INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY

ANALYTICAL CHEMISTRY DIVISION

TERMINOLOGY IN SOIL SAMPLING

(IUPAC Recommendations 2003)

Prepared for publication by
PAOLO DE ZORZI[1], SABRINA BARBIZZI[1], MARIA BELLI[1], GIOVANNI CICERI[2], ALES FAJGELJ[3], DAVID MOORE[4] AND UMBERTO SANSONE[1]

[1] *Agenzia per la protezione dell'ambiente e per i servizi tecnici-APAT, Servizio di Metrologia Ambientale, Via Castel Romano, 100-102, 00128 Roma, Italy*

[2] *CESI Spa – Business Unit Environment. Via R Rubattino, 54, 20134 Milano, Italy*

[3] *IAEA, Agency's Laboratories Seibersdorf and Vienna, Wagrammer Strasse 5, A-1400, Vienna, Austria*

[4] *Materials Dynamics, Los Alamos National Laboratory, Los Alamos, NM 87545 USA*

Terminology in Soil Sampling

(IUPAC Recommendations 2003)

Abstract

The need to be understood is the first objective of writers and speakers, be a poet or a scientist. But there is difference: the scientist must be sure that, within a stated context, the terms used in articles, publications or in the daily conversation among colleagues, are intended by all in the same precise way, without any possible ambiguity. As already pointed out by IUPAC recommendation 1990 "Nomenclature for Sampling in Analytical Chemistry", it is not acceptable that scientists are not able to orient themselves in a sampling or analytical process. This can occur if the terms used are not well defined. Moreover, to better appreciate the development of new theories or concepts, progressive update can be necessary. To this end, on the basis of the existing terminology documents and of the most recent knowledge in the field of soil sampling, an up-dated terminology in sampling (specifically soil sampling) is recommended.

CONTENTS

1.0 INTRODUCTION

Recently, there have been a growing number of efforts for overcoming confusion, ambiguity and contradictions in usage of terms and clarification of their definitions in the field of sampling. The IUPAC Recommendations, "Nomenclature for Sampling in Analytical Chemistry", published in 1990 and the ISO standard 11074-2, "Soil Quality – Vocabulary – Part 2, 1998", are the most widely used terminology documents related to soil sampling. However, recent developments and studies of various sampling aspects (i.e. uncertainty quantification, method validation, comparison of sampling tools and strategies) require new concepts to be developed and also some new terms to be introduced for their description.

One of the outcomes of the SOILSAMP international project, funded and coordinated by the

Italian Environmental Protection Agency (APAT, Italy) and aimed at assessing the uncertainty associated with soil sampling in agricultural, semi-natural, urban and contaminated environments, was an updated terminology in sampling.
This document is the result of that effort, and is intended to present terms and definitions to be used in the field of soil sampling and sampling uncertainty. A set of geostatistical terms, of interest in the context of soil sampling and sampling uncertainty estimation, are also illustrated together with the recommended definitions.

2.0 SOIL SAMPLING TERMINOLOGY

A wide variety of terms are used in the practice of sampling. However, many terms identify a specific (single) operation and in some case different terms are used to describe the same concept. Furthermore it is quite usual in the scientific literature that different authors used different terms to describe the same operation or concept. Also in works produced by standardisation bodies it is possible to find different terms for describing the same operation or concept as well as a different philosophy of organising the concepts themselves.

In this paper, selection of terms recommended for use in connection with soil sampling is made with the aim to improve the consistency of related sampling terminology. In addition, some new terms are proposed. To facilitate the understanding and real meaning of such terms, schemes illustrating the relationship of the operations related to sampling are presented (see Figure 1 and 2).

Sampling (Figure 1) starts from the selection of the "test area" which should include all of the potentially impacted "sampling sites". Samples are then collected at the different "sampling points" identified inside of a number of "sampling units". The number and shape of the "sampling units" depend on the scope of sampling and thus on the particular "sampling pattern/strategy" which is selected on a statistical basis.

"Increments" are taken from each "sampling point" to produce a "primary sample". "Increments" can be kept separate. When mixed together a "composite/aggregate sample" is obtained.

A "laboratory sample" is obtained either directly from the "primary sample" (usually from the "bulk sample"), or by "reduction" of the "primary sample" or of the "composite samples". During this phase "coning and quartering", "quartering", "riffling", "grinding", etc. may be necessary. The laboratory sample is then packed and shipped to the laboratory for required pre-treatment and characterisation (chemical, physical, biological, etc.).

