Designation: D 5922 – 96$^{\epsilon 1}$

AMERICAN SOCIETY FOR TESTING AND MATERIALS
100 Barr Harbor Dr., West Conshohocken, PA 19428

Standard Guide for Analysis of Spatial Variation in Geostatistical Site Investigations[1]

This standard is issued under the fixed designation D 5922; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

ϵ^1 NOTE—Paragraph 1.5 was added editorially October 1998.

INTRODUCTION

Geostatistics is a framework for data analysis, estimation, and simulation in media whose measurable attributes show erratic spatial variability yet also possess a degree of spatial continuity imparted by the natural and anthropogenic processes operating therein. The soil, rock, and contained fluids encountered in environmental or geotechnical site investigations present such features, and their sampled attributes are therefore amenable to geostatistical treatment. This guide is concerned with the analysis, interpretation, and modeling of spatial variation. The purpose of this guide is to offer guidance based on a consensus of views but not to establish a standard practice to follow in all cases.

1. Scope

1.1 This guide covers recommendations for analyzing, interpreting, and modeling spatial variation of regionalized variables in geotechnical and environmental site investigations.

1.2 The measures of spatial variation discussed in this guide include variograms and correlograms; these are fully described in **(1)**, **(2)**, **(3)**, and **(4)**.[2]

1.3 This guide is intended to assist those who are already familiar with the geostatistical tools discussed herein and does not provide introductory information on the analysis, interpretation, and modeling of spatial variation.

1.4 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

1.5 *This guide offers an organized collection of information or a series of options and does not recommend a specific course of action. This document cannot replace education or experience and should be used in conjunction with professional judgment. Not all aspects of this guide may be applicable in all circumstances. This ASTM standard is not intended to represent or replace the standard of care by which the adequacy of a given professional service must be judged, nor should this document be applied without consideration of a project's many unique aspects. The word "Standard" in the title of this document means only that the document has been approved through the ASTM consensus process.*

2. Referenced Documents

2.1 *ASTM Standards:*

D 653 Terminology Relating to Soil, Rock, and Contained Fluids[3]
D 5549 Guide for Reporting Geostatistical Site Investigations[4]
D 5923 Guide for Selection of Kriging Methods in Geostatistical Site Investigations[4]
D 5924 Guide for the Selection of Simulation Approaches in Geostatistical Site Investigations[4]

3. Terminology

3.1 *Definitions of Terms Specific to This Standard:*

3.1.1 *anisotropy, n—in geostatistics*, a property of the variogram or covariance stating that different spatial variation structures are observed in different directions.

3.1.1.1 *geometric anisotropy, n*—a form of anisotropy in which the variogram range changes with direction while the sill remains constant.

3.1.1.2 *zonal anisotropy, n*—a form of anisotropy in which the variogram sill changes with direction.

3.1.2 *correlogram, n*—a measure of spatial variation expressing the coefficient of correlation between two variables as a function of the lag separating their locations.

3.1.3 *drift, n—in geostatistics*, a systematic spatial variation of the local mean of a variable, usually expressed as a polynomial function of location coordinates.

---

[1] This guide is under the jurisdiction of ASTM Committee D-18 on Soil and Rock and is the direct responsibility of Subcommittee D18.01 on Surface and Subsurface Characterization.

Current edition approved April 10, 1996. Published June 1996.

[2] The boldface numbers in parentheses refer to a list of references at the end of the text.

[3] *Annual Book of ASTM Standards*, Vol 04.08.

[4] *Annual Book of ASTM Standards*, Vol 04.09.

3.1.4 *estimation, n*—a procedure by which the value of a variable at an unsampled location is predicted using a weighted average of sample values from the neighborhood of that location.

3.1.5 *lag, n—in geostatistics*, the vector separating the locations of two variables, as used in measures of spatial variation.

3.1.6 *nugget effect, n*—the component of spatial variance unresolved by the sample spacing including the variance due to measurement error.

