Designation: D 5549 – 94ε1

AMERICAN SOCIETY FOR TESTING AND MATERIALS
100 Barr Harbor Dr., West Conshohocken, PA 19428
Reprinted from the Annual Book of ASTM Standards. Copyright ASTM

Standard Guide for The Contents of Geostatistical Site Investigation Report[1]

This standard is issued under the fixed designation D 5549; 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 1.7 was added editorially October 1998.

INTRODUCTION

Geostatistics is a framework for data analysis and estimation in media whose measurable attributes show erratic spatial variability yet also possess a degree of spatial continuity imparted by the natural and anthropogenic process 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 reporting of such investigations.

1. Scope

1.1 This guide covers the contents required for a complete report of a geostatistical site investigation. A complete report is understood here to be one that contains all the information necessary to the understanding and evaluation of the geostatistical site investigation by other geostatisticians.

1.2 This guide does not discuss the reporting of supplementary information that may assist evaluation of the report.

1.3 While geostatistical methods are used in many fields, this guide is primarily intended for the reporting of environmental and geotechnical applications. Environmental applications of geostatistics are discussed in United States Environmental Protection Agency (EPA) documents **(1-7)**.[2]

1.4 The basic geostatistical methods referred to in this guide are fully described in texts by David **(8)**, Journel and Huijbregts **(9)**, Clark **(10)**, and Isaaks and Srivastava **(11)**. Olea **(12)** gives a thorough compilation of geostatistical terminology.

1.5 This guide does not discuss the reporting of multivariate, space-time, and other less-frequently used geostatistical methods; however this is not intended to reflect any judgment as to the validity of these methods.

1.6 Geostatistics is but one approach that can be used to understand and describe site conditions. Investigations should incorporate whatever supplementary knowledge of the site that may be available from other sources. As with classical statistical approaches, geostatistics is not intended to establish cause-and-effect relationships.

1.7 *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]

3. Terminology

3.1 *Definitions of Terms Specific to This Standard:*

3.1.1 *additivity, n*—a mathematical property of a regionalized variable stating that it can be combined linearly in order to define a similar variable on a larger support.

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

3.1.3 *covariance, n*—a measure of similarity between two variables defined as the expected value of their product minus the product of their expected values and often used to measure spatial continuity between two variables as a function of the lag distance separating their locations.

3.1.4 *cross validation, n—in geostatistics*, a procedure by which the suitability of an estimation method is assessed by

[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 June 15, 1994. Published September 1994.

[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.

sequentially removing and estimating sample data, and comparing the resulting pairs of true and estimated values.

3.1.5 *drift, n—in geostatistics*, a systematic spatial variation that is usually expressed as a polynomial function of location coordinates.

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

3.1.7 *field, n—in geostatistics*, the region of one-, two-, or three-dimensional space within which a regionalized variable is defined.

3.1.8 *geostatistics, n*—the application of random function theory to the analysis and estimation of natural phenomena characterized by the eratic spatial variability of their measurable attributes.

3.1.9 *kriging, n*—a linear estimation method where sample weights are obtained using a least-squares optimization procedure based on a mathematical model of spatial continuity and where the unknown variable may have a point or a block support.

3.1.10 *kriging variance, n*—the expected value of the squared difference between the value of an unknown variable and its kriging estimate, sometimes used as a measure of kriging precision.

3.1.11 *lag, n—in geostatistics*, the distance vector separating the locations of two variables, as used in measures of spatial continuity.

3.1.12 *point, n—in geostatistics*, the location in the field at which a regionalized variable is defined. It also commonly refers to the support of sample-scale variables.

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

3.1.14 *realization, n*—an outcome of a random function or a random variable.

3.1.15 *search neighborhood, n*—the region over which samples are considered for inclusion in the kriging process.

3.1.16 *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 the case of conditional simulation, the realizations are constrained to honor data values measured at sampled locations.

3.1.17 *support, n—in geostatistics*, the spatial averaging region over which a regionalized variable is defined, often approximated by a point for sample-scale variables.

3.1.18 *variogram, n*—a measure of spatial continuity defined as one half the variance of the difference between two variables and expressed as a function of lag distance. It is also sometimes referred to as the semi-variogram.

