent, excessive applications of water resulting in large volumes of drainage during relatively short times.

Similar influence of the root-zone upon variability of water and solute movement has been demonstrated by Bresler et al. (1979) and Wagenet and Rao (1982). It is clear that although soil hydraulic properties may be spatially variable by orders of magnitude, this does not always translate into orders of magnitude in variability of water and solute movement in cropped situations. Transient upward and downward movement of water and solute within the root zone accompanied by plant extraction produces much more homogeneous water and solute profiles than would be expected by considering only the natural soil variability of hydraulic properties. This is true so long as leaching fractions are minimal. It also appears that deterministic, numerical models of water and salt movement can represent this variability and provide reasonable simulations of water movement and solute leaching, so long as other mediating processes (e.g. biological conversion, chemical reactions) are well-understood.

Spatial variability of mediating processes

Whether we consider the transport, transformation and plant uptake of nitrogen and organic chemicals or the simultaneous movement, precipitation, dissolution, and adsorption/desorption of inorganic ions, it is clear that a number of processes other than convection and dispersion can influence solute leaching. All these processes are potentially spatially variable as each depends upon conditions of water content, substrate concentration, temperature, and soil structure that are also spatially variable. Information on spatially variable leaching process has to date been developed through the study of non-interactive ions, such as chloride or bromide, and at times considerable success in simulating such ions has been achieved (Smith et al. 1984). However, almost all cases of environmental or agricultural interest focus on interacting chemicals that can be degraded or transformed during leaching. The experiences gained from studies using non-interactive ions have enlightened us greatly on the variability of soil physical processes related to flow, but the implications of these studies with respect to the movement of pesticides, heavy metals, toxic organics, and inorganic salts remains to be determined. The recognition that spatial variability of leaching processes exists does not mean that this variability is the controlling factor in producing spatially variable water or solute fluxes. It could eventually be shown that spatial and/or temporal variability in, for example, a biological or chemical transformation process actually overwhelms variability of hydraulic properties in determining the distribution of a solute within the soil profile.

An example of such results has been provided with respect to measuring and modeling nitrogen transport and transformation in a field cropped with corn (Wagenet and Rao 1982). Extensive measurement was made of nitrate concentrations and water contents under several water and nitrogen application regimes. The spatial variability of the K(θ) relationship was studied in 100 locations at 7 depths (700 total estimates; Jones and Wagenet 1984), and was condensed and summarized using scaling methods (Warrick et al. 1977). A numerical model of nitrogen transport and transformation which represented nitrogen transformation by first-order kinetics, was used with this information in an attempt to simulate measured nitrate concentrations. It was found that predicted nitrate concentrations were relatively insensitive to the variable K(θ) relationship. The model correctly predicted the soil depth at which peak nitrate values would exist, yet grossly overpredicted the nitrate concentration. Lack of understanding of the rate of biological conversion was hypothesized as the limiting information to successful simulation of the nitrate distribution. In this case, spatial variability of the K(θ) relationship was treated well, but the influence of variable biological conversion processes was not well enough understood. As more field observations are made of spatially variable biologically active solutes, it will become increasingly important to understand the variability of transformation or retention processes. Only in this way can solute leaching be understood on a field scale, and managed within the context of the entire soil system.

Summary

Although we now appreciate that leaching processes are spatially variable, we have not yet developed a suitable technology for accurately measuring that variability, for interpreting data from spatially variable field sampling programs, or for predicting the distribution of water and solutes in spatially variable soil profiles. However, we are closer to accomplishing these tasks than we were ten years ago. Our ability to progress to further understanding and improved management will depend upon how well we have learned our lessons from past efforts. For example, we now recognize that the manner in which we collect samples can influence the degree of variability that we think are measuring. This should motivate us to determine appropriate sampling methodologies that focus upon the REV of a property and the sample size necessary to accurately measure variability. We also recognize that statistical techniques which consider the spatial dependence of field samples can often be helpful in quantifying variability, although we do not yet include such techniques in routine assessment of water and chemical movement. In fact, routine sampling of field soils employs very little of the rapidly growing literature on soil sampling in spatially variable soils, due to a combination of logistics, poor communication, and lack of generalizable design. In response, we must redouble our efforts to design sampling

programs that maximize the scientific return from every field study. We also do not yet appreciate the full significance of a growing crop's effect on the variability of the quantity and quality of water exiting the root zone and how this variability is related to water and amendment application schemes. Finally, when the objective is prediction, we can formulate stochastic and deterministic models that describe the movement of non-interacting ions in spatially variable field soils, but we are not yet capable of treating interacting or transforming solutes that are subject to the effect of multiple, mediating spatially variable processes.

The summary effect of these facts is that although we are closer than ever before to understanding and managing spatially variable soils, we are still far from devising a comprehensive strategy for doing so. The need for such a strategy is obvious, particularly if we consider important environmental problems that focus on the movement of water and solutes. One example of such an issue is groundwater quality. In the next decade, groundwater degradation from the leaching of pesticides, fertilizers and toxic wastes will become an increasingly sensitive public issue, and it will be important to manage the soil system to reduce chemical leaching. This management must be based not upon the mean values of solute concentration within the profile, but with concern for the occasional extreme values that evolve from spatially variable soil processes and properties. It is these statistical outlyers that in many instances result in unacceptable groundwater contamination. Understanding spatial variability of solute and water movement is a fundamental first step in devising a strategy for the proper management of such systems.

