INTERNATIONAL STANDARD

ISO 19258

First edition
2005-12-15

Soil quality — Guidance on the determination of background values

Qualité du sol — Guide pour la détermination des valeurs de bruit de fond

Reference number
ISO 19258:2005(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 19258 was prepared by Technical Committee ISO/TC 190, *Soil quality*, Subcommittee SC 7, *Soil and site assessment*.

Soil quality — Guidance on the determination of background values

1 Scope

This International Standard provides guidance on the principles and main methods for the determination of pedo-geochemical background values and background values for inorganic and organic substances in soils.

This International Standard gives guidance on strategies for sampling and data processing and identifies methods for sampling and analysis.

This International Standard does not give guidance on the determination of background values for groundwater and sediments.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 10381-1, *Soil quality — Sampling — Part 1: Guidance on the design of sampling programmes*

ISO 10381-5, *Soil quality — Sampling — Part 5: Guidance on the procedure for the investigation of urban and industrial sites with regard to soil contamination*

ISO 11074:2005, *Soil quality — Vocabulary*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 11074 and the following apply.

3.1
background content
content of a substance in a soil resulting from both natural geological and pedological processes and including diffuse source inputs

3.2
background value
statistical characteristic (3.8) of the background content

3.3
contaminant
substance or agent present in the soil as a result of human activity

NOTE There is no assumption in this definition that harm results from the presence of the contaminant.

3.4
diffuse source input
input of a substance emitted from moving sources, from sources with a large area or from many sources

NOTE 1 The sources can be cars, application of substances through agricultural practices, emissions from town or region, deposition through flooding of a river.

NOTE 2 Diffuse source input usually leads to sites that are relatively uniformly contaminated. At some sites, the input conditions may nevertheless cause a higher local input such as near the source or where atmospheric deposition/rain is increased.

[ISO 11074:2005]

3.5
pedo-geochemical content
content of a substance in a soil resulting from natural geological and pedological processes, excluding any addition of human origin

NOTE It may be hardly possible to determine the precise pedo-geochemical content of certain substances in a soil due to anthropogenic diffuse contamination.

3.6
pedo-geochemical background value
statistical characteristic (3.8) of the pedo-geochemical content

NOTE Any estimate of pedo-geochemical background value will be prone to a certain amount of error given the uncertainty associated with determining the pedo-geochemical content.

3.7
soil
upper layer of the Earth's crust composed of mineral parts, organic substance, water, air and living organisms

[ISO 11074:2005]

3.8
statistical characteristic
numerical value calculated from a variate of a chosen parameter of the population

EXAMPLE Examples of the statistical characteristics are the mean, the median, the standard deviation or the percentiles of the ordered frequency distribution.

3.9
study area
three-dimensional definition of the area where samples are to be obtained from and thus for which the background value(s) are to be estimated

3.10
support
size, shape and orientation of a soil sample

NOTE For the purpose of analysing spatial variation in soils geostatistically (by estimation of the variogram of a soil property), the support should be the same at each sampling site.

3.11
variate
set of observed values of a variable

EXAMPLE A variate could for instance be the series of numbers of the concentration of a substance in soil or numerous, individual soil samples.

4 General

Soils retain the evidence of their past history including impacts due to natural events or human activities. Chemical impacts related to human activities can be detected in soils all over the world, even in regions far from any source of contamination. For this reason, the background contents of inorganic and organic substances in soils consist of a pedo-geochemical fraction and an anthropogenic fraction. The ratio of these fractions varies widely depending on the type of substances, the type of soil and land use, and the kind and extent of external impacts.

For many inorganic substances, the background content of unpolluted soils is dominated by the pedo-geochemical content and consequently by the mineralogical composition of the soils parent material. Pedogenetic processes may lead to a redistribution (enrichment/impoverishment) and consequently to a horizon-specific differentiation of the substances within a soil profile. Persistent organic substances in soils originate more often from non-natural sources and therefore the background content of soils is governed by the kind and extent of diffuse contamination from non-soil sources.

In practice, it is often difficult to distinguish clearly between the pedo-geochemical and the anthropogenic fraction of the background content of soils. Nonetheless, a detailed knowledge of the background content as well as of its natural fraction for the substances of concern is essential both for any evaluation of the current status of soils for environmental or land use related aspects or just for scientific purposes within the scope of pedology or geochemistry. To this end, so-called background values in terms of the statistical characteristics of both, the pedo-geochemical and the anthropogenic fraction have to be determined.

A variety of different objectives can be identified for the determination of background values of inorganic and/or organic substances in soils. The objectives themselves provide insufficient information to define the technical programme that will produce the desired background values. Thus a number of technical approaches have to be defined which together form the basis of the technical programme.

This guidance provides essential aspects of sampling strategies and procedures, minimum requirements regarding the necessary steps and ways of sample pre-treatment, analytical methods and statistical evaluation procedures for determining sound and comparable background values.

Guidance is given for

a) evaluating existing data from different data sources and

b) setting up complete investigation programs aiming to compile background values for a clearly defined three-dimensional picture of the soil.

These situations are representing the two extreme starting positions for the process of compiling background values. In practice, a third intermediate situation may be dealt with when additional data need to be collected because the quantity or quality of the existing data is insufficient.

5 Procedures

5.1 General

The procedures to determine background values encompass aspects of sampling (strategy, procedure), soil analysis (pre-treatment, extraction and measurement), data processing and presentation. In general, two starting positions can be distinguished, namely

a) the evaluation of existing data mostly from different data sources, and

b) the collection of new data based on an appropriate investigation strategy.

5.2 Objectives and technical approaches

5.2.1 General

Before commencing any survey on background values in soils it is of crucial importance to define the objective of the survey and the related technical approach.

The objective is, in general terms, the definition of 'why' background values are to be determined. The technical approaches describe aspects like the 'where', 'what', 'how' and 'when'. Together the technical approaches determine the technical programme that will provide the required background values.

NOTE It should be noted that a technical approach that is fit for one objective, will often be unfit for other objectives.

The objectives for defining background values might be:

- to identify the current contents of substances in soils, e.g. in the context of soil-related directives;
- to assess the degree of contamination by human activities;
- to derive reference values for soil protection;
- to define soil values for reuse of soil material and waste;
- to calculate critical levels and tolerable additional critical loads;
- to identify areas/sites with atypically enhanced levels of element contents due to geogenic reasons or human impact;
- etc.

In order to meet the objective, the technical approaches might include the following.