2.1 Terms and definitions

Composite sample (average sample, aggregate sample) (ISO11074-2, 1998)
"*Two or more increments/sub-samples mixed together in appropriate proportions, either discretely or continuously (blended composite sample), from which the average value of a desired characteristic may be obtained*"**[1)]**;
Increment (IUPAC, 1990):
"*Individual portion of material collected by a single operation of a sampling device.*";**[2)]**

---

[1] IUPAC, 1990, considers the composite sample only with reference to sampling of bulk material when more increments from the bulk are combined to obtain a physically averaged sample. However, the term composite sample is not specified.
[2] Increments may be reduced and tested individually or combined with other increments, with the resulting composite reduced in size and tested as a single unit.

Informative judgmental sampling
"Sampling in which locations are chosen according to the judgement of an expert and partly in accordance with the statistical principles of sampling";
Judgemental sampling (ISO 11074-2)
"Sampling in which locations are chosen according to the judgement of an expert ";
Laboratory sample (IUPAC, 1990)
"*The sample or sub-sample sent to or received by the laboratory*"[3)];
Primary sample (IUPAC, 1990)
"*The collection of one or more increments or units initially taken from a population*";
Replicate (Duplicate) samples (ISO11074-2, 1998)
"*One of the two or more samples or sub-samples obtained separately at the same time by the same sampling procedure or sub-sampling procedure*"[4)];
Representative sample (IUPAC, 1990; ISO11074-2, 1998):
"*Sample resulting from a sampling plan that can be expected to reflect adequately the properties of interest in the parent population*";
Sample (IUPAC, 1990; ISO11074-2, 1998):
"*A portion of material selected from a larger quantity of material*";
Sample size (ISO 11074-2: 1998):
"Number of items or the quantity of material constituting a sample";
Sampler (ISO 11074-2: 1998)"[5)]
"Person carrying out the sampling procedures at the sampling locality";
Sampling (ISO11074-2, 1998):
"*Process of obtaining a sample*" ;[6)]
Sampling design (ISO 11074-2: 1998): [7)]
"Arrangement by which a sampling programme is to be conducted" ;
Sampling device (ISO11074-2, 1998)
"*Apparatus/tool to obtain a sample*";
Sampling pattern/sampling strategy

Increments are created by the sampling operation and are usually taken from parts of a lot separated in time or space.
Increments of a bulk population correspond to units of a packaged population.

[3] Even if the definitions in ISO11074-2 and IUPAC 1990 are slightly different, the notes attached are the same. Mainly, it is noted that the laboratory sample is the final sample from the point of view of the sample collection but it is the initial sample from the point of view of the laboratory.

[4] In IUPAC 1990, this term has a slightly different definition. In that case it is stressed that the comparability has to be assured in space and/or time. The Note is the same in both standards. Although the replicate samples are expected to be identical, often the only thing replicated is the act of taking the physical sample. A duplicate sample is a replicate sample consisting of two portions.

[5] Tools and other devices to obtain samples are sometimes also designated 'samplers'. In this case the terms 'sampling devices' or 'sampling equipment' should be used. The sampler should have specific knowledge and experience in soil sampling.

[6] Sampling, ordinarily ends, after pre-treatment steps, with the removal of the test (or analytical) portion from the test (or analytical) sample. For "test portion" and "test sample" see the related definitions (3.1). 'Sampling' also relates to the selection of locations for the purpose of *in situ* testing carried out in the field without removal of material.

[7] The purpose of designing a sampling programme is to provide the most efficient methods to reach valid and relevant conclusions from the investigations of soil, with due regard to cost or resource use commensurate with sampling programme objectives. The design is a function of many considerations such as the aim of the investigation, the (degree of) heterogeneity of the material under consideration and the cost of performing the investigation.

"The result of the subdivision of the test area into sampling sites and sampling units" ; [8)]

Sampling point

"The place where sampling occurs within the sampling unit" ;

Sampling plan (IUPAC, 1990; ISO11074-2, 1998)

"*Predetermined procedure for the selection, withdrawal, preservation, transportation and preparation of the portions to be removed from a population as a sample*";

Sampling procedures (ISO11074-2, 1998)

"*Operational requirements and/or instructions relating to the use of a particular sampling plan*";

Sampling site

"A well delimited site inside the test area, where sampling operations take place" ;[9)]

Sampling techniques (ISO11074-2, 1998)

"*All appropriate procedures and sampling devices used to obtain and describe samples of soil, either in the field or during transportation and in the laboratory*";

Sampling unit

"At the sampling site, the portion of soil surface, defined in physical term, inside which is located the sampling point" ;

Simple random sample (ISO11074-2, 1998)

"*Sample of n items taken from a population of N items in such a way that all possible combinations of n items have the same probability of being taken*";