References

Amoozegar-Fard, A., D.R. Nielsen and A.W. Warrick. 1982. Soil solute concentration distributions for spatially varying pore water velocities and apparent diffusion coefficients. Soil Sci. Soc. Amer. J. 46:3-9.

Anderson, J.L. and J. Bouma. 1973. Relationships between saturated hydraulic conductivity and morphometric data of an argillic horizon. Soil Sci. Soc. Amer. Proc. 37:408-412.

Baker, F.G. and J. Bouma. 1976. Variability of hydraulic conductivity in two subsurface horizons of two silt loam soils. Soil Sci. Soc. Amer. J. 40:219-222.

Bear, J. 1979. Hydraulics of Groundwater. McGraw-Hill, NY. 567 pp.

Biggar, J.W. and D.R. Nielsen. 1976. Spatial variability of the leaching characteristics of a field soil. Water Resour. Res. 12:78-84.

Bouma, J. 1977. Soil survey and the study of water in unsaturated soil. Soil Survey Papers No. 13. Netherlands Soil Survey Institute, Wageningen. 107 pp.

Bouma, J. 1983. Use of soil survey data to select measurement techniques for hydraulic conductivity. Agric. Water Mgt. 6:177-190.

Bresler, E., H. Bielorai and A. Laufer. 1979. Field test of solution flow models in a heterogeneous irrigated cropped soil. Water Resour. Res. 15:645-652.

Bresler, E. and G. Dagan. 1979. Solute dispersion in unsaturated heterogeneous soil at field scale: II. Applications. Soil Sci. Soc. Amer. J. 43:467-472.

Bresler, E. and G. Dagan. 1981. Convective and pore scale dispersive solute transport in unsaturated heterogeneous fields. Water Resour. Res. 17:1683-1693.

Cavallo, H.O., D.K. Cassel, J. Hammond and A. Bauer. 1976. Spatial variability of in-situ unsaturated hydraulic conductivity of Maddock sandy loam. Soil Sci. 121:1-8.

Dagan, G. and E. Bresler. 1979. Solute dispersion in unsaturated heterogeneous soil at field scale: I. Theory. Soil Sci. Soc. Am. J. 43:461-467.

Dudley, L.M., R.J. Wagenet and J.J. Jurinak. 1981. Description of soil chemistry during transient solute transport. Water Resour. Res. 17:1498-1504.

Fryer, J.D. and K. Kirkland. 1970. Field experiments to investigate long-term effects of repeated applications of MCPA, triallate, simazine and linuron: report after 6 years. Weed Res. 10:133-158.

Gelhar, L.W., P.J. Wierenga, K.R., Rehfeldt, C.J., Duffy, M.J., Simonett, T.-C. Yeh, and W.R. Strong. 1983. Irrigation return flow water quality monitoring, modeling and variability in the middle Rio Grande Valley, New Mexico. EPA-600/2-83-072. USEPA/ORD, Ada, OK.

Harris, C.I., E.A. Woolson and B.E. Hummer. 1969. Dissipation of herbicides at three soil depths. Weed Sci. 17:27-31.

Hassan, H.M., A.W. Warrick and A. Amoozegar-Fard. 1983. Sampling volume effects on determining salt in a soil profile. Soil Sci. Soc. Amer. J. 47:1265-1267.

Hawley, M.E., A.H. McCuen, and T.J. Jackson. 1982. Volume - accuracy relationship in soil moisture sampling. J. Irrig. and Drainage Div., Proc. ASCE., 108:1-11.

Hornsby, A.G., J.M. Davidson, D.K. Cassel and A.A. Bruce. 1983. Regional field study of the spatial variability of selected soil physical properties. Agr. Water Mgt. 6:269-276.

Hornsby, A.G., P.S.C. Rao, P. Nkeddi-Krizza, W.B. Wheeler, and A.B. Jones. 1984. Fate of aldicarb in Florida citrus soils: I. Field and laboratory studies. In: D.M. Nielsen (ed) Proc. Conf. "Characterization and Monitoring of the Vadose (unsaturated) Zone" sponsored by the National Well Water Association and US/EPA, Las Vegas, NV. Dec. 8-10.

Jones, A.J. and R.J. Wagenet. 1984. In-situ estimation of hydraulic conductivity using simplified methods. Water Resour. Res. 11:1620-1626.

Jury, W.A. 1982. Simulation of solute transport using a transfer function model. Water Resour. Res. 18:363-368.

King, J.J. and D.P. Franzmeier. 1981. Estimation of saturated hydraulic conductivity from soil morphology and genetic information. Soil Sci. Soc. Amer. J. 45:1153-1156.

Lauren, J.G. 1985. Field and laboratory characterization of hydraulic parameters in a heavy clay soil (M.S. thesis in progress, Department of Agronomy, Cornell University).

Luxmoore, R.J. 1982. Physical characteristics of soils of the Southern region: Fullerton and Sequoia series. Env. Sci. Div. Publ. No. 1890 (ORNL-5868), Oak Ridge National Laboratory, Oak Ridge, TN. 82 pp.