- Definition of the substances and parameters
  - For example, the background values to be estimated may be the total heavy metal content or the bioavailable heavy metal content. (See 5.2.2)
- Definition of the study area
  - The (three-dimensional) definition of the area where samples are to be obtained from. This has to be a detailed description of what is to be considered as the study area, and what is not. (See 5.2.3.)
- Definition of the time period of interest:
  - Are the historical or current contents relevant for the objective? (See 5.2.4.)
- Definition of the size and geometry (support) of the area sampled at a sampling location. (See 5.2.5.)

5.2.2 Substances and parameters

Background values can be determined for all kinds of inorganic and organic substances in soils as well as soil characteristics. In practice, the more persistent and immobile compounds are of primary interest because of their potential to adsorb and accumulate in soil, whereas remobilization and intrinsic biodegradation are of less significance.

As well as the substances of concern, basic soil parameters and site characteristics (see 5.4.1.3) need to be provided to assist in interpretation of the contents of substances. A number of so-called basic soil parameters influence soil processes that affect the contents of inorganic and organic substances. Table 1 lists these parameters which should be analysed according to the given International Standards.

Within the group of inorganic substances, trace elements (e.g. heavy metals, micronutrients) are most often analysed (Table 2). Concerning the analytical methods, a distinction has to be drawn between different extraction/preparation methods (Table 2), whereof very few determine the total content which may be needed for instance for calculating element stocks. Besides total contents, the (eco-) toxicologically more relevant mobile fractions (Table 2) are of increasing interest, e.g. if pathway-related questions are to be examined. Analysis of parameters in Table 2 should be carried out according to International Standards given in Table 2.

Table 1 — Basic soil parameters

Parameter	Method	ISO International Standard
Texture	Sieving, sedimentation	ISO 11277
Fraction of coarse material	Sieving	ISO 11277
Amount of non-soil material	Sieving/visual control	ISO 11259, ISO 11277
Bulk density	Direct measurement of undisturbed soil samples, estimation form soil water retention curves	ISO 11272
pH	pH-electrode	ISO 10390
Content of organic carbon	Dry combustion	ISO 14235
Cation exchange capacity, exchangeable cations	*BASECOMP* $BaCl_2$	ISO 11260 ISO 13536
Carbonate content	CO_2-evolution	ISO 10693

Table 2 — Examples for the analysis of inorganic substances

Parameter	Speciation/form	Extraction/preparation Method	ISO International Standard	
			Extraction/preparation	Determination
Metalloids, e.g. arsenic and selenium	Total	Alkaline fusion + X-ray fluorescence HF + $HClO_4$	ISO 14869-1 ISO 14869-2	ISO 14869-1 ISO 11047
Metals,	Pseudo total	aqua regia	ISO 11466	ISO 11047
barium, cadmium, chromium, cobalt,	Complexing	EDTA DTPA	 ISO 14870	 ISO 11047
copper, iron, lead, manganese, mercury, molybdenum, nickel, thallium, zinc	Exchangeable	$NaNO_3$ NH_4NO_3 $CaCl_2$ KCl		
Cyanides	Water soluble	H_2O, leaching tests	See NOTE.	See NOTE.

NOTE There are a variety of extraction and analytical methods for soil-water in the series of International Standards on water quality which may also be applicable. However, it is important to confirm that they will work with the extracts obtained form particular soil material.

Surveys on organic substances usually refer to persistent compounds. The persistent organic contaminants listed in Table 3 are some of the more commonly encountered, but the list is not complete. Analysis should be carried out according to International Standards listed in Table 3.

Various methods are used for the analysis of organic substances. The aim of these methods is usually to extract the greatest possible quantity of organic substances from soils. It is important to recognize that organic compounds may be extracted from naturally occurring organic materials (e.g. organic matter, decaying vegetation, peat, charcoal), and that non-specific analyses in particular may, therefore, give misleading results.

Table 3 — Examples for the analysis of organic substances

Substance/groups of substances	Method	ISO International Standard
PAH	Soxhlet/HPLC/UV Thin-layer chromatography RP C-18/HPLC GC/MS	ISO 13877 ISO 7981-1 ISO 7981-2 ISO 18287
Dioxins/Furane		
Chlorophenols	Hexane/GC/ECD	ISO 8165-1
Chlorpesticides	RP C-18/HPLC/UV	ISO 11369
PCBs	GC-ECD	ISO 10382
Chlornaphthalene		
Chlorparafin		
Bromodiphenylethers		
NOTE There are a variety of extraction and analytical methods for water in the series of International Standards on water quality which may also be applicable. However it is important to confirm that they will work with the extracts obtained from a particular soil material.		

When collecting new data for determining background values, it is recommended that the investigation program be designed with regard to additional questions that could arise in future. In most cases, carrying out new sampling campaigns is much more expensive than analysing additional substances in the first place. To this end, a suitable storage of soil samples for subsequent analyses of organic or inorganic substances is of crucial importance. Besides the substances of concern (Tables 2 and 3) and additional soil parameters (Table 1), it is essential to provide a comprehensive site description (see 5.4.1.3) for interpretation purposes. The documentation of all the actions taken is of utmost importance if the data measured is to be of use for other assessments in future investigations.

5.2.3 Study area

The definition of the study area (3.9) can be based on two different principles, that is:

— a purely spatial definition (X, Y, Z), defining the contours of the study area by the coordinates within which the study area lies. Apart from the definition in a horizontal plane, the soil depth that is to be studied should also be defined;

— a typological definition of the study area, based on one or more characteristic(s), e.g. soil type (for example, the A-horizon of a specific soil type), land use (also considering the potential effects on the background values), elevation level, etc.

Of course, it is possible to mix the spatial and typological definition of the study area.

EXAMPLE Examples of a mix of the spatial and typological definition of the study area might be:

— the grassland in a county or province;

— the A-horizon in an area defined by X- and Y-coordinates.

The definition of the study area must be detailed at a level where there cannot be any misinterpretation on what is and what is not part of the study area. For an unambiguous definition of the study area, all actual point and diffuse sources within the study area need to be defined. As the general objective is to determine background values, a safety zone around that (type of) source might be defined and thereby excluding parts of the more generally defined study area. It might also result in specific zones for which the data is to be considered separately from the rest of the study area.

The definition of the study area as described is independent of whether the soil samples are still to be taken, or whether already existing soil samples (or results) are to be used. In the latter situation, the detailed definition of the study area will define which samples/results are to be included or excluded.

5.2.4 Time period

Background values are influenced both by the natural processes (pedogenesis, biogeochemical cycles) as well as by diffuse source input. Two different time scales can be distinguished:

— the period in which the background value may significantly vary due to natural processes;

— the period in which the background value will most probably only change due to human influences (except for large scale natural phenomena).

The second period is generally smaller than the first one.

It might be that a specific historic period is of interest when measuring background values. When a soil layer is formed during this same period, it is indeed possible to determine background values for a certain time period.