Spot sample (ISO 11074-2)

"Sample of specified number or size taken from a specified place in the material or at a specified place and time in a stream of material and representative of its own immediate or local environment." ;

Stratification (ISO 3534-1)

"The division of a population into mutually exclusive and exhaustive sub-populations (called strata), which are thought to be more homogeneous with respect to the characteristics investigated than the total population" ;

Stratified sample (IUPAC, 1990; ISO11074-2, 1998)

"*Sample obtained from strata or subparts, putatively homogeneous of the parent population*"[10)]

Sub-sample (ISO11074-2, 1998)

"Sample obtained by a procedure in which the items of interest are randomly distributed in part of equal or unequal size"[11)];

Sub-sampling (sample division) (ISO11074-2, 1998)

"*Process of selection one or more sub-samples from a sample of a population*" ;[12]

---

[8] Sampling patterns are selected (using a statistical approach) according to the scope of sampling and to the need of obtaining representative samples

[9] Depending to the scope of the sampling, one or more "sampling sites" may be identified within the "Test area" on the basis of available data or after area recognisance.

[10] The objective of taking stratified samples is to obtain a more representative sample than that which might otherwise be obtained by random sampling. A mean value of a measurand for the whole lot may be determined on a blended composite sample from individual increments or mathematically obtained from separate determinations on each increment.

[11] The Note in ISO11074-2, 1998 is totally corresponding to the definition of "sub-sample" in IUPAC, 1990. That is "A sub-sample may be: a) a portion of the sampled obtained by selection and division; b) an individual unit of the lot taken as part of the sample; c) the final unit of the multistage sampling (ISO11074-2, def. 3.6)

[12] In IUPAC, 1990 there is not the term "sub-sampling" but at point 2.4 (Sample preparation) are reported the definitions related to different procedures for the sample division such as "reducing", "coning and quartering", "riffling". Some of these definitions are in compliance.

Systematic sampling (ISO11074-2, 1998)
"*Sampling by some logical and organized method*";
Stratified sampling (ISO 3534-1)
"*In a population which can be divided into mutually exclusive and exhaustive sub-populations (called strata), sampling carried out in such a way that specified proportions of the sample are drawn from the different strata and each stratum is sampled with at least one sampling unit* [10)]
Test area
"The area to be characterized according to regulation rules or for scientific purposes" ; [14)]
Unit, item, portion, individual (IUPAC, 1990; ISO11074-2, 1998) [15)]
"*Each of the discrete, identifiable portions of material suitable for removal from a population as a sample or as a portion of a sample, and which can be individually considered, examined, tested or combined*";

3.0 OPERATIONS RELATED TO SAMPLE PRE-TREATMENT AND ANALYSIS

Pretreatment can be performed either at the sampling site in order to obtain the "Laboratory sample" or at the laboratory in order to obtain a "Test Sample" from the "Laboratory sample" for analytical purposes.

At the laboratory, the "laboratory sample" is split (by "reducing", "milling" etc.) into a number of "test samples" to be used for the analytical characterisation.

According to the particular analytical test, a "test portion" is obtained from the "test sample" (Figure 2) and submitted to the various analytical steps. For example, if dissolution/extraction and clean-up operations are required, an aliquot of the treated test solution is finally sent to the measurement. If no treatment is foreseen, the test portion (solid) is directly measured.

3.1 Terms and definitions

Aliquot (ISO11074-2, 1998)
"Known amount of a homogeneous material, assumed to be taken with negligible sampling error" ;
Coning and quartering (IUPAC, 1990)
"*The reduction in size of a granular or powdered sample by forming a conical heap which is spread out into a circular, flat cake. The cake is divided radially into quarters and two opposite quarters are combined. The other two quarters are discarded. The process is repeated as many time as necessary to obtain the quantity desired for some final use (e.g. as the laboratory sample or as the test sample)*"
Quartering
"*The reduction in size into quarters of a granular or powdered sample. Two opposite quarters are combined, while other two quarters are discarded. The process is repeated as many time as necessary to obtain the quantity desired for some final use (e.g. as the laboratory sample or as the test sample)*";

---

[14] Borders of this area should be selected in order to include all the possible zones that may be reached by the contamination. In the first instance the borders are usually administrative borders or factory perimeters
[15] The term "unit" is also used to identify the portion of soil surface at the "sampling site" in which the "sampling point" is located: for this meaning it is advisable to use the term "sampling unit" (see Figure 1)