3.1.7 *range, n—in geostatistics*, the maximum distance over which a variable exhibits spatial correlation in a given direction.

3.1.8 *regionalized variable, n*—a measured quantity or a numerical attribute characterizing a spatially variable phenomenon at a location in the field.

3.1.9 *sill, n—in geostatistics*, a stable level of spatial variation observed for lags greater than the range.

3.1.10 *simulation, n—in geostatistics*, a Monte-Carlo procedure for generating realizations of fields based on the random function model chosen to represent a regionalized variable. In addition to honoring a random function model, the realizations may also be constrained to honor data values observed at sampled locations.

3.1.11 *structure, n—in geostatistics*, a source of spatial variability with a characteristic length scale.

3.1.12 *variogram, n*—a measure of spatial variation defined as one half the variance of the difference between two variables and expressed as a function of the lag; it is also sometimes referred to as the semi-variogram.

3.1.12.1 *experimental variogram, n*—an experimental measure of spatial variation usually calculated as one half the average squared difference between all pairs of data values within the same lag.

3.2 For definitions of other terms used in this guide, refer to Terminology D 653 and Guides D 5549, D 5923, and D 5924. A complete glossary of geostatistical terminology is given in Ref **(5)**.

4. Summary of Guide

4.1 This guide presents advice on three separate but related components of the study of spatial variation: the analytical tools that are used; the interpretation of the results; and the development of an appropriate mathematical model.

4.2 For the analysis of spatial variation, this guide emphasizes the use of variograms and correlograms on both transformed and untransformed variables since these are the most common and successful analytical tools in most practical situations. Other methods exist and may enhance the development of an appropriate model of spatial variation.

4.3 For the interpretation of spatial variation, this guide emphasizes the importance of site-specific quantitative and qualitative information. Quantitative information includes the number and configuration of the available data, their precision, and their univariate statistics; qualitative information includes items such as local geology and geomorphology, site usage, and history. All of these are necessary for a sound interpretation of spatial variation.

4.4 For the modeling of spatial variation, this guide recommends attention to the short-scale behavior of the mathematical model of spatial variation and to its anisotropy as reflected in the directional changes in the range.

5. Significance and Use

5.1 This guide is intended to encourage consistency in the analysis, interpretation, and modeling of spatial variation.

5.2 This guide should be used in conjunction with Guides D 5549, D 5923, and D 5924.

6. Analysis of Spatial Variation

6.1 The principal tools for analyzing spatial variation are the variogram and the correlogram; whenever possible, both should be used.

NOTE 1—Features that appear on both the variogram and correlogram are usually worthy of interpretation and should be reflected in the mathematical model for spatial variation. Features that appear on one but not the other may reflect artifacts of the calculation or peculiarities of the available data and their configuration; such features require further investigation before a decision can be made on whether they should be reflected in the mathematical model for spatial variation.

6.2 If univariate data analysis has revealed that the data have a skewed distribution or if study objectives require that the data be transformed, then the analysis of spatial variation should be performed on an appropriate transform of the data.

NOTE 2—One of the most important aspects of a mathematical model of spatial variation is the direction and degree of anisotropy. This is often much better revealed by variograms and correlograms of transformed data values, such as logarithms or normal scores. Even if the study ultimately makes use of the original data values in estimation or simulation, the analysis of spatial variation on transformed data values often leads to the development of a more appropriate model of spatial variation.

6.3 The choice of lag spacing and tolerance should take into account the data configuration, particularly the minimum spacing between the available data and the average spacing between the available data. Whenever possible, the choices of lag spacing and tolerance should ensure that at least 20 paired data values will be available for each lag.

NOTE 3—With data configurations that are pseudo-regular, it is common to use the spacing between the columns and rows of the sampling grid as the lag spacing and to use half of this distance as the lag tolerance. If the data configuration is irregular, then the lag spacing and tolerance may also need to be irregular (see Refs **(3)**, and **(6)**).