3.2 For definitions of other terms used in this guide, refer to Terminology D 653.

4. Summary of Guide

4.1 This guide is summarized in Table 1 that shows the suggested table of contents for a geostatistical site investigation report and the corresponding sections of the standard that apply to each heading. For certain headings, additional references can be found in standard guides in preparation.

TABLE 1 Suggested Table of Contents for a Geostatistical Site Investigation Report

Report Heading	Section	Other References
Objectives	6.2	none
Site description	6.3	none
Data sources	6.4	Guide in preparation
Definition of field	6.5	none
Choice of regionalized variable	6.6	none
Exploratory data analysis	6.7	none
Spatial continuity analysis	6.8	Guide in preparation
Estimation	6.9	Guide in preparation
Simulation	6.10	Guide in preparation
Conclusions and recommendations	6.11	none
References	6.12	none
Computer software	6.13	Guide in preparation

5. Significance and Use

5.1 This guide is intended to encourage consistency and thoroughness in the reporting of geostatistical site investigations by describing the basic information required in a complete report.

5.2 Referring to the table of contents suggested in Table 1, this guide may be used as a template by those preparing reports or as a checklist for review and auditing purposes by qualified nonparticipants in the study.

6. Contents of a Geostatistical Site Investigation Report

6.1 The following subsections describe the contents of a complete geostatistical site investigation report, in the order suggested in Table 1.

6.2 *Objectives*—The report of a geostatistical site investigation should contain a section stating study objectives. Throughout the report, selected procedures, assumptions, and approximations should be justified in terms of these objectives. Objectives must be shown to be consistent with the data quality, sample spacing, and site coverage available.

6.3 *Site Description*—The report of a geostatistical site investigation should contain a section summarizing site history and conditions, both physical and chemical, relevant to the study.

6.3.1 Maps should be used to show the location of the site relative to state or county boundary lines; roads; population centers; bodies of water, and other physiographic features.

6.3.2 Detailed site plans should be provided if the site has numerous cadastral, physical, or biological features that influence access to sample data locations or future survey locations related to the geostatistical study. Plans should be provided at the same scale and degree of accuracy and precision as that provided with sample location and site boundaries map(s).

6.3.3 Graphical displays such as maps or sections should be used to show data locations, site boundaries, and all other horizontal and vertical controls affecting boundaries or data locations.

6.4 *Data Sources and Accessory Information*—The report of a geostatistical site investigation should contain a section describing all qualitative and quantitative information needed for a geostatistician to evaluate the impact of the data as they will affect the study results.

6.4.1 Sources of all physical and chemical data relevant to the geostatistical site investigation should be reported, identifying sample data in accordance with the sampling method, time of collection, sample gathering agency, and laboratory.

6.4.2 Complete tables listing the data selected for the geostatistical site investigation should be included in the report. If the combined data set is large, use of an appendix or electronic storage media is recommended. The source reference should be identified for each separate data table page.

6.4.3 All quantitative and qualitative errors and potential out-of-control events identified in the data source(s), including those pertaining to sample location, should be summarized in an accuracy and precision statement.

6.4.4 Results of previous studies where relevant data were collected but not used in the current study should be summarized, providing a rationale for their exclusion.

6.5 *Definition of the Field*—The report of a geostatistical site investigation should contain a section describing the region in physical space and the period in time that define the geostatistical field.

6.5.1 The decision to use a one-, two-, or three-dimensional geostatistical field should be justified in terms of the study objectives, the choice of regionalized variable, the sampling technique, and the amount of data available.

6.5.2 The field must be situated in space and its boundaries, either physical or jurisdictional, should be specified. Field boundaries should be displayed in graphical form, showing data locations and site boundaries.

6.5.3 The field should be situated in time, indicating the period over which sample data were obtained.

6.5.4 If a site is partitioned into different fields for separate geostatistical analysis, the reasons for the decision should be documented.

6.6 *Choice of Regionalized Variable*—The report of a geostatistical site investigation should contain a section documenting the choice of regionalized variable with reference to the study objectives and the sampled attributes of the medium.

6.6.1 The report must demonstrate that all data for a regionalized variable have the same support size in the geostatistical field.

6.6.2 The report must demonstrate the additivity of the regionalized variable if study objectives require procedures involving a change of support.

6.6.3 The report must demonstrate that all data for a regionalized variable are time-invariant or were measured synchronously at all sample locations.