Homogeneity/heterogeneity (IUPAC, 1990; ISO11074-2, 1998)
"*Degree to which a property or a constituent is uniformly distributed throughout a quantity of material*";
Milling, Grinding (IUPAC, 1990; ISO11074-2, 1998)
"*Mechanical reduction of the particle size of a sample by attrition, impact or cutting*";
Mixing (IUPAC, 1990; ISO11074-2, 1998)
"*Combining of components, particles or layers into a more homogeneous state*";
Reducing (IUPAC, 1990; ISO11074-2, 1998)
"*Decreasing the size of the laboratory sample or individual particles, or both*";
Riffling (IUPAC, 1990; ISO11074-2, 1998)
"*Separation of a free-flowing sample into (usually) equal parts by means of a mechanical device composed of diverter chutes*";
Sample division (ISO 11074-2)
"*Process of selecting one or more sub-samples from a sample of population"*;
Sample pretreatment (ISO11074-2, 1998)
"*Collective noun for all procedures used for conditioning a soil sample to a definite state which allows subsequent examination or analysis or long-term storage*";
Test portion (IUPAC, 1990; ISO11074-2, 1998)
"*Quantity of material, of proper size for measurement of the concentration or other property of interest, removed from the test sample"* ; **[16)]**
Test sample (IUPAC, 1990; ISO11074-2, 1998)
"*Sample, prepared from the laboratory sample, from which the test portions are removed for testing or for analysis*".

4.0 SAMPLING UNCERTAINTY TERMINOLOGY

All measurements are affected by uncertainty. This reflects the property of a measurement to determine a range of values that can be reasonably attributable to the quantity subject of the measurement. Many sources could influence the overall measurement uncertainty and in general all of them necessarily should be identified and, if possible, quantified.

Although scientists have, in general, already accepted these concepts, and a lot of effort has been made by standardization bodies to develop guides and rules, there is still confusion on colloquial usage of terms related to uncertainty. For example, terms such as error are normally used instead of uncertainty, contributing in some cases to possible misunderstanding, and the precise limits of application of the terms are not always well understood. In general the error can be detected but not quantified, as the true value of a measurement can never be determined. On the contrary, uncertainty is a measurable parameter, being based on the evaluation of the statistical distribution of the results of a series of measurements, described by standard deviations and variances.

In the frame of soil sampling and analysis, the uncertainty and all related terms have to be the basis of a common language. If in the analytical field a framework of the uncertainty terminology has been defined, the same effort on sampling has not yet produced a consensus. Sampling affects the analytical results, as well as sample preparation and treatment, contributing to the total measurement uncertainty. Sampling uncertainty can be properly quantified, following different

[16] The portion may be either combined or kept separate. If combined and mixed to homogeneity, it is a blended bulk sample

approaches and considering different situations (matrices, environment, parameter, concentration). Sampling uncertainty can be assumed as a parameter of the quality of sampling (and then of the measurement), in order to compare different sampling strategies/devices, to assess the sampling performance and, finally, to select an appropriate sampling technique and protocol for stated objectives.

Tools and parameters normally related to the analytical field and QA/QC schemes, such as reference materials, can find some analogues in the sampling field. From this point of view, terms such as "reference sampling" and "reference site" need to be explained. Through a reference sampling on a selected area, which is an intensive sampling performed by a single operator/sampler using a single sample device and following a defined pattern and protocol and the subsequent analysis of the samples collected, a well-characterized reference site can be obtained. Therefore, the quantity of a soil property (i.e. trace element concentration or pedo-chemical parameters, etc.) in each possible sample location can be well known with its uncertainty. The reference site is nothing more than a natural matrix reference material for sampling. When not all elements of interest are naturally present in this reference site, but one or more selected elements (of suitable, known, quantity and concentration) are added and homogenised into the soil, a synthetic reference material for sampling is obtained.

4.1 Terms and definitions

Characteristic (IUPAC, 1990; ISO11074-2, 1998)
"*Property or attribute of a material that is measured, compared and noted*";
Uncertainty (of measurement) (ISO, 1993)
"*A parameter associated with the result of a measurement that characterises the dispersion of the values that could be reasonably attributed to the measurand*";
Measurand (ISO, 1993)
"*Particular quantity subject to measurement*";
Measurement (ISO, 1993)
"*Set of operations having the object of determining the value of a quantity*";
Error (of measurement) (ISO, 1993)
"*The result of a measurement minus a true value of the measurand*";
Accuracy (ISO, 1993)
"*The closeness of the agreement between the result of a measurement and a true value of the measurand*";
Precision (ISO3534-1, 1993; ISO, 1993)
"*The closeness of the agreement between independent test result obtained under stipulated conditions*";
Bias (ISO3534-1, 1993)
"*The difference between the expectation of the test results and an accepted reference value*";[17].
Sampling error (IUPAC, 1990; ISO11074-2, 1998)
"*The part of the total error (the estimate from a sample minus the population value) associated with only a fraction of the population and extrapolating to the whole, as distinct from analytical or test error*";
Sampling uncertainty

[17] Bias is the total systematic error contrasted to random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the accepted reference value is reflected by a larger bias value.

