AGRO 896/ MSYM 898 Understanding and Managing Variability in Soils

Lecture 6: Quantifying spatial variability

Instructor: Achim Dobermann, Professor, Soil Science and Nutrient Management
Dept. of Agronomy and Horticulture, University of Nebraska-Lincoln
253 Keim Hall, PO Box 830915, Lincoln, NE 68583-0915, USA
Phone: +1-402-472-1501, E-mail: adobermann2@unl.edu

Objectives:

1. Understand the principles of sampling
2. Understand the use of descriptive statistics for characterizing variability in spatial data
3. Understand how geostatistics can be used to characterize the structure of spatial variation and to predict values at unvisited locations (interpolation to create maps)

Content:

1. General approach for assessing spatial variation

- Collection of prior information,
- Sampling
- Statistical analysis
- Prediction (mapping).

2. Sampling

a. Considerations for designing a sampling scheme:
   Purpose: target area, variables, accuracy required
   Constraints: financial, logistical, operational
   Sampling design: no. of samples, how selected, composite samples per location
   Method of taking samples (optimal time, size, depth, devices)
   Method of determination (field, laboratory)
   Protocols for data recording and fieldwork
   Methods of statistical analysis
b. Sampling designs used in precision agriculture:
   Adaptive sampling by strata (management zones)
   Grid sampling
   Point sampling
   Automated mapping (on-the-go)
c. Considerations for composite samples:
   Sample several single cores, put them together, mix, sub-sample for analysis
   Reduces micro-variability found among single cores
   Reduces cost
   Suitable for obtaining a representative sample for a larger block of soil, e.g., grid cell
   Less suitable for unraveling micro-variability and detecting extreme values
   Less suitable for detecting boundaries or abrupt soil changes in a field

Three sources of error: (i) a single core can inflate the value of the composite sample, (ii) imperfect mixing, and (iii) error due to sub-sampling

d. Some recommendations:

- Use other available information to decide on the most suitable sampling design (EC maps, imagery, DEM, yield maps)
- Grid sampling is often not economical for variable rate applications of fertilizers; grid sizes of > 1-2 acre tend to mask much spatial variation.
- Invest in more detailed point or on-the-go sampling, but it is not necessary to repeat this frequently. Spatial patterns of many soil properties of agronomic interest (SOM, texture, P, K, pH, CEC) are relatively consistent for periods of at least 5 years.
- In between, particularly for dynamic properties such as residual soil nitrate-N, use annual composite sampling by management zones.
- Stick to a constant sampling method: core diameter, depth, device used. Sample at the depth for which prescription algorithms have been calibrated (0-8" for pH, P, K, S, Zn; 1-3 ft for nitrate).
- Minimize compaction of soil in the core to always have a constant depth. Don't sample when it is too dry or too wet.

3. Descriptive statistics

a. Order statistics

Median

Percentiles (e.g., min., 25%, 75%, max.)

b. Measures of central tendency

Mean

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + ... + x_n}{n}$$

c. Measures of dispersion

Variance and standard deviation

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\sigma = \sqrt{\sum_{i=1}^{n}\frac{(x_i - \mu)^2}{(n-1)}}$$

Coefficient of variation

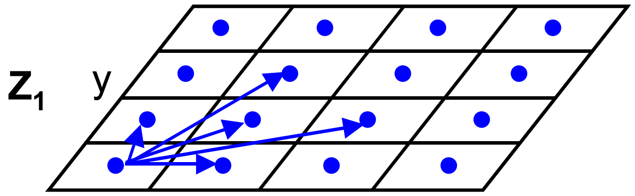
$$\mathrm{CV}\ (\%) = \frac{\sigma}{\mu}\times 100$$

Skewness

Kurtosis

4. Geostatistics

Geostatistics: mathematical procedures to recognize and describe spatial relationships (spatial correlation) that might exist in a field.



a. Exploratory data analysis

Visual investigation of patterns & descriptive statistics: postplots and histograms

b.Variography: analysis of the spatial correlation structure using semivariograms.