When background values are to be re-determined after a period of time in order to determine if changes occur, the time period between measurements should be based on (see also ISO 16133):

— the expected enrichment of substances in soils (accumulation for example due to diffuse source input);

— the expected loss of substances in soils (for example, due to leaching, biodegradation or plant uptake);

— changes in concentration level that can be determined both analytically and statistically.

5.2.5 Scale of sampling (Support)

Variability in concentrations is by definition a scale-related characteristic. Depending on the volume for which an analytical result is to be considered representative, the variability in concentrations encountered might be different. The scale — or in more technical terms the (geo-statistical) support (3.10) — is therefore an important technical aspect on which a decision is to be made prior to data collection.

For (mainly) two-dimensional surveys, the support is the size (and geometry) of the area sampled at a sampling location.

The study will always involve a certain soil layer of depth. However, as in the horizontal plane, the dimensions are much larger than in the vertical plane, the support in soil surveys is most often defined in a two-dimensional way.

More information on support is given in Annex A.

5.3 Evaluation of existing data

5.3.1 General

When using existing data, specific care must be taken concerning the quality and comparability of data particularly if the data originate from different sources. Data with appropriate information have to be harmonized in a step-wise procedure with regard to the specific evaluation objectives. In general, the harmonization of data sets results in a more or less significant reduction of the respective variate. Nonetheless, the procedure of harmonization of data sets is inevitable to produce a sound and reliable evaluation. The respective harmonization strategy should encompass aspects like

a) the check of the completeness of the data sets related to minimum requirements,

b) the harmonization of different sampling strategies, references, nomenclatures and analytical procedures,

c) the identification and elimination of contaminated samples (excluded from the population of background values by definition).

5.3.2 Completeness of data sets/minimum requirements

In order to ensure a minimum level of data quality, it is essential to provide sufficient and sound information of the data, for instance

— the date of sampling,

— the procedure used to select sampling locations (plots),

— the scale of sampling (e.g. support),

— the site location (coordinates),

— the sampling depth intervals,

— the number and configuration of samples (e.g. regular grid or random sampling) taken at a sampling location (plot),

— the method used to extract and analyse the components (including quality assurance and detection limits),

— the site-specific information (e.g. pedology/lithology, land use).

This information can be used to screen the data on their suitability for the objective of compiling background values.

The definition of minimum requirements on information of the data set depends, amongst others, on the substances of concern, the area and spatial reference to be considered and the approach pursued to achieve an adequate spatial representation of the point-related data.

Apart from the information listed above, the type and degree of accuracy, e.g. of site-specific information depends on soil and other parameters influencing the behaviour and hence the contents of the substances in soils. For instance, inorganic substances need to be related at first priority to lithogenic soil properties due to their predominant geogenic origin, whereas the content of organic substances of soils is more strongly correlated to, e.g. land-use-related parameters.

5.3.3 Comparability of data (Sampling, nomenclatures, analyses)

Different sampling strategies may have a crucial impact on the comparability of data sets. Problems arise here in particular through the comparison of horizon versus depth level-related samples and that of mixed versus individual samples. Further on, the representative nature of the variate for a sample population with regard to the same support for an area needs to be taken into account. Also, an uneven spatial distribution of the sampling points within an area may cause biased estimates of the parameters of the frequency distribution due to an overestimation of some parts of the study area. Block-kriging is recommended to deal with this problem. It is strongly recommended to carefully balance the possible inaccuracies introduced by merging data sets from different campaigns, versus the advantage of an increasing number of samples and consequently an increasing representation of a population.

The extent to which different sample pre-treatments and analytical procedures (extraction, measurements) can be compared and harmonized has to be evaluated in each individual case, e.g. against the intended accuracy of the background value. For inorganic substances, the analytical results originating from different analytical procedures may be transformed by applying regression functions or constants provided the respective relations are known. Alternatively, the analytical procedures may be grouped roughly according to the operationally defined extracted fractions (see Table 2). The broader the ranges of classified background values as target variables are, the lower may be the demand of data comparability. Nonetheless, the assessor should bear in mind, that merging data sets analysed by different analytical procedures invariably requires compromises to be made.

5.3.4 Elimination of outliers

According to definition 3.1, the background content of substances in soils includes the moderate diffuse input into the soil. Therefore, locally contaminated sites are excluded from the population of background contents. Consequently, data obviously stemming from locally contaminated sites have to be identified and eliminated from the respective data set. To this end several statistical tests for identifying outliers are applicable, e.g. test on distribution of the data, exploratory data analysis (boxplots), principle component analysis, etc. (See also 5.5.1.2.1 and Annex B.)

NOTE The removal of outliers has a significant effect on the resulting definition of the background value. The statistical identification of an outlier by itself is insufficient for removing a high (or low) measurement out of the database describing the background value. The statistical test does however provide a good method of defining which data should be investigated in more detail, in order to see if an explanation can be found for the high value to be an outlier. If such an explanation is found, the value is indeed an outlier and should be eliminated.

5.4 Collection of new data

5.4.1 Sampling

5.4.1.1 Sampling strategy

5.4.1.1.1 General considerations

The natural pedo-geochemical content and the usual content of substances in soil vary according to soil parent material. They also depend on soil horizons, as pedogenic processes modify and redistribute components in soils, leading generally to the formation of several soil horizons that may exhibit different compositions.

Land use and distance to contamination sources also influence the content of substances in soils. Human activity modifies soil composition through agriculture, waste spreading, building, atmospheric deposition from industry, households, traffic, etc. A sampling site is considered here as a small portion of land, from a few square metres to about 1 ha, where one sample of each of the soil layers or horizons of interest is collected.

This section presents two strategies for selecting the sampling sites within the study area: the systematic approach and the typological approach. The choice of one of them is generally influenced by the degree of pre-existing knowledge about the soil and land use. When relatively little is known, the systematic approach is often more appropriate. However, these two approaches can be considered as typical ones in the continuum of all possible strategies. Therefore, it is possible to build an intermediate strategy, mixing some aspects of the systematic approach with others from the typological one.

5.4.1.1.2 Systematic approach

The sampling sites are located using a grid. The interval between the grid points is dependent on the resolution desired for the determination of the pedo-geochemical and/or background content. In principle, the interval between the sampling sites should be such that the minimum number of samples can be collected to represent each of the defined soil units. A square grid can be used, with cells varying in size (available monitoring recommendations should be considered).

For instance, square cells with a 16 km, 5 km or 2,5 km site can be used at the scale of a country, whilst square cells of a few hundred meters are more appropriate for the study at the level of a small area.