"*The part of the total measurement uncertainty attributable to sampling*";

Reference sampling

"Characterization of an area, using a single sampling device and a single laboratory, to a detail allowing the set-up of a distribution model in order to predict element concentrations at any sampling point";

Reference site

"Area, one or more of whose element concentrations are well characterised in terms of spatial/time variability".[18]

5.0 GEOSTATISTICAL TERMINOLOGY

Many of the elaborations made by classical statistics are based on a root assumption: the data must be independent. The real situation, mainly in the soil environment, is that this is not true. A state of non-independence is the most common condition between soil samples and the effect of spatial (or temporal) fluctuation and variability has to be accounted. Geostatistics has been demonstrated in the recent past to be a useful system for estimation of concentration of elements in soil (at unsampled locations). At the same time, it represents a powerful tool suitable for tackling problems associated with models describing the distribution of elements in the soil. This technique enables the data (and the collected samples), taking into account the correlation existing between different soil samples in space (or in time) to be assessed. In Figure 3, the flow-chart of a hypothetical geostatistic process to be applied to analytical data is illustrated. A preliminary screening, to detect error(s) (human or instrumental) and mistakes, is needed to avoid negatively influencing the subsequent spatial analysis. Summary statistics allow the main characteristics of the data and their statistical distribution to be understood. Data transformation could be required in case of data not normally distributed (commonly, a log-normal distribution characterizes the environmental data). Structural analysis allows an experimental (or sample) variogram to be computed from the data. That is, the spatial correlation of the data is analysed. Since data can have some long-range (distance) trend over an area, this must be detected and, eventually, removed. Similarly, the spatial variation is probably not the same in all directions. This so-called anisotropy must be taken into account, before fitting a suitable model for the variogram. In a spatial sense it is, from the data collected, at the end possible to predict the values of a property at unsampled points (kriging) and to display the results on a map.

Soil sampling uncertainty, as a component of the total measurement uncertainty, can be described and quantified by information obtained through some geostatistical parameters combined with parameters (standard deviation and variance) from classical statistics. Sill, nugget, range, covariance (see Figure 4) are geostatistical terms that help to account properly for the spatially and non-spatially correlated components of the variance. The non-spatially correlated component includes the variance due to analytical operations, sampling, sample preparation/reduction and other unexplained sources of spatially uncorrelated variance.

5.1 Terms and definitions

Anisotropy

"A property of a variogram to have different spatial variation structures depending on direction and distance";

Covariance

" A measure of similarity between two variables defined as the expected value of their predict

[18] The definition originates from the ISO-30 "reference material" term.

minus the product of their expected values and often used to measure spatial variation between two variables as a function of the lag distance separating two locations";[19)]

Detrending

"The process of removing the trend";

Geostatistics

" Statistical methodology based on the use of spatial coordinates and able to define a model of estimation and prediction";

Kriging

" A general interpolation method applied in geostatistics for the estimation of unknown values of a variable at unsampled locations"; [20)]

Lag

"The vector that separates two locations";

Nugget

"A parameter of the variogram/semivariogram, described in term of variance, representing the spatially uncorrelated component of the variance"; [21)]

Partial sill

"The part of total sill, described in terms of variance, minus the nugget"; [22)]

Prediction

"The process of forming a statistic from observed values to predict random variables at an unsampled location;

Range

"A parameter of the variogram/semivariogram representing the distance beyond which there is little or no correlation between semivariance and lag";

Sill

"A parameter of the variogram/semivariogram describing the value that the variogram/semivariogram tends to at large distance"; [23)]

Spatial variation

"A property of samples in consequence of which samples collected near to each other are more similar than samples farther apart";

Support

"Physical size of the sample referred to a particular area, depth, and/or volume of the soil";[24)]

Trend/Drift

---

[19] The definition originates from the ASTM D5549-94 standard.

[20] Different models of Kriging exist and these are used depending on the objective of the study. To study spatial correlation between multiple variables Co-kriging is applied. Disjunctive Kriging is the interpolation model that enables estimation of the probability that at a stated point a value of a variable is below or above a threshold value. If the objective of the study is the estimation of a mean value of a variable on an aerial support, Block-Kriging is normally applied. Universal Kriging is the interpolation method used in the presence of a trend of the values of the variables observed.