McKeague, J.A., C. Wang and G.C. Topp. 1982. Estimating saturated hydraulic conductivity from soil morphology. Soil Sci. Soc. Amer. J. 46:1239-1244.

Russo, D. and E. Bresler. 1982. Soil hydraulic properties as stochastic processes: II. Errors of estimates in a heterogeneous field. Soil Sci. Soc. Amer. J. 46:20-26.

Smith, S.J., L.R. Ahuja and J.D. Ross. 1984. Leaching of a soluble chemical under field crop conditions. Soil Sci. Soc. Amer. J. 48:252-258.

Taylor, A.W., H.P. Freeman and W.M. Edwards. 1971. Sample variability and the measurement of dieldrin content of a soil in the field. Agr. and Food Chem. 19:832-836.

Vachaud, G., A.P. DeSilans, P. Balabanis and M. Vauclin. 1984. Temporal stability of spatially measured soil water probability density function. (Submitted to Soil Sci. Soc. Amer. J.).

Van DePol, R.M., P.J. Wierenga and D.R. Nielsen. 1977. Solute movement in a field soil. Soil Sci. Soc. Amer. J. 41:10-13.

Vauclin, M., J. Imbernon, G. Vachaud, C. Dancette. 1983. Experimental description and stochastic modeling of transfers using a scaling factor for the hydrodynamic properties of soils (in French). pp. 103-124. In: Isotope and Radiation Techniques in Soil Physics and Irrigation Studies 1983. International Atomic Energy Agency, Vienna.

Wagenet, R.J. and B.K. Rao. 1983. Description of nitrogen movement in the presence of spatially variable soil hydraulic properties. Agric. Water Mgt. 6:227-242.

Walker, A. and P.A. Brown. 1983. Spatial variability in herbicide degradation rates and residues in soil. Crop Protection. 2:17-25.

Warrick, A.W. and D.R. Nielsen. 1980. Spatial variability of soil physical properties in the field. Chapter 13. In Applications of Soil Physics (D. Hillel, ed.). Academic Press, N.Y.

Warrick, A.W., G.J. Mullen and D.R. Nielsen. 1977. Predictions of the soil-water flux based upon field measured soil-water properties. Soil Sci. Soc. Amer. J. 41:14-19.

Warrick, A.W., G.J. Mullen and D.R. Nielsen. 1977. Scaling field-measured soil hydraulic properties using a similar media concept. Water Resour. Res. 13:355-362.

Wauchope, R.D., J.M. Chandler and K.E. Savage. 1977. Soil sample variation and herbicide incorporation uniformity. Weed Sci. 25:193-196.

Discussion

L. Wilding: The question of intrinsic vs. extrinsic variability. We must not forget that the biochemical system of soils, which is an intrinsic property, controls many of the physical-chemical reactions in soils such as biodegradation of pesticides, redox potential, dentrification, rhizosphere effects on phosphorus and symbiotic microorganism impacts.

R. Wagenet: I agree. My purpose in distinguishing between extrinsic and intrinsic variability is to recognize that we impose certain variations in soil distribution of water and solutes by our human activities (extrinsic variation), and there is another variation (intrinsic) that is produced quite naturally by biological, physical and chemical processes.

W. Schuh: When speaking of elementary sample volumes - how much can we expect the definition of our population to affect the required volume? Does this point to a need for examination of optimal classification units - not necessary but possibly spatial - for the desired parameter?

R. Wagenet: There are many possible scales of observation that can be used to characterize variability, from the smallest local volume to field scale and on to watershed or larger scales. It does not seem logical to me to ever use a smaller scale than one which incorporates in an average macroscopic manner all the possible microscopic variations. This is my concept of a representative elementary volume (REV). The population defined by it consists of measurements on soil volumes that vary by the property being measured. I believe it will be useful on any scale and will vary according to the measurement techniques and purpose. However, we should be careful that the density of variation of the factors that affect the property is the same in both the sample volume and the landscape. The minimum sample volume that meets this criteria is then both "representative", and "elementary".

L. Wilding: Doesn't the representative elementary volume depend on the property under concern (i.e. the adsorbed versus non-adsorbed elements)? For example, do we wish to look at the soil as the plant root sees it? If so we may be examining

too large a volume of soil rather than too small (consider Eh for example). On the other hand, in hydrological work perhaps our sampling volume is too small.

R. Wagenet: The point is well taken, and the answer is yes, the REV depends on the measurement being made. My answer here is similar to that given to Mr. Schuh, and can be quickly restated as "The REV, to be both representative and elementary, must contain variation in the factors affecting the property at the same density as is present in the natural system under study." In this way the REV will vary according to the property being measured.

J. Parker: We should be wary of attempts to define a single unique REV. The size of an "REV" is contingent on the scale of the continuum we seek to model. We may describe flow at the pore scale (in concept) by solution of the Navier Stokes equation given detailed knowledge of the pore structure and a big enough computer! Or as Lynn Gebhar's work elegantly shows we may be able to define a REV on the order of many meters or kilometers when, for example, an asymptotic microdispersion tensor may be obtained. Of course, as the REV size increase the variance between samples diminished while more variability is incorporated within the REV.