If sampling at a given grid point is rendered impossible due to buildings, roads, water surface or any other reason, a new location may be chosen using a systematic procedure. For instance, a deviation may be permitted from the initial point by steps of a definite distance north, then east, then south and finally west.

For each selected site, consider moving the sampling area if it is potentially highly contaminated by near-by point sources, or in a pedo-geochemical way, if any source could compromise the purpose of the study (e.g. overhead power-lines should be avoided if the zinc content of soils is of interest).

Samples are taken from soil layers of definite depths or from a defined pedological horizon or horizons. If the surface layers are contaminated by diffuse sources, the contents determined indicate the background content in these soils. For relatively immobile substances (e.g. heavy metals), the deep layers and particularly those at a depth below 40 cm are generally uncontaminated (provided local contamination by point sources can be excluded), and the respective substance content can be considered as the pedo-geochemical content.

A comprehensive site and soil description (Table 5) should be done at the same time as the samples are taken.

5.4.1.1.3 Typological approach

In the typological approach, the soil is stratified according to soil parent material (for inorganic substances), soil type and land use. Sites potentially highly contaminated by adjacent point sources are rejected.

The typological approach needs detailed information about the area to be investigated. Information (such as on geology, pedology, land use and sources of possible contamination) has to be gathered and evaluated in order to elaborate the sampling scheme.

For inorganic substances, the first step of the stratification of the area refers to the soil parent materials. Within each type of parent material stratum, the soil is stratified again on the basis of pedogenesis, if this is considered to have markedly modified the distribution of substances in soil. Further stratification, e.g. for organic substances is related to land use; it is recommended to distinguish between cultivated and forest soils and soil under meadow or spontaneous vegetation. On a local scale, the best stratification is based on pedology, e.g. that of the soil series, as this taxonomic level generally explains most of the variation of the soil properties. Finally, the horizon to be sampled has to be chosen.

Within one stratum, the sampling sites should generally be chosen in such a way that the area is covered representatively. The choice of the sampling sites can be carried out within each stratum using a random or systematic sampling scheme.

NOTE Apart from the degree of pre-existing knowledge, the type of question largely determines the choice between a random or a systematic sampling scheme. To estimate parameters of a frequency distribution of the background contents, a random sampling approach is most appropriate. For mapping background contents, in general, a (centred) regular grid is more appropriate.

5.4.1.2 Number of sampling sites

Background values cannot be summarized in a central parameter such as the mean. It is necessary to describe the variability of a given content in soil as precisely as possible. In the case of a normal probability distribution, the number of samples necessary for the estimation of the standard deviation is independent from the standard deviation of the population. It can be determined using Table 4, which shows that a minimum number of 30 samples is necessary to estimate the standard deviation of a normal population.

Table 4 — Maximum relative error on the estimation of the standard deviation of a normal population, with a = 0,05; n: number of samples

n	e_r (%)
10	57
20	35
30	27
40	23
50	21

However, probability distributions of substances in soils are rarely normal. They are often positively skewed but not necessarily log-normal. The estimation of the required number of samples to assess variability of such a distribution can then be equal to the number of samples necessary to draw a representative histogram or to calculate representative percentile. To this end, a minimum number of 30 samples is recommended.

5.4.1.3 Soil description

The interpretation of background contents of soils requires general information about the study area. The most relevant parameters for the soil description of the study area are listed in Table 5. It is important to bear in mind that the reliability of data interpretation strongly depends on a profound knowledge of the study area, hence collection of parameters, indicated in Table 5, should be as comprehensive as possible.

Table 5 — Parameters for site and soil description

Parameter		ISO International Standard
Landform and topography	Topography, landform, land element, position coordinates, slope microtopography	ISO 11259
Land use and vegetation	Land use, human influence, vegetation	ISO 11259
Geology and lithology	Kind of parent material, effective soil depth	ISO 11259
Surface characteristics	Rock outcrops, surface coarse fragments, erosion phenomena, surface sealing, surface cracks, other characteristics	ISO 11259
Soil-water-relationship	Surface water balance, rainfall, evapotranspiration, groundwater recharge, presence and depth of water table, site drainage, moisture conditions	ISO 11259
Soil type/soil profile description	Soil unit in regards of the classification system used	ISO 11259
	Sequence and depth of diagnostic horizons, kind of boudaries	
	Soil colour (matrix, mottling)	
	Organic matter	
	Texture, coarse elements, presence of non-soil material pedofeatures	
	Carbonates, field-pH, electrical conductivity	
	Structure, voids, fracturing, inhomogeneties	
	Compactness and consistence	
	Total estimated porosity	
	Roots, worm channels, biological activity	

5.4.1.4 Sampling depth

Background contents and values vary with soil depth due to pedogenesis, use and type and source of contamination. Sampling can be carried out on a fixed depth basis (i.e. layer) or according to definite horizon types. Sampling according to the depth is easier as it does not need the identification of the horizon type. But it will give a less precise measurement of the background contents, as it does not enable control of the variability due to horizon differentiation and use.

As anthropogenic contamination mainly enters soil at the surface, the concentration measured in the upper layers or horizons is regarded as a background content for those substances which are introduced in soil as a consequence of human activity. The determination of these substances in the deep layers or horizons gives an estimation of their pedo-geochemical content. For the substances that are not introduced in soils by human activity, the analysis of any of the layers or horizons of this soil gives an estimation of the pedo-geochemical content of this layer or horizon.

Note that in some cases, a contamination may also enter the soil from underground (e.g. by contaminated ground water).

5.4.1.5 Sampling period

Sampling should be spread over a period of time chosen to limit the temporal variation of the background contents.

When soil parameters do not vary within one year, which is the common case for most soil substances, sampling can be carried out at any period of the year.

Practical aspects, mainly concerning the access to the soil horizon or layer, have to be considered before choosing the sampling period. For instance, it may be difficult to sample deep horizons during a wet season, due to the presence of a water table close to the surface. On the contrary, sampling during the dry season may be rendered difficult by drought, which makes soil hard to penetrate with the sampling tools. Access to cultivated plots may be difficult because of growing crops. In this case, it is advisable to sample shortly after harvest and/or immediately after sowing.

5.4.1.6 Sampling technique

Sampling should be performed in accordance with ISO 10381-1 and ISO 10381-5. The following recommendations may generally be taken into account. The sampling techniques depend on the depth or horizon that has to be sampled. If only the surface horizon or layer has to be sampled, a hand corer or equivalent tool can be used.

If deep horizons or layers have to be represented, a powered corer would be preferable. All sampling tools must be designed and/or used in order to avoid cross contamination among layers or horizons.