6.6.4 If logarithmic, indicator, gaussian, uniform, or other mathematical transformations are applied to the original regionalized variable, the transformation procedure and the reasons for performing such a transformation should be documented.

6.6.5 If a transformed regionalized variable was used throughout a geostatistical study, the report must demonstrate that a back-transformation, if applied, does not introduce any unwanted bias in the spatial statistics of the restored variable.

6.6.6 If the regionalized variable is a numerical quantity assigned to qualitative or subjective attributes of the medium, the numerical coding should be described.

6.7 *Exploratory Data Analysis*—The report of a geostatistical site investigation should contain a section on exploratory data analysis. Exploratory data analysis refers to the univariate, bivariate, or multivariate statistical analyses of the regionalized variable generally conducted at the beginning of the geostatistical study.

6.7.1 Any grouping of data into subsets for separate geostatistical analysis should be justified with respect to the results of exploratory data analyses and study objectives.

6.7.2 The report should present a table of summary statistics and a histogram or box-plot of the regionalized variable, for each data subset identified.

6.7.3 The treatment of below-detection-limit or outlier data values should be documented.

6.7.4 The influence of sample location clustering on the results of data analyses should be discussed and addressed if necessary.

6.7.5 The identification and, if necessary, removal of a drift in the data should be documented and justified with respect to any assumption of nonstationarity.

6.8 *Spatial Continuity Analysis*—The report of a geostatistical site investigation should contain a section presenting an analysis of spatial continuity for the regionalized variable. The most common measures of spatial continuity include the variogram, covariance, and correlogram.

6.8.1 The particular measure(s) of spatial continuity should be reported along with a rationale for the choice and an assessment of the sensitivity of the measure(s) to clustering of data locations, sampling error, and extreme values.

6.8.2 The determination of the lags used in the estimation of a measure of spatial continuity from irregularly spaced data should be clearly explained.

6.8.3 The experimental values of the spatial continuity measure should be tabulated for each lag, along with the average lag distance and direction of data pairs contributing to that lag, and the mean, variance, and the number of data contributing to that lag.

6.8.4 The experimental values of the spatial continuity measure should also be displayed graphically as a plot of experimental values against lag distance. If plots for multiple directions are superimposed, each should be clearly identified along with its associated directional tolerance. Any smoothing procedure applied to the experimental values should be documented.

6.8.5 Any mathematical function fit to the experimental continuity measure should be stated along with its parameters and should be displayed on the same plot as the experimental values. Any such model should be rationalized by establishing a clear link between it and supporting geological or qualitative information.

6.8.6 If cross-validation is used to assess the mathematical model of spatial continuity used in kriging, results should be displayed in graphical form, showing the magnitude and location of cross-validation errors.

6.9 *Estimation*—The report of a geostatistical site investigation may contain a section on estimation. If estimation is performed rather than simulation, the rationale for this choice should be clearly stated with reference to study objectives. Kriging, in its various forms, is the most common estimation method used in geostatistical investigations.

6.9.1 The particular form of kriging that is chosen should be

documented and justified with respect to study objectives and assumptions.

6.9.2 The support, point or block, of the quantity estimated by kriging should be stated, specifying dimensions and numerical discretization in the case of blocks.

6.9.3 If block-support quantities are estimated from point-support values, the additivity of the regionalized variable should be demonstrated.

6.9.4 The origin, orientation, and spacing of the kriging grid should be documented in graphical form, showing sample locations and boundaries of the geostatistical field.

6.9.5 The size, shape, and orientation of the kriging sample search neighborhood should be documented along with other information pertaining to the sample search strategy of the kriging software employed.

6.9.6 If cross-validation is used to evaluate the estimation procedure, results should be displayed in graphical form, showing the magnitude and location of cross-validation errors.

6.9.7 Kriging estimates for the gridded area should be presented in tabular form and in graphical form showing sample data locations.

6.9.8 If used, kriging variances for the gridded area should be presented in tabular form and in graphical form showing sample data locations.

6.10 *Simulation*—The report of a geostatistical site investigation may contain a section on simulation. If simulation is performed rather than estimation, the rationale for this choice must be clearly stated with reference to study objectives.

6.10.1 The choice of simulation method and whether it is conditional or nonconditional should be documented and justified with respect to study objectives.

6.10.2 If simulated values are to have a different support than the data, the report should demonstrate that conditioning and change-of-support processes ensure that the correct spatial variability of simulated values is reproduced.