R. Wagenet: I agree. My reply to Mr. Schuh and Mr. Wilding amplifies on my agreement.

K. Flach: Many measurements pedologists make routinely go back to a list prepared by Richard Bradfield in 1935. Feedback from modelers is needed to update the list of measurements that are being made and to improve the methods of sampling and measurement.

R. Wagenet: I hope the types of studies underway, in which we would like to relate hydraulic properties to soil morphology, will motivate the pedologists interested in soil structure to develop a quantitative expression of soil structure. Descriptive, somewhat qualitative terms like "blocky" and "prismatic" do not help us much as we move the more quantitative statistical and mathematical treatments of soil variability and leaching. This is my first "feedback" to my fellow pedologists.

C. Topp: The classically defined soil structure concepts have not been defined for interpretations of flow and transport properties such as K_{sat}. Therefore it is not surprising that present concepts of soil morphology do not yield direct information on K_{sat}. The soil morphologist has a highly refined concept of observation which can be "calibrated" and will result in a revised approach to describing and quantifying soil morphology. Our experience of using soil morphology to predict K_{sat} indicates that a revised approach can be developed to give reliable and quantified K_{sat}. This requires that a measurement technique must be used t "calibrate" the morphological analysis. The first evidence of this is available from McKeague et al. (1982) in SSSAJ.

R. Wagenet: I am aware of that work, and while it is focused on the general direction that I suggest, I would hope that we can eventually go even further toward defining relationships between morphology and flow properties in both quantitative and mechanistic terms.

J. Starr: Others are also concerned about the effect(s) of sample volume on variance, etc., in that at least two papers were presented earlier this week by myself and my colleagues on this subject.

Did you imply that an REV can be identified for a given parameter by its resulting in a spatial structure when smaller samples do not.

R. Wagenet: It appears from our studies that using sampling volumes less than REV precludes the measurement of any spatial dependent in hydraulic conductivity, while using samples of at least the size of a REV allows us to calculate spatial dependency in the samples. However, we define the REV using elementary units of soil structure, not in terms of the volume required for estimation of spatial structure. If a relationship can be established between the REV based on soil structure and the ability to determine spatial dependence it will help tremendously. As yet, I do think this relationship has been established.

M. Nash: If you find a different range for each variable, how can you design the soil boundary for all the variables?

If you found through principal component analysis that the first principal component held the highest variation, can you use that in kriging for mapping design?

R. Wagenet: If each soil property is separately analyzed by geostatistical methods, it is probable that a different range of spatial dependence will be found for each property. Condensing these into one collective region by defining a soil boundary that includes at once all ranges is of course not possible. However, we should consider developing separate maps of each property rather than seeking condensation of several properties. After all, the soil survey already accomplishes the collective description of soil variability. There is no reason to retreat to that position once we have taken the time to study the variability of each property.

I am not familiar with the details of principal component analysis, but those knowledgeable on the subject indicate that it could be used in conjunction with kiriging.

R. Hill: How would you suggest determining a relative elementary volume of an experimental unit and determining the field location of that experimental unit so that the effects of a treatment on physical properties may be evaluated?

R. Wagenet: If working on the scale of an experimental field plot or relatively small agricultural field, I would proceed much as described in my paper relative to in-situ measurement of hydraulic conductivity. If you are examining treatment effects on physical properties (e.g., the effect of additions of organic matter upon hydraulic conductivity), I would first determine the REV for hydraulic conductivity, and would then initiate a second study that imposed organic matter treatments in a serial manner over space, with measurement of K( ) made using the right REV. This assumes that the size of the REV is not affected by the treatment, and unless the treatment is quite extreme relative to the microscopic variability, this seems a reasonable assumption.

F. Morkoc: In order to determine the REV in your infiltration study, you have done extensive sampling.

Would you speculate on how you can determine it easier?

R. Wagenet: We are hoping to relate the REV to soil morphology so that we can use the soil survey as a guide to appropriate measurement volumes. However, it may also be true that the REV is a function of measurement method. I hope it is not. We need a few more studies on such issues.

E. Flaig: My comment considers the topic of extrinsic and intrinsic soil variability. Our evidence from steady rate irrigation with fertilization of nitrate illustrates that the spatial pattern of nitrate in soil solution at 0.5 m depth was not found to reflect the spatial pattern of the depth of irrigation water or applied nitrate at the surface. However, the pattern of nitrate leaching was influenced by the location of a localized traffic pan.

R. Wagenet: It is interesting that the presence of the traffic pan apparently overwhelms much of the other dynamics of the system. Certainly such dominating physical features will affect our assessment of variability, both in intrinsic and extrinsic cases.