Sampling in a soil pit is generally recommended because it provides a clear distinction of the soil horizons and other soil characteristics. It allows the soil description to be carried out at the same time of sampling for analysis. Sampling should then be performed from the bottom to the top in order to avoid cross contamination by soil material falling from the upper horizons on the lower ones.

Mixing several cores taken on the site area, according to a systematic or a random design can make composite samples of each of the layers or horizons. When sampling in a soil pit, it is recommended to clear a sufficient surface of the sampled horizon in order to take several cores to be mixed together.

All of the materials used for sampling, transport, labelling and storage of the samples must not release significant amounts of the elements or substances that are to be determined.

5.4.2 Soil analysis

5.4.2.1 General considerations

Two sets of parameters can be measured on the samples. The first one is made of the concentrations of the substances of interest. These may be, for instance, trace elements, major elements or organic compounds (see Table 2 and Table 3). The second set of basic soil parameters as listed in Table 1 should be available in order to interpret the substances of interest including their variability in soils.

5.4.2.2 Pre-treatment

Sample pre-treatment should be carried out in accordance with the standard dealing with this step of analysis, as well as with the requirements of the analytical methods.

5.4.2.3 Analysis

Analysis should be performed in accordance with the relevant standards for analytical methods, if available (see Table 1, Table 2 and Table 3).

If possible, analytical techniques should have detection limits significantly below the lowest natural pedo-geochemical content of the soil under consideration.

It is recommended that the analysis should be performed under a quality control procedure (see Clause 6).

5.4.2.4 Storage of the samples

It may be useful to store aliquots of the soil samples in order to analyse them for the determination of other parameters, which could become of interest several years after the sampling campaign. When looking for temporal trends, in the case of repetitive investigations, stored aliquots of soil samples from former campaigns are important in order to judge on any drift in analytical methodologies or to allow for new techniques. The dried samples should be stored in sealed containers that do not release any substance of interest. The containers should be placed in a room with low humidity and sheltered from dust, light and large temperature variations. Such storage does not prevent soil material from changing by chemical evolution. However, it should not significantly modify the total content of non-volatile element and persistent organic compounds.

NOTE The total content of organic compounds may get modified significantly depending on specific soil and substance properties (e.g. volatility, persistence). This effect can be minimized when storing soil samples at about −140 °C.

5.5 Data processing and presentation

5.5.1 Statistical evaluation of data

5.5.1.1 General considerations

The pedo-geochemical or background content can be regarded as a statistical population. The objective of the data processing is to represent and characterize this population, using a variate of n individual values. Therefore, the data processing is done in order to

a) select the data corresponding to pedo-geochemical or background content from the whole data set (e.g. test for outliers), and

b) characterize the distribution in such a way that a frequency or probability can be associated, at least approximately, to any of the selected values.

A systematic approach, if conducted in a heterogeneous area with little knowledge about the soils, needs precise data processing in order to select the different population samples and to define *a posteriori* different populations of background contents. On the other hand, results from a typological approach based on soil categories excluding highly contaminated sites, need less processing than those from a systematic approach, as the *a priori* defined strata are considered homogeneous and not contaminated by close point sources.

The following subclauses propose examples for selecting representative data and to characterize pedo-geochemical and background contents. The examples are applicable to both starting positions, e.g. evaluating existing data or collecting new data. They are indicative and other techniques can be found in the specialized literature.

5.5.1.2 Selection of representative data

5.5.1.2.1 Systematic approach

The histogram is a powerful tool to visualize the frequency distribution of the data. It gives information on the skewness and the kurtosis of the distribution, as well as on its homogeneity. It assists in distinguishing the possible different populations that comprise the total data set and, eventually, to separate them before further processing.

For soil parameters, homogeneous populations are mono-modal. They are rarely normal and often present a positive skewness that can rarely be modelled by a log-normal function. Therefore, the statistical methods to detect outlying values based on the assumption of normality of the population must be used with great care, after checking the reliability of this assumption.

It is generally preferable to use a method selecting the data that does not rely on assumptions about normality of the population. Exploratory data analysis (EDA) is an example of a technique that relies solely on the inherent structure of the data. Examples are given in Annex B.

An efficient way of selecting homogeneous statistical samples for pedo-geochemical or background content is to gather values measured on objects supposed to belong to the same category or stratum. This can be gained by selecting and gathering the data according to stratification by land use, distance to point sources, soil parent material, soil and/or horizon characteristics. This can be done to a certain extent from a sampling strategy based on a systematic approach if sufficient information is collected when sampling in order to define *a posteriori* different homogeneous strata, and also if sufficient data are available to represent each stratum. For instance, samples can be gathered according to soil type, texture class, pH, organic matter content or total iron content, etc. Then, the selection of the data of each *a posteriori* stratum can be carried out as described before.

5.5.1.2.2 Typological approach

The typological approach provides a statistical sample for each investigated stratum. As these strata are *a priori* defined and generally homogeneous, the number of outlying values in each stratum should be nil. In other words, once a stratum is precisely defined, all the concentrations measured on it are supposed to represent its composition and none of them should be eliminated. However, there could be outlying values in a data set representing a stratum, for instance because of sampling in the wrong stratum, the existence of an unexpected contamination, an analytical mistake, etc.

It is therefore necessary to check the homogeneity of each statistical sample representing each stratum. This can be done using histogram, EDA or principal component analysis, as described for the systematic approach (see also Annex B). A simple way to detect anomalous contents is to plot the substance content against a soil characteristic known to affect the distribution of the considered substance, like clay content, cation exchange capacity, total iron content, etc. An outlying value will locate away from the cloud formed by the rest of the values.

5.5.1.3 Distribution of the population of background values

Once the statistical sample is established, the distribution of the population can be summarized by various parameters. In the case of normal distribution, an estimation of arithmetic mean can be calculated, together with an estimated variance or standard deviation.

As probability distributions are frequently not normal or log-normal, it is recommended to use percentiles as background values. For instance, the 10th, 25th, 50th, 75th, 90th percentiles give a synthetic appreciation of an observed distribution.

A percentile can be interpreted in the following manner. The probability for a sample from the total population to show a concentration lower than the x^{th} percentile is x %.

The extreme values, the maximum and the minimum are generally not representative of the most probable values that can be encountered in a definite soil.

5.5.2 Data presentation and reporting

Background values can be presented in various manners, for instance in tables, boxplots or histograms.

Background values can be represented on a map of the investigated area. Values obtained from a typological approach can be reported on a soil map or on a map of soil parent material. The location of the sampling sites also enables the data from a systematic approach to be mapped. Various techniques exist to draw maps from spatial data, which can be found in specialized literature.

It is important to document all steps of the investigation including the results in a comprehensive report. An example of the structure and the main contents of such a report is given in ISO 10381-5.