Structure, range, direction, strength of spatial variation; single variables or multivariate spatial processes (correlated soil & environmental information)

Let **u** be a position in 1, 2, or 3 dimensions. The value of variable *Z* at **u** is given by:

$$Z(\mathbf{u}) = \mathrm{m}(\mathbf{u}) + \varepsilon(\mathbf{u}), \text{ or}$$

$$Z(\mathbf{u}) = \mathrm{m}(\mathbf{u}) + \varepsilon'(\mathbf{u}) + \varepsilon''$$

m(**u**) – deterministic function, mean or a ‘trend’ (‘drift’) = systematic variation

ε(**u**) – stochastic component: fluctuations around the trend

ε’(**u**) – local variation, spatially correlated

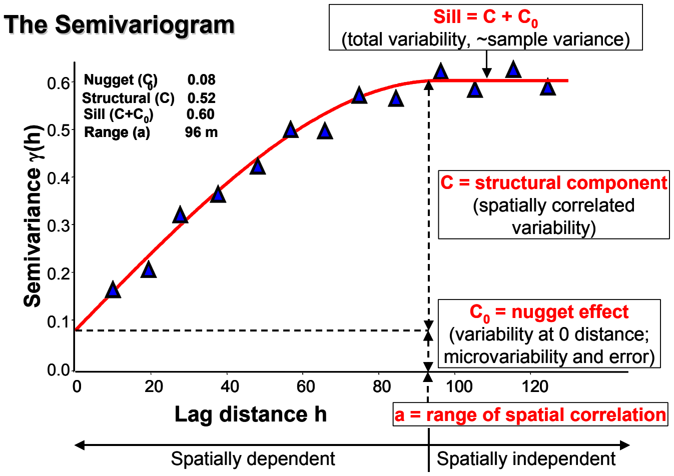
ε”(**u**) – local random variation, spatially uncorrelated (includes microvariability and measurement error)The semivariance γ(*h*) [“Gamma”] is a function describing the spatial correlated component of the variation of a soil property:

$$\gamma(h) = \frac{1}{2n}\sum_{i=1}^{n}\{Z(\mathbf{u}_i) - Z(\mathbf{u}_i + h)\}^2$$

where n is the number of pairs of sample points separated by the distance h (also called the **lag distance**).

γ(*h*) is calculated for all possible lag distance classes in a data set

A plot of γ(*h*) against the various *h* is known as the sample **semivariogram**.



The sill indicates the overall spatial variability. If the data are close to normal distribution, sill equals the variance.
Compare semivariograms of different soil properties or different fields; relate their structure (ranges) to other available information; deduct hypotheses about causes of variability.
Use the range to determine maximum sampling distances: sample at intervals shorter than the range.
Use the fitted semivariogram model for optimal interpolation of values in unsampled areas: **kriging**.

c. Prediction: creating maps by interpolation

Inverse distance weighted interpolation:

$$z^*(x_j) = \frac{\sum_{i=1}^{n} z(x_i)\, d_{ij}^{-p}}{\sum_{i=1}^{n} d_{ij}^{-p}}$$

z_i - locations of measured data points
d_{ij} - distance from each point to z_j
p – weight (smoothing) factor

Kriging: a family of generalized least-squares regression algorithms for estimating the value of a regionalized variable $Z(\mathbf{u})$ at unsampled locations $\mathbf{u_0}$.
All kriging estimates are best linear unbiased estimates (BLUE):

$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i)$ with $\sum_{i=1}^{n} \lambda_i = 1$ The weights $\lambda_\alpha(\mathbf{u})$ are chosen are chosen so that the estimate $Z^*(\mathbf{u})$ is unbiased and the estimation variance (kriging errror) is minimized, which is defined as

$\sigma_e^2 = \sum_{i=1}^{n} \lambda_i \gamma(x_i, x_0) + \psi$ where γ is the semivariance between the sampling point $\mathbf{u_i}$ and the unvisited point $\mathbf{u}_0$, and ψ is a Lagrange factor required for minimization.
Ordinary kriging vs. multivariate extensions (regression kriging)

d. Assessment, interpretation, decisions
Deduct hypotheses about causes of variability
Quantify map precision and uncertainties
Management decisions