List of participants

Austria

C. Kirda. Wagramerstrasse 5, P.O. Box 100, A-1400 Vienna

Belgium

F. J. H. De Smedt. Dept. of Hydrology, University of Brussels, Pleinlaan 2, Brussels 1050

R. J. Tavenier. State University of Gent, Geo Institute, Krijgslaan 28, B-9000 Gent

Canada

S. Abboud. Terrain Science Dept., Alberta Research Council, Admonton AB T6H5R7

T. L. Chow. Res. Station Agric. Canada, P.O. Box 20280, Fredericton NB E3B4Z7

D.E. Elrick. Dept. of Land Res. Science, University of Guelph, Guelph ON N1G2W1

J. A. Fischer. University of Guelph, Guelph ON N1G2W1

P. H. Groenevelt. RR No.5, Rockwood ON NOB2KO

P. M. Huang. Dept. of Soil Science, University of Saskatchewan, Saskatoon SK S7N0W0

R.G. Kachanoski. Dept. of Soil Science, University of Saskatchewan, Saskatoon SK S7N0W0

R. Protz. Dept. of Land Resource Science, University of Guelph, Guelph, ON N1G2W1

J. E. Smith. Dept. of Land Resource Science, University of Guelph, Guelph ON N1G2W1

L. Smith. Dept. of Geo. Sciences, University of British Colombia, Vancouver BC 6VT2B4

G. C. Topp. Central Experimental Farm, Land Resource Res. INstitute, Ottawa ON K1A0C6

C. Wang. Central Experimental Farm, Land Resource Res. Institute, Ottawa ON K1A0C6

Chile

L. Gurovich. Univ. Catolica de Chile, P.O. Box 114-D, Santiago

Federal Republic of Germany

G. M. Richter. Bodentechnologisches Institut, Friedrich-Missler-Strasse 46-50, Bremen

France

J. Bottraud. 262 rue des Marguerites, St. Geldu Fesc 34980
A. Ruellan. ORSTOM, 24 rue Bayard, Paris

Israel

E. Bresler. Institute of Soils & Water, Volcani Center, P.O. Box 6, 50-250 Bet Dagan

Japan

M. Kiuchi. 2-46-10 Akasutsumi, Setagaya-ku, Tokyo 156

Netherlands

H. F. M. ten Berge. Dept. of Soils and Fertilizers, Agricultural University, De Dreyen 3, 6703 BC Wageningen
J. Bouma. Soil Survey Institute, P.O. Box 98, 6700 AB Wageningen
J. J. De Gruijter. P.O. Box 297, 2501 CG The Hague
B. Overmars. Dept. of Hydraulics, Agricultural University. Nieuwe Kanaal 11, 6709 PA Wageningen
L. Stroosnijder. Agricultural University, De Dreyen 3, 6703 BC Wageningen

New Zealand

B. E. Clothier. DSIR Plant Physiology Division, Palmerston North

Nigeria

R. Lal. IITA. PMB 5320, Ibadan

Philippines

K. A. Gomez. IRRI, P.O. Box 933, Manila

Switzerland

B. Buchter. Zurlindenstrasse 214, 8003 Zürich
P. J. Greminger. EAFV, Zurcher Strasse 111, 8903 Birmensdorf

Taiwan R.O.C.

C. Lin. Dept. of Soil Science, National Chung Hsing University, 400 Taichung
M. S. Lin. Taiwan Agricultural Research Institute, Wu-feng,Taichung

Thailand

S. Akratanakul. Dept. of Soil Science, Kasetsart University, Kamphaeng Saen, Nakorn Pathom 73140

United States of America

P. O. Aina, Land, Air & Water Resources, University of California, Davis, CA 95616
E. E. Alberts. 3005 Lynnwood Dr., Columbia, MO 65201
J. C. Allen. Dept. of Geography, University of California, Santa Barbara, CA 93106
R. R. Allmaras. Columbia Plateau Cons. Res. Center, P.O.Box 370, Pendleton, OR 97801
D. L. Anderson. IFAS/ARES, University of Florida, Belle Glade, FL 33430
W. P. Ball. Dept. of Civil Engineering, Stanford University, Stanford, CA 94305
R. F. Bauer. 32 Corthell Rd., Laramie, WY 82070
J. W. Biggar. Land, Air & Water Resources, University of California Davis, CA 95616
C. W. Boast. Agronomy Dept., University of Illinois, 1102 South Goodwin Ave., Urbana, IL 61801
L. L. Boersma. Dept. of Soil Science, Oregon State University, Corvallis, OR 97331
D. B. Bottcher. Agricultural Engineering, Rogers Hall, University of Florida, Gainesville, FL 32611
R. S. Bowman. 4331 East Broadway Rd., U.S. Water Conservation Laboratory, Phoenix, AZ 85040
D. G. Boyer. P.O. Box 867, Airport Rd., Beckley, WV 25801
R. B. Brown. Soil Science Dept. G-159 McCarty, University of Florida, Gainesville, FL 32611
R. R. Bruce, Box 555, So. Piedmont Cons. Res. Center, Watkinsville, GA 30677
D. S. Burden. 1836 Pickett Ave., Baton Rouge, LA 70808
A. E. Carey, U.S. Environ. Protection Agency, TS-768-C, OPP/BUD/SSB Washington, DC 20460
D. K. Cassel. Dept. of Soil Science, North Carolina State University, Raleigh, NC 27650
R. G. Chase. Trop. Soils-Dan-1311, Dept. of State, Niamey ID, Washington, DC 20520
S. W. Childs. Dept. of Soil Science, Oregon State University, Corvallis, OR 97331
S. Chong. Dept. of Forestry, Southern Illinois University, Carbondale, IL 62901
C. E. Clapp. 329 Soil Science Building, University of Minnesota, USDA-ARS, St. Paul, MN 55108
M. Clark. Soil Conservation Service, Box 3164, Palmer, AK 99645
D. E. Clay. 2125 W. Highway 36, St. Paul, MN 55113
T. L. Coleman, Box 183, Normal, AL 35762
M. E. Collins. Soil Science Dept., G-159 McCarty Hall, University of Florida, Gainesville, FL 32611