6.4 Spatial variation should be analyzed in different directions; the choice of directions and directional tolerances should reflect the configuration of the available data and should also take into account qualitative information about the physical and chemical characteristics of the regionalized variable being studied.

NOTE 4—Omni-directional variograms or correlograms often are appropriate for refining decisions on lag spacing and lag tolerance; they also provide preliminary insight into the range of correlation and the short-scale variability of the data. However, omni-directional calculations of spatial variation do not constitute a thorough analysis of spatial variation since they offer no insight into directional anisotropies that commonly occur in geologic data. For two-dimensional (2D) problems, contour maps of the spatial variation for all possible distances and directions can assist with the identification of directions of maximum and minimum continuity (see Ref **(2)**).

6.5 Once the directions of maximum and minimum continuity have been identified, the experimental variogram values and correlogram values should be plotted as a function of distance for these two directions. These plots should be accompanied by tables that provide for each lag the following information:

6.5.1 The number of data pairs within the lag,

6.5.2 The average separation distance between the data pairs in the lag,

6.5.3 The average squared difference between the paired data values in the lag, and

6.5.4 The coefficient of linear correlation between the paired data values in the lag.

6.6 Supplementary information that often assists with an appropriate interpretation of the pattern of spatial variation includes:

6.6.1 The average separation direction between the data pairs in the lag,

6.6.2 The average difference between the paired data values in the lag,

6.6.3 The variance of the differences between the paired data values in the lag,

6.6.4 The average of the data values within the lag, and

6.6.5 The variance of the data values within the lag.

7. Interpretation of Spatial Variation

7.1 If there are fewer than 20 pairs of data within a lag, the experimental variogram and correlogram values for this lag should not influence the interpretation of spatial variation.

NOTE 5—For lags that contain fewer than 20 pairs, the variogram and correlogram values should still be tabulated and plotted, but should be annotated or marked in such a way that it is clear that the lag contains insufficient data to influence any interpretation of spatial variation.

7.2 If the variance of the data within each lag depends on the average distance of the paired data values within the lag, then the experimental variogram should not influence the interpretation of spatial continuity.

NOTE 6—With data sets in which there is intentional clustering of additional samples in areas with high data values, the variance of the most closely spaced data pairs is often much higher than that of the overall data set. In such situations, the fluctuations in the variogram often mirror the changes in the lag variance and do not directly convey meaningful information about spatial variation. The correlogram is often more easily interpreted since it takes into account fluctuations in the lag means and the lag variances.

7.3 The nugget effect shown by the experimental variogram and correlogram should be consistent with the precision of the data values.

NOTE 7—The nugget effect reflects a combination of the very short scale variability and variance due to various sampling errors. Even in situations where there is no short scale variability, the nugget effect of an experimental variogram must be at least as large as the variance of the errors that accumulate during sample collection, preparation, and analysis.

7.4 Directional anisotropy in the ranges shown by the experimental variogram and correlogram should be supported by information on the physics and chemistry of the regionalized variable under study or by information on site history and usage.

7.5 If the experimental variogram or correlogram does not reach a stable sill, then the report on spatial variation should specifically address the issue of whether the data values show a trend or drift.

NOTE 8—A drift often causes the sill of an experimental variogram to increase steadily rather than stabilize. Both estimation and simulation studies will usually be improved by an explicit attempt to remove any drift, and to add it back after the residuals have been geostatistically analyzed and modelled.

7.6 If the sill of the experimental variogram or correlogram oscillates in a periodic manner, then the report on spatial variation should specifically address the issue of whether chemical or physical processes have imparted any cyclicity or periodicity to the data values.

NOTE 9—Temporal fluctuations may manifest themselves as periodic fluctuations in vertical profiles through soils and sediments. Such periodic fluctuations often cause the sill of an experimental variogram to show a dampened oscillation. Both estimation and simulation studies will usually be improved by an explicit attempt to remove any periodicity as a drift, and to add it back after the residuals have been geostatistically analyzed and modeled.