Often, the reports presenting the results of data analyses and compilation of background values will be scrutinized by regulators and other interested parties including the general public. It is important, therefore, that such reports are of a high technical standard but also take account of the diverse and often non-technical readership. Use should, therefore, be made of tabular summaries, graphical and other means to present the data in ways that will make the data and conclusions as easy as is practicable to assimilate and assess.

6 Data handling/quality control

This International Standard provides guidance on the types of data that are required for compiling background values and indicates for which parameters or procedures there are International Standards available. The assessor must choose those parameters that are appropriate to the task in hand.

Before any compilation of background values can be made, the sufficiency of data to be used must be evaluated. The data must be sufficient in terms of

— type,

— quantity, and

— analytical/testing quality.

In the context of data quality, it is essential to

— define the objectives of the survey,

— establish a sampling strategy in terms of types of samples to be obtained, sampling locations, and how samples are to be handled consistent with these objectives,

— establish an analytical and testing strategy taking into account the guidance in this and other relevant International Standards,

— set data quality objectives consistent with the assessment procedure to be used.

It is essential to have sufficient and reliable data. The confidence that can be attached to any judgements made, for example through comparison with the requirements of a published standard (the requirements in such standards regarding sampling must always be followed), is no greater than the confidence there is in the representativeness of the data.

NOTE The assessor will need to bear in mind the disproportionate costs and time delays that may result if it is necessary to carry out an additional sampling exercise if, for example, a particular parameter is not determined when the opportunity is available (e.g. in order to reduce the cost of the investigation).

The quality of the data to be used may be assured by:

— setting formal data quality objectives (e.g. for accuracy, reproducibility, etc);

— using standardized analytical and testing methods such as those listed in this International Standard, or where International Standard methods are not available, those published by national standardization or equivalent bodies;

— using laboratories applying methods accredited under ISO/IEC 17025;

— using laboratories accredited under ISO 9001 (for example);

- using laboratories that take part in relevant proficiency testing schemes of international level. If the analyses are performed by different laboratories, it should be made sure that they participate to the same proficiency test in order to control precision and accuracy, and thus quality and comparability of the results;
- the commissioning agent employing its own quality assurance procedures.

It should be stressed that the use of any or all of the given criteria does not exclude the possible presence of excessive analytical variability (so called “analytical errors”). These may be present for individual samples, for several samples, for one or several of the parameters analysed.

Annex A
(informative)

Scale of sampling

In the sampling of granular materials like soil, at least three different scales can be identified, as illustrated in Figure A.1.

The first level of information is the scale of the individual increments or samples (the data scale). It follows that this is the smallest scale on which information on the sampled soil is available.

Most often, the data scale is used purely as a means of defining the quality of a larger amount of material in which the sampling points are situated. This gross volume might be the volume/area that is sampled resulting in one mixed sample. This is defined as the object scale.

In specific situations there can even be a third level of information, the random test scale, when the study area is investigated by means of a random test where more 'units' are investigated at the object scale.

EXAMPLE 1 A study area of 100 000 m^2 has to be investigated.

For the investigation, increments of 1 kg of soil are taken. As for legislation, a soil surface of 1 000 m^2 is considered important, the increments are taken distributed within five individual volumes of 1 000 m^2 randomly chosen within the study area.

In this example, the data scale is 1 kg, the object scale is 1 000 m^2 and the random test scale is 100 000 m^2.

EXAMPLE 2 A study area of 100 000 m^2 has to be investigated.

For the investigation, samples of 0,4 kg of soil are taken. The samples are randomly distributed over the study area.

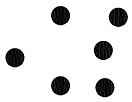
In this example, the data scale is 0,4 kg, and the object scale is 100 000 m^2. There is no random test scale in this example.

The amount of (spatial) variability cannot be defined without definition of the scale (support) on which that degree of variability occurs. It is therefore of vital importance that the scale on which the spatial variability is encountered is mentioned simultaneously when talking about spatial variability.

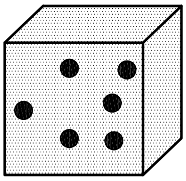
When sampling a particulate material, like soil, the final analytical result will always be an estimate of the mean concentration of a certain soil volume. Based on these results, other statistical parameters might be determined, but it is relevant to realize that the underlying data are always estimates of the mean concentration for a certain soil volume (support).

NOTE When an analytical sample of 1 g is taken and analysed completely, the resulting value will be the true mean for the different soil particles that were present in that analytical sample.

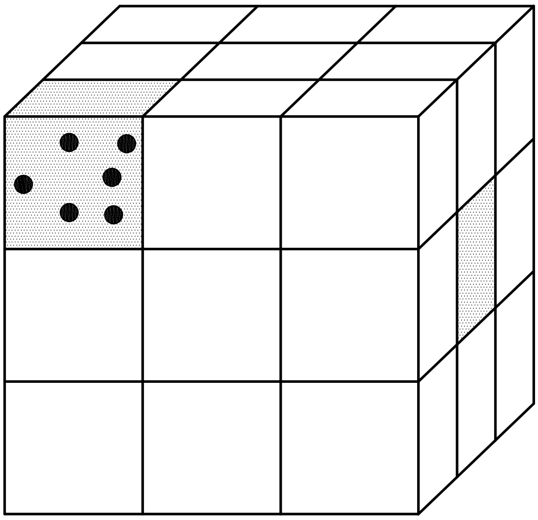
When the support in a study to determine the background values is defined as 1 ha, a sampling strategy has to be chosen that will result in an expected representative sample for this soil volume (support). One cannot derive information on a smaller scale from these samples, but as the sample support has been chosen after due consideration, that should be no problem; one is not interested in the variability on a smaller scale than 1 ha.



a) Net sampled soil volume — Data scale



b) Gross sampled soil volume — Object scale



c) Random test — Random test scale

Figure A.1 — Conceptual picture of the different scales relevant for sampling soil

Annex B
(informative)

Outlier tests

The removal of outliers has a significant effect on the resulting definition of the background value. The performance of an outlier test is therefore an essential step in data processing/data analysis.

The statistical identification of an outlier by itself is insufficient for removing a high (or low) measurement from the database describing the background value. The statistical test does however provide a good method of defining which data should be investigated in more detail in order to see if an explanation can be found for the high value to be an outlier. If such an explanation is found, the value is indeed an outlier.

There are different reasons why outliers can be present in the database, including:

- an administrative error;
- a measurement error that is confirmed to be an error based on new analysis;
- errors in the procedure resulting in the measurement; thus resulting from sampling, sample pre-treatment, destruction or extraction and analysis.