N. B. Comerford. Soil Science Dept., G-159 McCarty Hall, University of Florida, Gainesville, FL 32611
J. L. B. Culley. Dept. of Soil Science, 1529 Gortner Ave., St. Paul, MN 55108
J.H. Cushman. Agronomy Dept., Purdue University, West Lafayette, IN 47907
J. H. Dane. Dept. of Agronomy & Soils, Auburn University, Auburn, AL 36849
L. A. Daugherty. Crop & Soil Science Dept., Box 30, New Mexico State University, Las Cruces, NM 88003
B. Davidoff. Agronomy Dept., Louisiana State University, Baton Rouge, LA 70803
J. M. Davidson. IFAS, 1022 McCarty Hall, University of Florida, FL 32611
S. J. Deverel. P.O. Box 273, Davis, CA 95616
J. A. Doolittle. SCS, Federal Building, 401 SE 1st Ave., Room 248, Gainesville, FL 32601
R. H. Dowdy. Soil Science Dept., 1529 Gortner Ave., University of Minnesota, St. Paul, MN 55108
C. J. Duffy. Utah Water Res. Laboratory, UMC82, Utah State University, Logan, UT 84322
M. G. Durant. Dept. of Civil Engineering, Stanford University, Stanford, CA 94305
P. B. Durgin. 4608 Dows Prairie Rd., McKinleyville, CA 95521
H. S. Elabd. Dept. of Soil & Environm. Science, University of California, Riverside, CA 92521
W. E. Emmerich. 2000 E. Allen Rd., Tucson, AZ 85719
C, G, Enfield. RSKERL, P.O.Box 1198, Ada, OK 74820
M. S. Erich. Box 42, Bradfield Hall, Cornell University, Ithaca, NY 14853
H. Eswaran. P.O. Box 2890, USDA Soil Conservation Service, Washington, DC 20013
M. J. Fayer. 1713 Gray St., Richland, WA 99352
C. M. Feldhake. USDA, P.O.Box 867, Beckley, WV 25801
A. H. Ferguson. Dept. of Plant & Soil Science, Montana State University, Bozeman, MT 59717
S. E. Fisher Jr. 9270 W. 9th Ave. Lakewood, CO 80215
J. G. A. Fiskell. G-183 McCarty Hall, University of Florida, Gainesville, FL 32611
K. W. Flach. USDA, Soil Conservation Service, South Building, P.O. Box 2890, Washington, DC 20013
E. G. Flaig. 20-3 N.W. 8th St., No. 2, Gainesville, FL 32601
G. T. Flatman. U.S. EPA EMSL-LV, 3583 Haverford Ave., Las Vegas, NV 89121
R. Fujii. 400 Coloma Way, Sacramento, CA 95819
C. J. Gantzer. Dept. of Agronomy, 143 C Mumford, University of Missouri, Columbia, MO 65211
H. G. Gauch Jr. Dept. of Agronomy, Cornell University, Ithaca, NY 14853
G. W. Gee. Geoscience Res. & Engr. Dept., Batelle N.W. Labs. P.O. Box 999, Richland, WA 99352
L. W. Gelhar. M.I.T., 48-331, Cambridge, MA 02139