[21] Nugget effect is due to different components of variance, independent to each other, and distinct from the variance linked to spatial variability. Nugget is an estimation of the spatially uncorrelated variance components, comprehensive of variance due to errors in sampling, measurement, micro-scale variabilities and other unexplained spatially uncorrelated sources of variance.

[22] The partial sill represents the variance attributable only to the spatially correlated components.

[23] The sill is described as a variance that assumes the maximum value at the distance where the variables considered are not more correlated.

[24] The concept of support is similar to unit, if this is characterised in terms of volume. For example, a support can be a unit of stated length and diameter core. Different supports can produce effects on sample distribution and statistics.

"Systematic spatial variation of the local mean of variable, expressed as a polynomial function of location coordinates"; [25)]

Variogram/Semivariogram

"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";" [26)]

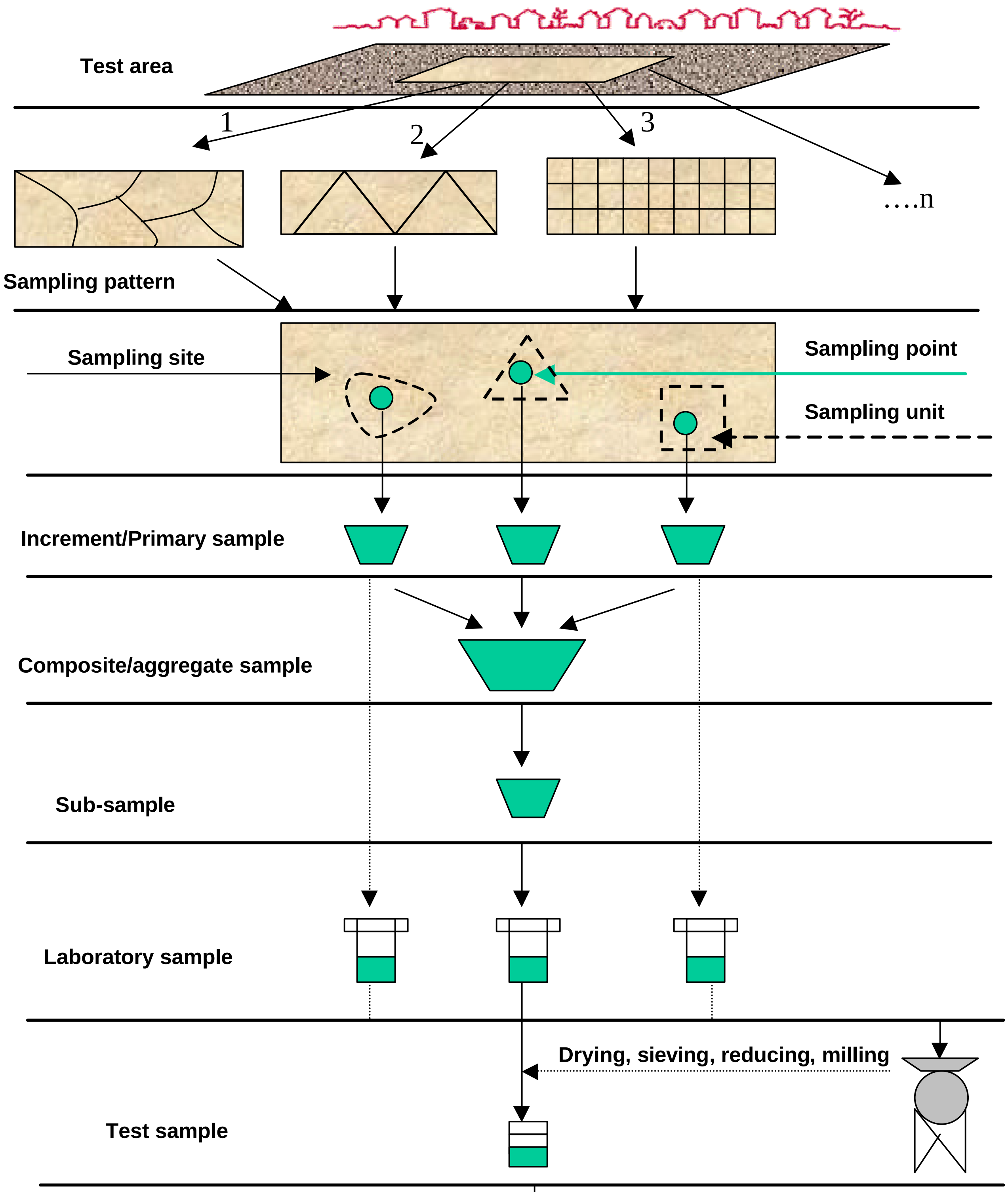
6.0 REFERENCES

1. ASTM D5549-94, 1994. Standard guide for the contents of geostatistical site investigation report. Annual Book of ASTM Standard.
2. ISO11074-2:1998. Soil quality – Vocabulary. Part 2: Terms and definitions related to sampling. International Organization for Standardization (ISO), Geneva.
3. ISO3534-1:1993. Statistics – Vocabulary. Part 1: Probability and general statistical terms. International Organization for Standardization (ISO), Geneva.
4. ISO Guide 30:1992. Terms and definitions used in connection with reference materials. International Organization for Standardization (ISO), Geneva.
5. W, Horwitz. Nomenclature for sampling in analytical chemistry (Recommendations 1990). Pure Appl. Chem 62-6, 1193-1208 (1990)

[25] The trend is a surface defined by a function of parameters of x- and y-coordinates. The definition originates from the ASTM D5549-94 standard.

[26] The variogram is mainly described by three parameters: sill, nugget and range. The definition originates from the ASTM D5549-94 standard.