If an outlier is explained by one or more of these errors, the measurement should be discarded from the database completely.

Other explanations for the occurrence of outliers are

- a difference in the (historical) soil use,
- a local case of soil pollution, and
- a difference in soil type.

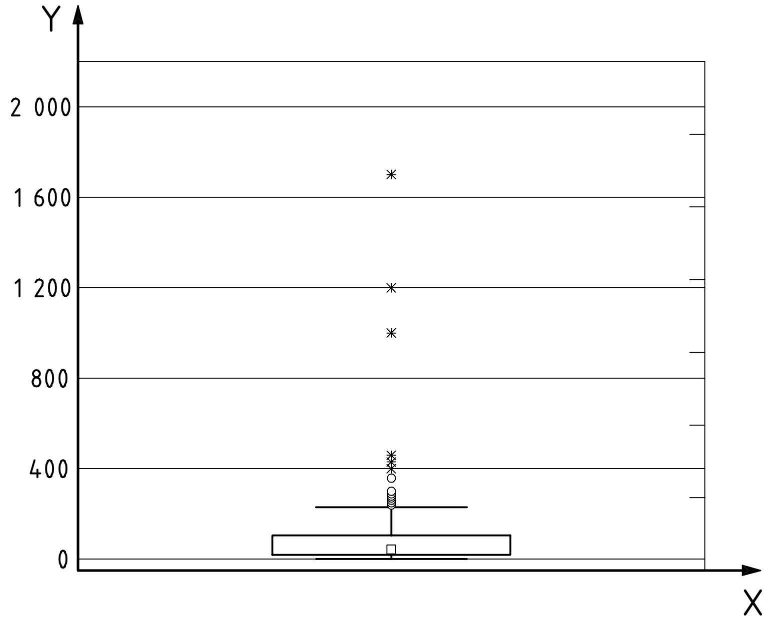
In all three cases, the outlier should not be included in the (part of the) database describing the background soil quality, but should be kept in the overall database as these high values do describe the local soil quality.

For the detection of outliers, numerous tests can be applied. Per definition, it is necessary to assume an underlying distribution as only the fact that a measurement does fall outside the expected width of that distribution, makes the value an outlier.

In order to detect outliers in a simple and visualised manner, often the 'box and whisker plot' is used (often called only 'box plot'). The box and whisker plot graphically displays statistical information on the distribution of the data. Different types of box and whisker plots exist. Often a box and whisker plot is used where different percentile values of the distribution are displayed. The 'box' describes the part of the distribution between the 25th percentile and 75th percentile of the distribution, thus displaying the 'central' 50 % of the data. Within the box, the median is displayed. The 'whiskers' run from the 25th percentile and 75th percentile respectively up to the value equal to k times the inter-quartile range (the difference between the 25th percentile and 75th percentile). k often equals 1,5. Values larger than the whiskers are detected as outliers. Especially the value (here 1,5) used for positioning the whiskers, is based on an assumption of the underlying distribution of the measurements.

Figure B.1 shows a box and whisker plot for lead concentrations found in the study area. The outliers in the figure are divided into two groups, the 'outliers' and 'extremes', where the extremes represent measurements over 2 times the inter-quartile range. Other factors for the distinction between outliers and extremes (like a value of 3) are also used.

As shown by Figure B.1 a number of outliers are detected. However, it should be investigated if these values are truly outliers.