P. F. Germann. Dept. of Environm. Sciences, Clark Hall, University of Virginia, Charlottesville, VA 22903
A. Goldin. 2212 Dean Ave. Bellingham, WA 98225
M. N. Goltz. Dept. of Civil Engineering, Stanford University, Stanford, CA 94305
D. W. Goss. SCS Federal Building, Room 393, Lincoln, NE 68508
R. E. Green. Agronomy & Soils, 3190 Maile Way, University of Hawaii, Honolulu, HI 96822
R. B. Grossman. Fed. Office Bldg. P.O.Box 82502, SCS Nat. Soil Survey Laboratory, Lincoln, NE 68508
W. D. Guenzi. USDA-ARS, P.O. Box E, Fort Collins, CO 80522
A. Gutjahr. Dept. of Mathematics, New Mexico Tech. Socorro, NM 87801
G. A. Halvorson. P.O. Box 459, North Dakota State University, Mandan, ND 58554
S. P. Hamburg. Dept. of Biological Sciences, Stanford University, Stanford, CA 94305
R. D. Hammer. Dept. of Soil Science, Box 1071, University of Tennessee, Knoxville, TN 37916
L. C. Hammond. 2169 McCarty Hall, University of Florida, Gainesville, FL 32611
C. A. Harrington. Southern Forest Expt. Sta., Box 3516 UAM, Monticello, AR 71655
A. R. Harris. North Central Forest Exp. Sta., Forestry Science Lab., 1831 Highway 169E, Grand Rapids, MN 55744
H. H. Hatzell. Division of Agriculture, Arizona State University, Tempe, AZ 85287
J. M. Hendrickx, Dept. of Agronomy, Box 3Q, New Mexico State University, Las Cruces, NM 88003
W. N. Herkelrath. U.S. Geological Survey, 345 Middlefield Rd., Menlo Park, CA 94025
R. C. Herriman. 931 Lobster Trap Lane, Tempe, AZ 85283
R. L. Hill. Dept. of Agronomy, University of Maryland, College Park, MD 20742
D. I. Hillel. Dept. of Plant & Soil Sciences. University of Massachusetts, Amherst, MA 01003
K. C. Hinkley. 4097 Braxton Rd., Chantilly, VA 22021
G. Holmgren. 4240 Normal Blvd., Lincoln, NE 68506
J. E. Hook. P.O.Box 748, Tifton, GA 31794
A. G. Hornsby. Soil Science Dept. 2169 McCarty Hall, University of Florida, Gainesville, FL 32601
R. Horton Jr. Agronomy Dept., Iowa State University, Ames, IA 50011
E. H. Horvath. Applications Branch, EROS Data Center, Sioux Falls, SD 57198
Y. P. Hsieh. P.O. 239, Florida A & M University, Tallahassee, FL 32307
G. L. Hutchinson. P.O. Box E, Fort Collins, CO 80522
B. A. Ibrahim. Dept. of Agronomy and Soils, 305 Johnson, Washington State University, Pullman, WA 99164
J. D. Istok. Dept. of Soil Science, Oregon State University, Corvallis, OR 97331
D. B. Jaynes. U.S. Water Cons. Lab. 4331 East Broadway Rd., Phoenix, AZ 85040
J. A. Johnson. 860 Spruce St., Riverside, CA 92507

A. J. Jones. Western Triangle Research Center. P.O. Box 1474, Conrad, MT 59425
W. A. Jury. Dept. of Soil & Environ. Science, University of California, Riverside, CA 92521
S. R. Kaffka. Dept. of Agronomy, Bradfield, Cornell University, Ithaca, NY 14853
W. M. Klittich. 201 Agronomy Dept. Iowa State University, Ames, IA 50011
G. J. Kluitenberg. 4856 N. Glenn Ave., Apt. E., Fresno, CA 93704
R. E. Knighton, 301 Maple Ave., Apt. H-2, Ithaca, NY 14850
J. B. Kool. Agronomy Dept., VPI & SU, Blacksburg, VA 24061
R. J. Kuehl. 307 Abraham, Ames, IA 50010
D. A. Lammers. 1062 Gates Dr. (61-1), Roosevelt, UT 84066
K. J. Lang. Western Energy Company, Box 67, Colstrip, MT 59323
L. A. Lanyon. 102 Cherry Ridge Rd., State College, PA 16803
R. Lascano. Univ. Experiment Station, Texas A&M, Rt. 3, Lubbock, TX 79401
G. A. Lehrsch. Dept. of Agronomy, P.O. Box 5248, Mississippi State University, Mississippi State, MS 39762
D. R. Linden. Soil Science Building, University of Minnesota, St. Paul, MN 55108
L. E. Loomis. Dept. of Range Science, Texas A&M University, College Station, TX 77843
R. J. Luxmoore. P.O. Box X, Oak Ridge National Laboratory, Oak Ridge, TN 37830
J.L. MacIntyre. 1910 Lehigh Dr., Davis, CA 95616
R. S. Mansell. Soil Science Dept., G-149 McCarty Hall, University of Florida, Gainesville, FL 32611
M. J. Mausbach. Federal Building, Room 345, 100 Centennial Mall North, Lincoln, NE 68508
P. K. McConnaughey. Dept. of Agronomy, Box 5248, Mississippi State University, Mississippi State, MS 39762
D. E. McCormack. 3160 Holmes Run Road, Falls Church, VA 22042
D. H. McNabb. 1301 Maple Grove Dr., Forestry Intensified Res., Medford, OR 97501
R. J. Miles. Agronomy Dept., 138 Mumford Hall, University of Missouri, Columbia, MO 65211
M. P. Miller. Land, Air & Water Resources, University of California, Davis, CA 95616
A. R. Mosier. P.O.Box E, Fort Collins, CO 80522
D. J. Mulla. Dept. of Agronomy & Soils, Washington State University Pullman, WA 99164
I. P. Murarka. 3412 Hillview Ave., Electric Power Res. Institute, Palo Alto, CA 94304
D. E. Meyers. Dept. of Mathematics, University of Arizona, Tucson, AZ 85721
D. L. Myhre. 2280 N.W. 21st Place, Gainesville, FL 32605
D. D. Myrold. Dept. of Soil Science, Oregon State University, Corvallis, OR 97331
M. S. Nash. New Mexico State University, P.O.Box 3Q, Las Cruces, NM 88003
M. H. Nash. Agronomy Dept., New Mexico State University, P.O. Box 3Q, Las Cruces, NM 88003