4. Summary

- Utilize available information to carefully design the sampling scheme, depending on the purpose.
- Assess your data first by plotting the measured values on a map and computing histograms and descriptive statistics.
- If the sampling scheme allows it, use a geostatistical approach for quantifying the spatial correlation structure.
- Use map patterns and the semivariogram for identifying possible causes of spatial variability.
- Always try and compare different interpolation methods and settings for each method.
- Use strongly smoothing interpolators only for variables that show smooth trends (climate, elevation).
- Be cautious about using inverse distance interpolation, particularly for scarce data.
- Kriging provides the most flexibility and an estimate of map precision, but it requires detailed sampling to be able to model the semivariogram.
- Fine-resolution secondary information can increase the precision of maps of soil properties, if it is readily available and correlated with primary variables.

Advanced reading materials

Beckett, P.H.T., and R. Webster. 1971. Soil variability: a review. Soils Fert. 34:1-15.

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

Burgess, T.M., and R. Webster. 1980. Optimal interpolation and isarithmic mapping of soil properties. II. Block kriging. J. Soil Sci. 31:333-341.

Cahn, M.D., J.W. Hummel, and B.H. Brouer. 1994. Spatial analysis of soil fertility for site-specific crop management. Soil Sci. Soc. Am. J. 58:1240-1248.

De Gruijter, J.J. 2002. 1.4 Sampling. p. 45-79. *In* J.H. Dane and G. Clarke Topp (ed.) Methods of soil analysis. Part 4. Physical methods. Soil Science Society of America, Madison,WI.

Dobermann, A. 1993. Geostatistical methods for analysis of soil heterogeneity. p. 47-66. *In* R.J.K. Myers and C.R. Elliott (ed.) Interpreting data from field experiments on upland soils. IBSRAM Technical Notes no. 6. IBSRAM, Bangkok.

Dobermann, A., P. Goovaerts, and T. George. 1995. Sources of soil variation in an acid Ultisol of the Philippines. Geoderma 68:173-191.

Oliver, M.A., and R. Webster. 1991. How geostatistics can help you. Soil Use Manag. 7:206-218.

Goovaerts, P. 1997. Geostatistics for natural resources evaluation. Oxford University Press, New York.

Gotway, C.A., R.B. Ferguson, G.W. Hergert, and T.A. Peterson. 1996. A comparison of kriging and inverse-distance methods for mapping soil parameters. Soil Sci. Soc. Am. J. 60:1237-1247.

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

McBratney, A.B., M.L. Mendonca Santos, and B. Minasny. 2003. On digital soil mapping. Geoderma 117:3-52.

Nielsen, D.R., and O. Wendroth. 2003. Spatial and temporal statistics - sampling field, soils and their vegetation. Catena Verlag, Reiskirchen,Germany.

Trangmar, B.B., R.S. Yost, and G. Uehara. 1985. Application of geostatistics to spatial studies of soil properties. Adv. Agronomy 38:45-93.

Webster, R. 1994. The development of pedometrics. Geoderma 62:1-15.

Webster, R. 2000. Is soil variation random? Geoderma 97: 149-163.

Webster, R., and M.A. Oliver. 2000. Geostatistics for environmental scientists. Wiley, Chichester.

Wendroth, O., P. Jurschik, K.C. Kersebaum, H. Reuter, C. van Kessel, and D.R. Nielsen. 2001. Identifying, understanding, and describing spatial processes in agricultural landscapes--four case studies. Soil Tillage Res. 58:113-127.

Wollenhaupt, N.C., D.J. Mulla, and C.A. Gotway. 1997. Soil sampling and interpolation techniques for mapping spatial variability of soil properties. p. 19-53. *In* F.J. Pierce and E.J. Sadler (ed.) The state of site-specific management for agriculture. ASA-CSSA-SSSA, Madison,WI.