7.7 If the experimental variogram and correlogram show directional anisotropy, then the report on spatial variation should specifically address the reasons for improved continuity in certain directions.

NOTE 10—The direction of maximum continuity is usually identified as the direction in which the experimental variogram or correlogram shows the longest range. In situations where the experimental variogram or correlogram shows several components or structures, each with a different range, the direction of maximum continuity is usually identified as the one in which the experimental variogram or correlogram shows the lowest slope at the origin.

8. Modeling of Spatial Variation

8.1 The mathematical model selected to approximate the behavior of the experimental correlogram should be positive definite (see Ref **(1)**).

NOTE 11—In practice, this requirement of positive definiteness is usually satisfied by fitting the experimental variogram or correlogram with one of a handful of well-studied functions that are known to be positive definite. These include the spherical, exponential, and gaussian variogram models described in any of the references given in the reference list at the end of this guide. Linear combinations of positive definite functions are also positive definite, and many practical applications require the mathematical model of spatial variation to include two or more components to adequately describe the spatial variation.

8.2 The mathematical function of spatial variation should provide a good approximation of the short-scale behavior of the experimental variogram or correlogram and of any directional anisotropies.

NOTE 12—Geostatistical simulation and estimation algorithms that make use of the mathematical model of spatial variation are usually most sensitive to the behavior of the model at very short distances and to the direction and degree of anisotropy. The mathematical model should therefore capture the nugget effect and any directional anisotropy in the slope near the origin of the experimental variogram or correlogram.

8.3 If the pattern of spatial variation of a regionalized variable is directionally anisotropic, as is usually the case, the mathematical function of spatial variation should use a consistent shape and sill in all directions and should express the directional anisotropy solely through appropriate changes in the range(s) of the different components.

NOTE 13—Usually, the experimental sill is the same in all directions although the ranges may vary. This is known as a "geometric anisotropy." Even where the experimental sill appears to change with direction, a situation described as "zonal anisotropy," an appropriate mathematical model can still be fit to the experimental data by expressing the directional anisotropy solely through appropriate changes in the ranges of the different components, as described in Ref **(2)**.

8.4 The results of automatic fitting procedures should be reviewed visually.

NOTE 14—Though automatic fitting procedures often expedite the process of developing a mathematical model of spatial variation, their results should not be used without careful review. The results of such a procedure should be plotted along with the experimental variogram or correlogram and any significant differences should be reconciled. Though additional components in the mathematical model may improve an automatic fit, the need for these additional components should be reconciled with the objective of the study and with qualitative information; if additional components cannot be explained by the physics or chemistry of the attribute and they cannot be justified by the goals of the study, then a simpler model with fewer components should be used.

9. Keywords

9.1 correlogram; geostatistics; variogram

REFERENCES

(**1**) Journel, A. G., and Huijbregts, C., *Mining Geostatistics*, Academic Press, London, 1978.

(**2**) Isaaks, E. H., and Srivastava, R. M., *An Introduction to Applied Geostatistics*, Oxford University Press, New York, 1989.

(**3**) Deutsch, C. V., and Journel, A. G., *GSLIB Geostatistical Software Library an User's Guide*, Oxford University Press, New York, 1992.

(**4**) Srivastava, R. M., and Parker, H. M., "Robust Measures of Spatial Continuity," *Geostatistics*, M. Armstrong, ed., Reidel, Dordrecht, 1988, pp. 295–308.

(**5**) Olea, R. A., ed., *Geostatistical Glossary and Multilingual Dictionary*, Oxford University Press, New York, 1991.

(**6**) Yates, S. R., and Yates, M. V., *Geostatistics for Waste Management: A User's Manual for the GEOPACK (Version 1.0) Geostatistical Software System*, EPA/600/8-90/004 (NTIS PB90-186420), 1990.