R. W. Nelson. Battelle Northwest, PSL Building, P.O. Box 999, Richland, WA 99352

D. R. Nielsen. Land, Air & Water Resources, University of California, Davis, CA 95616

G. A. Nielsen. Plant & Soil Science Dept., Montana State University Bozeman, MT 59717

P. Nkedi-Kizza. Soil Science Dept., McCarty Hall, University of Florida, Gainesville, FL 32611

D. L. Nofziger. Dept. of Agronomy, Oklahoma State University, Stillwater, OK 74074

T. Ohno. U.S. Plant, Soil & Nutr. Lab., Tower Rd., Ithaca, NY 14853

J. C. Parker. 334 B Smyth Hall, Agronomy Dept., VPI & SU, Blacksburg VA 24061

T. B. Parkin. Building 007, Room 249, Beltsville Agr. Res. Center, Beltsville, MD 20705

C. L. Ping. Agricultural Exper. Station, 533 E. Fireweed, Palmer, AK 99645

G. J. Post. 911 Driftwood Dr., Lincoln, NE 68510

D. E. Radcliffe. 325 1/2 Oakland Ave., Athens, GA 30606

P. S. Rao. Soil Science Dept., 2169 McCarty Hall, University of Florida, IFAS, Gainesville, FL 32611

J. A. Rehage. 1500 Hawk Tree, College Station, TX 77840

R. C. Rice. 4331 E. Broadway Rd., Phoenix, AZ 85040

J. A. Robinson. The Upjohn Co., Gull Rd. Kalamazoo, MI 49001

D. E. Rolston. Land, Air & Water Resources, University of California, Davis, CA 95616

E. C. A. Runge. Soil & Crop Science Dept., Texas A&M University, College Station, TX 77843

M. P. Russelle. Soil Science Dept., University of Minnesota, St. Paul, MN 55108

N. M. Safaya. Reclamation Division, State Capitol Building, Bismarck, MD 58505

W. H. Schuh. 5912 Horseshoe Bend, Mandan, ND 58554

H. M. Selim. Agronomy Dept., Louisiana State University, Baton Rouge, LA 70803

A. J. Sexstone. Plant & Soil Science, Box 6057, West Virginia University, Morgantown, WV 26506

M. J. Shaffer. USDA-SEA-AR, 313 Soil Science Building, University of Minnesota, St. Paul, MN 55108

P. J. Shouse. Blackland Research Center, Box 748, Temple, TX 76501

K. B. Showers. Rodale Research Center, Rt. 1, Box 323, Kutztown, PA 19530

R. Shumway. Div..of Statistics, University of California, Davis, CA 95616

M. L. Sipin. Mississippi State University, P.O. Box 2519, Mississippi State, MS 39762

J. B. Sisson. Agronomy Dept., Throckmorton, Kansas State University Manhattan, KS 66506

E. L. Skidmore. USDA-ARS, Waters Hall, Room 148, Kansas State University, Manhattan, KS 66506

J. P. Slater. P.O.Box 3163, El Macero, CA 95618

C. M. Smith. USDA, 4605 Dolphin Lane, Alexandria, VA 22309

A. J. M. Smucker. Dept. of Crop & Soil Sciences, Michigan State University, East Lansing, MI 48824
P. N. Soltanpour. Agronomy Dept., Colorado State University, Fort Collins, CO 80523
J. L. Starr. USDA-ARS BARC-W, Building 007, Room 253, Beltsville, MD 20705
G. P. Streile. Dept. of Soil & Envir. Science, University of California, CA 92521
M. L. Thompson. Agronomy Dept., Iowa State University, Ames,IA 50011
P. M. Tillotson. 2937 Janet Dr., West Sacramento, CA 95691
R. I. Turner. 2521 Marilynn Ave., Lincoln, NE 68502
R. S. Turner. Oak Ridge National Laboratory, Environmental Science Division, P.O.Box X, Oak Ridge, TN 37830
G. Uehara. Dept. of Agronomy & Soils, University of Hawaii, Honolulu, HI 96822
M. T. van Genuchten. U.S. Salinity Laboratory, 4500 Glenwood Dr. Riverside, CA 92501
R. J. Wagenet. Dept. of Agronomy, Cornell University, Ithaca, NY 14853
D. M. Walters. U.S. EPA RSKERL, P.O. Box 1198, Ada, OK 74820
A. W. Warrick. 521 Agricultural Sciences, University of Arizona, Tucson, AZ 85721
K. W. Watson. Oak Ridge National Laboratory, Environmental Science Division, P.O.Box X, Oak Ridge, TN 37830
R. C. Wendt. USDA-ARS, 207 Business Loop 70 East, Columbia, MO 65201
F. D. Whisler. Dept. of Agronomy, P.O. Box 5248, Mississippi State MS 39762
P. J. Wierenga. Dept. of Agronomy, Box 3Q, New Mexico State University, Las Cruces, NM 88003
L. P. Wilding. Soil & Crop Sciences Dept. Texas A&M University, College Station, TX 77843
A. L. Wood. Environmental Protection Agency, P.O. Box 1198, Ada, OK 74820
S. R. Yates. 1408 E. Duke Dr. Tucson, AZ 85719
R. S. Yost. Dept. of Agronomy & Soil Science, 3190 Maile Way, Honolulu, HI 96822
R. E. Zartman. Dept. of Plant & Soil Science, Texas Tech. University, Lubbock, TX 79409
T. M. Zobeck. USDA-ARS, Wind Erosion & New Crops, P.O.Box 909, Big Spring, TX 79720