Fig. 1 Relationships of sampling operations



Analytical Operations

Fig. 2 Relationships of analytical operations

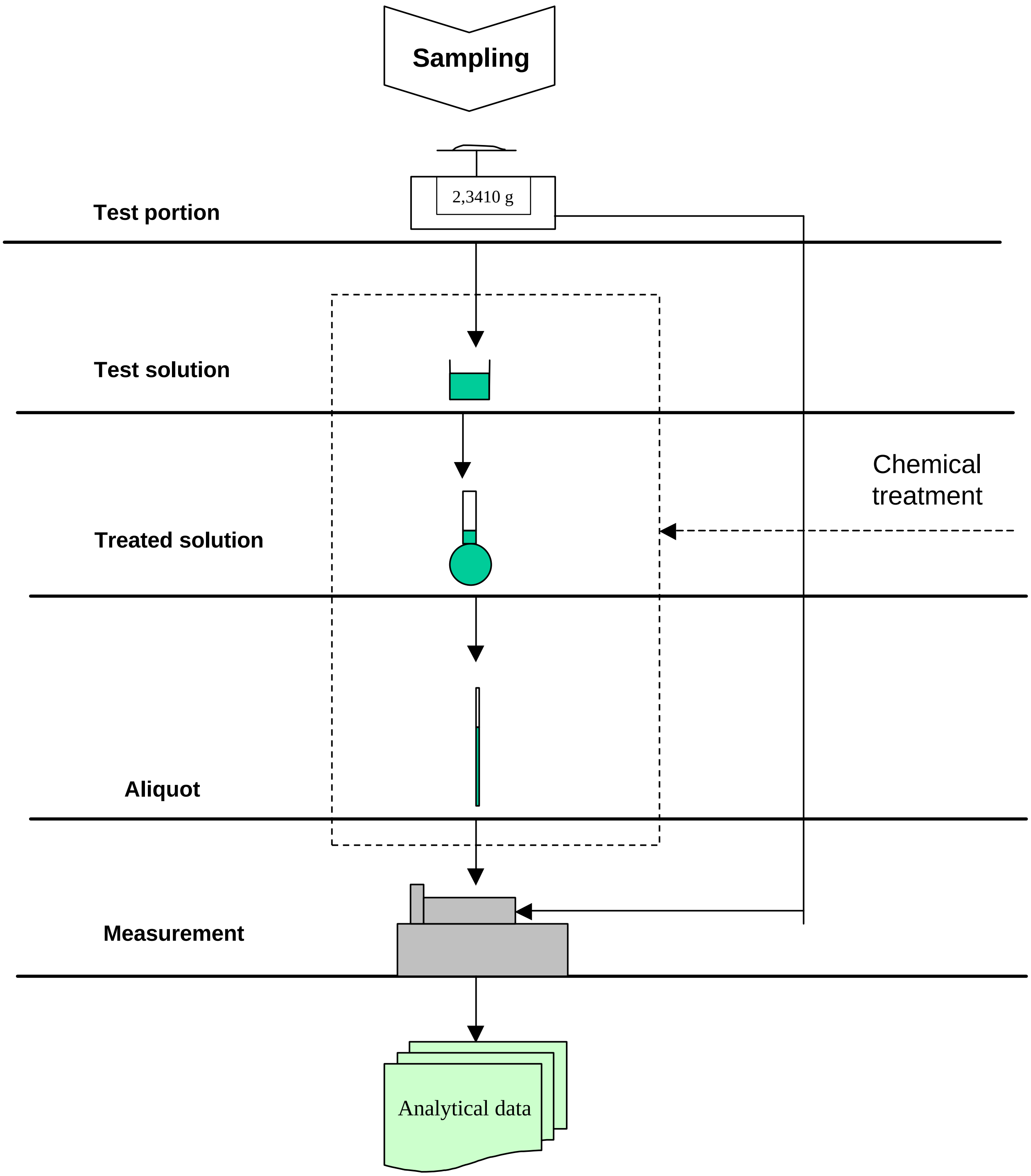


Fig. 3 Geostatistical process

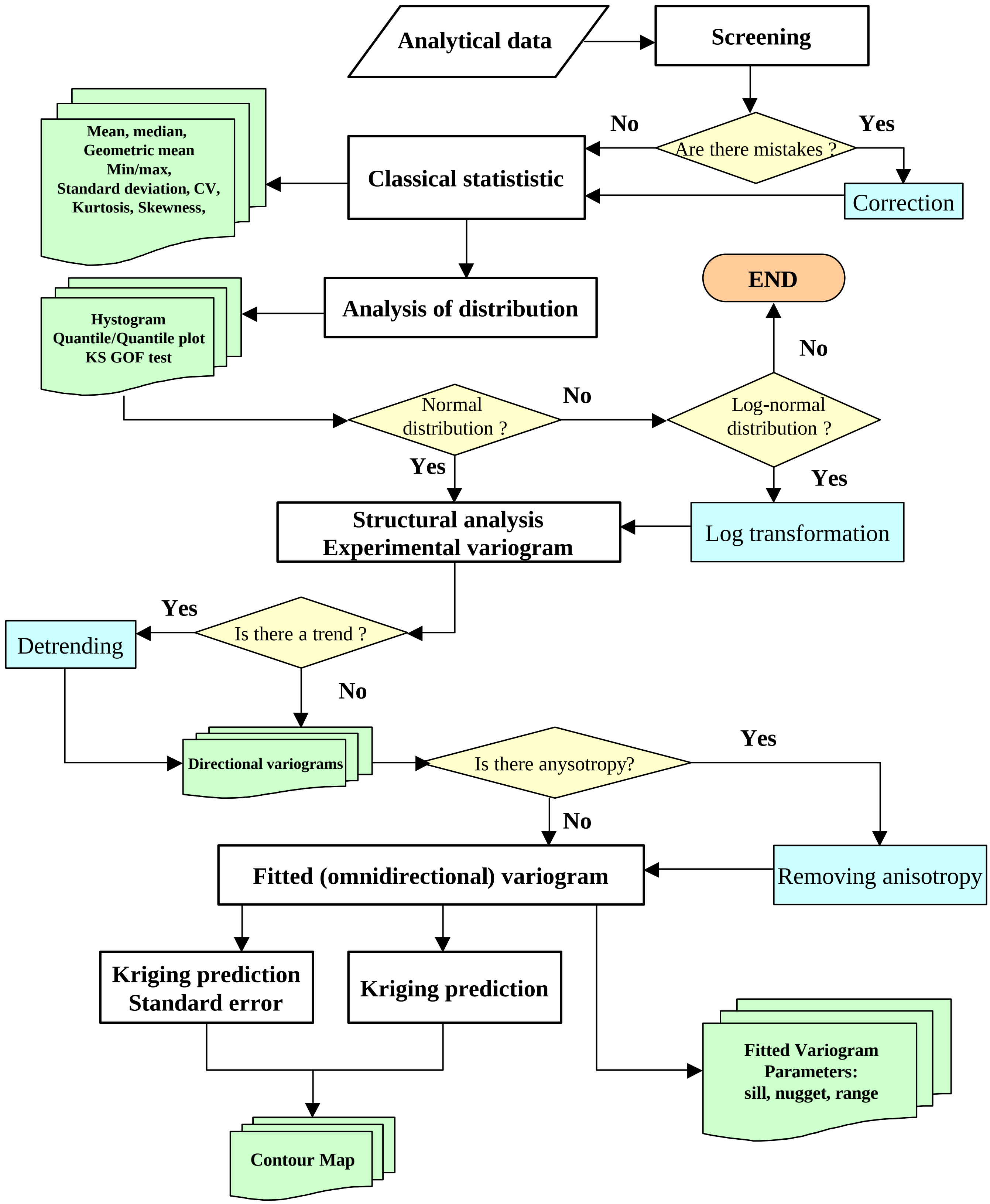


Fig.4 Variogram and main parameters