6.10.3 The report should list information used to condition the simulation together with a rationale for its use and documentation of the degree to which it was honored by the simulation.

6.10.4 The report should state clearly the univariate distribution model that is to be honored and address the issue of whether the simulated regionalized variable has the same distribution as the sample data.

6.10.5 The report should state clearly the spatial continuity model that is to be honored and address the issue of whether the simulated regionalized variable has the same spatial continuity as the sample data.

6.10.6 The simulated field, or some suitable subset should be displayed in graphical form showing sample data locations.

6.10.7 If simulation is used to generate multiple realizations, the report should document the selected number of realizations along with justification as to why this particular number is sufficient.

6.11 *Conclusions and Recommendations*—The report of a geostatistical site investigation should contain a section of conclusions and recommendations.

6.11.1 Each conclusion should be supported by clearly identified results presented within the study and their associated degree of uncertainty expressed quantitatively or qualitatively.

6.11.2 The sensitivity of each conclusion to sources of error or uncertainty affecting the results should be discussed.

6.11.3 Each conclusion should respond to the original legal, regulatory, or technical question(s) proposed in the study objectives.

6.11.4 Each conclusion must be accompanied by a description of the spatial and temporal boundaries of the site to which it applies.

6.11.5 Recommendations should follow the reporting of all conclusions. Each recommendation should be accompanied by a rationale. If further experimentation is recommended, either by additional data collection or by further mathematical application, revised objectives should be proposed.

6.11.6 Recommendations should identify assumptions that are most likely to be revised in the future following the interpretation and analysis of any additional data.

6.12 *References*—The report of a geostatistical site investigation should contain a section providing references for all prior studies and other information sources used to prepare the report. References should be given within the report text where cited and should be keyed to a complete reference list at the end of the report. A bibliography of other information and data sources not used, but pertaining to the study site should also be provided.

6.13 *Appendix: Computer Software*—The report of a geostatistical site investigation should contain an appendix describing the computer implementation of statistical and geostatistical methods used in the study.

6.13.1 If commercial or public domain software is used, the authors and source references, such as user's guides should be noted in the report. Information concerning the capabilities, limitations and popular acceptance of the software should also be provided, if available.

6.13.2 If self-developed software is used, its theoretical basis and numerical implementation should be described. Capabilities, limitations, and checks of the software should also be documented.

7. Keywords

7.1 environmental assessment; geostatistics; kriging; simulation; variogram

REFERENCES

(**1**) U.S. Environmental Protection Agency (EPA), *Data Quality Objectives for Remedial Response Activities: Development Process*, EPA/540/G-87/003 (NTIS PB88-131370), OSWER Directive 9355.0-7B, March 1987.

(**2**) U.S. Environmental Protection Agency (EPA), *Data Quality Objectives for Remedial Response Activities: Example Scenario: RI/FS Activities at a Site with Contaminated Soils and Groundwater*, EPA/540/G-87/004 (NTIS PB88-131388), OSWER Directive 9355.0-7B, March 1987.

(**3**) Englund, E. J., and Sparks, A. R. *Geo-EAS (Geostatistical Environmental Assessment Software) User's Guide*, EPA/600/4-88/033 (NTIS PB89-151252), Environmental Monitoring Systems Laboratory, Las Vegas, NV, March 1988.

(**4**) Barth, D. S., Mason, B. J., Starks, T. H., and Brown, K. W., *Soil Sampling Quality Assurance User's Guide*, 2nd ed., EPA/600/889/046 (NTIS PB89-189864), Environmental Monitoring Systems Laboratory, Las Vegas, NV, March 1989.

(**5**) 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.

(**6**) U.S. Environmental Protection Agency (EPA), *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Interim Final Guidance*, EPA/530/SW-89/026 (NTIS PB89-151047), September 1989.

(**7**) U.S. Environmental Protection Agency (EPA), *Guidance for Data Usability in Risk Assessment*, EPA/540/G-90/008 (NTIS PB91-921208), Office of Solid Waste and Emergency Response, OSWER Directive 9285.7-05, 1990.

(**8**) David, M., *Geostatistical Ore Reserve Estimation*, Elsevier, Amsterdam, 1977.

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

(**10**) Clark, I., *Practical Geostatistics*, Applied Science Publishers, London, 1979.

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

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