Key

Y is the mass fraction of Pb expressed in milligrams per kilogram, mg/kg

non-outlier
non-outlier min

75 %
25 %

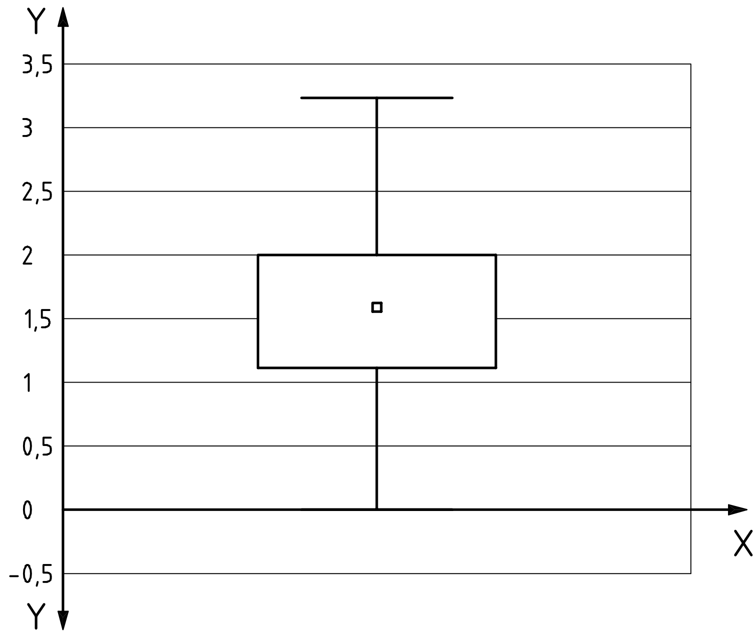
median

outliers

extremes

Figure B.1 — Lead concentrations (original data) in a box plot graphic

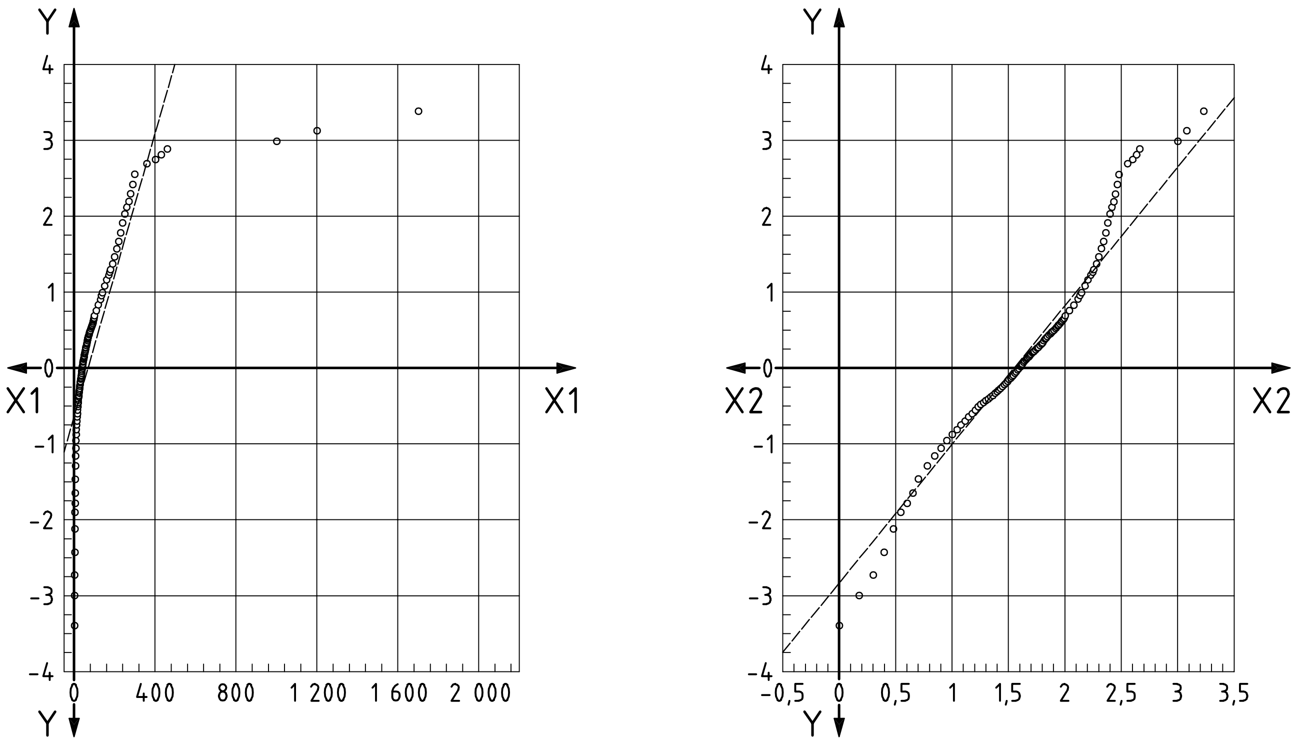
Figure B.2 displays the same lead measurements, but after logarithmic transformation. Now it appears that there are no outliers present. This is also supported by the 'normal probability' plots for the measurements as well as the log-transformed measurements as shown in Figure B.3.



Key

Y is the mass fraction of Pb expressed logarithmically in milligrams per kilogram, log mg/kg

Figure B.2 — Lead concentrations (log transformed data) in a box plot graphic



Key

Y is the expected normal value

X1 is the observed Pb value expressed in milligrams per kilogram, mg/kg

X2 is the observed Pb value expressed logarithmically in milligrams per kilogram, log mg/kg

Figure B.3 — Normal probability plots for the contents (left) and log transformed lead contents

This clarifies that the measurements are positively skewed as appears in most environmental data. After logarithmic transformation the log-transformed data fit the expected values (displayed as the line in Figure B.3) reasonably well. This does not imply that no explanation for the outliers has to be sought, but it can provide part of the explanation why the measurements do belong in the database describing the background value.

The statistical outlier test simply identifies what data should be investigated in more detail. It can however also provide ‘proof’ for the fact that the data are high, but do describe a part of the distribution of background values. In addition to plotting the data according to an expected distribution, as shown in the example, this exploratory data analysis (EDA) can also exist out of the spatial display of the measurements in order to visualize spatial patterns. For example, spatial relations can be present due to specific forms of land use or spatial structures in the study area.

Correlation matrices displaying the correlation between the investigated components can also be useful in explaining and understanding the available measurements.

Another approach to select the values corresponding to pedo-geochemical or background content is to apply principal component analysis. Plotting the scores of the samples against the three principal components of the total variance generally shows the differences between the populations.

An alternative for the detection of outliers is to use specific statistical outlier tests. These tests however are less simply to interpret (visually) and it is essential that the tested data meet the conditions for applying the test (like normality of the data). If these conditions are not met, measurements can be falsely detected as outliers.

Bibliography

[1] ISO/DIS 7981-1, *Water quality — Determination of polycyclic aromatic hydrocarbons (PAH) — Part 1: Determination of six PAH by high-performance thin-layer chromatography with fluorescence detection after liquid-liquid extraction*

[2] ISO/DIS 7981-2, *Water quality — Determination of polycyclic aromatic hydrocarbons (PAH) — Part 2: Determination of six PAH by high-performance liquid chromatography with fluorescence detection after liquid-liquid extraction*

[3] ISO 8165-1, *Water quality — Determination of selected monovalent phenols — Part 1: Gas-chromatographic method after enrichment by extraction*

[4] ISO 9000, *Quality management systems — Fundamentals and vocabulary*

[5] ISO 9001, *Quality management systems — Requirements*

[6] ISO 10382, *Soil quality — Determination of organochlorine pesticides and polychlorinated biphenyls — Gas-chromatographic method with electron capture detection*

[7] ISO 10390, *Soil quality — Determination of pH*

[8] ISO 10693, *Soil quality — Determination of carbonate content — Volumetric method*

[9] ISO 11047, *Soil quality — Determination of cadmium, chromium, cobalt, copper, lead, manganese, nickel and zinc — Flame and electrothermal atomic absorption spectrometric methods*

[10] ISO 11259, *Soil quality — Simplified soil description*

[11] ISO 11260, *Soil quality — Determination of effective cation exchange capacity and base saturation level using barium chloride solution*

[12] ISO 11271, *Soil quality — Determination of redox voltage — Field method*

[13] ISO 11272, *Soil quality — Determination of dry bulk density*

[14] ISO 11277, *Soil quality — Determination of particle size distribution in mineral soil material — Method by sieving and sedimentation*

[15] ISO 11369, *Water quality — Determination of selected plant treatment agents — Method using high performance liquid chromatography with UV detection after solid-liquid extraction*

[16] ISO 11466, *Soil quality — Extraction of trace elements soluble in aqua regia*

[17] ISO 13536, *Soil quality — Determination of the potential cation exchange capacity and exchangeable cations using barium chloride solution buffered at pH = 8,1*

[18] ISO 13877, *Soil quality — Determination of polynuclear aromatic hydrocarbons — Method using high-performance liquid chromatography*

[19] ISO 14235, *Soil quality — Determination of organic carbon by sulfochromic oxidation*

[20] ISO 14869-1, *Soil quality — Dissolution for the determination of total element content — Part 1: Dissolution with hydrofluoric and perchloric acids*

[21] ISO 14869-2, *Soil quality — Dissolution for the determination of total element content — Part 2: Dissolution by alkaline fusion*

[22] ISO 14870, *Soil quality — Extraction of trace elements by buffered DTPA solution*

[23] ISO 16133, *Soil quality — Guidance on the establishment and maintenance of monitoring programmes*

[24] ISO 18287, *Soil quality — Determination of polycyclic aromatic hydrocarbons (PAH) — Gas-chronomatographic method with mass spectrometric detection (GC-MS)*

[25] ISO/IEC 17025, *General requirements for the competence of testing and calibration laboratories*

ICS 13.080.99

Price based on 24 pages