INTERNATIONAL STANDARD

ISO 19258

Second edition
2018-08

Soil quality — Guidance on the determination of background values

Qualité du sol — Recommandations pour la détermination des valeurs de fond

Reference number
ISO 19258:2018(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 190, *Soil quality*, Subcommittee SC 7, *Impact assessment*.

This second edition cancels and replaces the first edition (ISO 19258:2005), which has been technically revised. The main changes compared to the previous edition are as follows:

— Clauses 2 and 3, and subclauses 5.3, 5.4, 5.4 and Annex B (formerly Annex C) have been completely technically revised;

— 5.2.2 has been revised and the structure of its subclauses has been changed to 5.2.2.1, *Basic parameters*, 5.2.2.2, *Persistent compounds* (split up into 5.2.2.2.1, *Inorganic substances*, and 5.2.2.2.2, *Organic substances*), and 5.2.2.3 *Non persistent compounds* (added);

— text has been added to 5.2.5;

— "typological" has been replaced by "judgemental" throughout the document;

— "scale of sampling" has been deleted from Annex A;

— the Bibliography has been updated.

Soil quality — Guidance on the determination of background values

1 Scope

This document gives guidelines for the principles and main methods for the determination of background values for inorganic and organic substances in soils at a local/regional scale. The site scale is excluded.

It gives guidelines for sampling and data processing strategies. It identifies methods for sampling and analysis.

This document does not apply to the determination of background values for groundwater and sediments.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 11074, *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.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
background concentration
concentration of an element or a substance characteristic of a soil type in an area or region arising from both natural sources and anthropogenic diffuse sources such as atmospheric deposition

[SOURCE: ISO 11074:2015, 3.5.1, modified — In the definition, "an element or" has been introduced before "a substance" and "anthropogenic" has replaced "non-natural". Note 1 to entry has been removed.]

3.2
background value
statistical characteristics (3.8) of the total (natural pedo-geochemical and anthropogenic) content of a substances in soil

Note 1 to entry: Commonly expressed in terms of average, typical, median, mode, a range of values or a background value.

[SOURCE: ISO 11074:2015, 3.5.2, modified — Note 1 to entry has been added from ISO 11075:2014, 3.5.1.]

3.3
diffuse source input
input of a substance emitted from moving sources, from sources with a large area or from many sources

Note 1 to entry: In practice, two situations are commonly recognized: rural areas with diffuse source inputs typically from land spreading and aerial deposition; and urban areas where the diffuse source inputs come typically from traffic and industrial activities.

Note 2 to entry: Diffuse source input usually leads to sites that are relatively uniformly contaminated. At some sites, the input conditions can nevertheless cause a higher local input, such as near the source or where atmospheric deposition/rain is increased. Two types of main diffuse source input can be considered: one in rural areas (e.g. atmospheric deposits, spreading); and one in urban areas (e.g. traffic, industries).

[SOURCE: ISO 11074:2015, 3.3.9, modified — Note 1 to entry has been replaced with new text. The last sentence in Note 2 to entry has been added.]

3.4
pedo-geochemical concentration
concentration of a substance in a soil resulting from natural geological and pedological processes, excluding any addition of anthropogenic origin

Note 1 to entry: It is difficult to determine the precise pedo-geochemical concentration of certain substances in soil due to the presence of anthropogenic diffuse contamination.

3.5
pedo-geochemical background value
statistical characteristic (3.8) of the *pedo-geochemical concentration* (3.4)

Note 1 to entry: Any estimate of pedo-geochemical background value is prone to certain errors given the uncertainty associated with determining the pedo-geochemical concentration.

[SOURCE: ISO 11074:2015, 3.5.9, modified — In the definition, “concentration” has replaced “content”.]

3.6
anthropogenic concentration
concentration of a substance in a soil resulting from anthropogenic origin

3.7
anthropogenic background value
statistical characteristic (3.8) of the anthropogenic *background concentration* (3.1) of a substance in soils

3.8
statistical characteristic
numerical value calculated from a *variate* (3.10) of a selected parameter of the population

EXAMPLE Mean, median, standard deviation, standard error, percentiles of the ordered frequency distribution.

[SOURCE: ISO 11074:2015, 3.5.11, modified — “selected” has replaced “chosen” and “standard error” has been added in the example.]

3.9
study area
three-dimensional definition of the area where samples are to be obtained from and, thus, for which the *background values* (3.2) are to be determined

[SOURCE: ISO 11074:2015, 5.2.29]

3.10
variate
set of observed values of a variable

EXAMPLE Series of numbers of the concentration of a substance in soil; 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 determination of background values of inorganic and organic substances in soils consists 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 concentration is dominated by the pedo-geochemical concentration and, consequently, by the mineralogical composition of the soils' parent material. Pedogenetic processes can 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. Therefore, the background concentration 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 concentration of soils. Nonetheless, a detailed knowledge of the background concentration and its natural fraction for the substances of concern is essential for any evaluation of the current status of soils for environmental or land use-related aspects, as well as 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 should 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 should be defined, which together form the basis of the technical programme.

This guidance identifies:

— essential requirements of sampling strategies and procedures;

— minimum requirements regarding the necessary steps and ways of sample pretreatment;

— analytical methods;

— statistical evaluation procedures for determining sound and comparable background values.

Guidance is given on:

a) evaluating existing data from different data sources;

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

These situations represent the two extreme starting positions for the process of compiling background values. In practice, there is also a third intermediate situation in which additional data are collected because the quantity or quality of the existing data are insufficient.

5 Procedures

5.1 General

The procedures to determine background values encompass aspects of sampling (e.g. strategy, procedure), soil analysis (e.g. pretreatment, extraction, 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;

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 that are determined. The technical approaches describe aspects such as the “where”, “what”, “how” and “when”. Together, the technical approaches determine the technical programme that will provide the required background values.

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 could be:

— to identify the current concentrations of substances in soils, e.g. in the context of soil-related regulations;

— 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 chemical substance contents due to geogenic reasons or human impact, etc.

In order to meet the objective, the technical approaches could include:

— definition of the substances and parameters, e.g. the background values to be estimated may be the total metal concentration or the bioavailable metal concentration (see 5.2.2);

— definition of the study area, e.g. the (three-dimensional) definition of the area where samples are obtained from, including 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, e.g. whether the historical or current concentration is relevant for the objective (see 5.2.4);

— definition of the size and geometry of the area to be sampled at a sampling location (see 5.2.5);

— definition of the pretreatment of the sample (see 5.4.2.2) and the fraction of soil to be analysed.

5.2.2 Substances and parameters

5.2.2.1 Basic 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 substances are of primary interest because of their potential to adsorb and accumulate in the soil. Substances in which the concentration can be influenced by 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) should be provided to assist in the interpretation of the concentrations of substances. A number of so-called basic soil parameters influence soil processes that affect the concentrations of inorganic and organic substances. Table B.1 lists these parameters, which should be analysed in accordance with the given International Standards.

5.2.2.2 Persistent compounds

5.2.2.2.1 Inorganic substances

Within the group of inorganic substances, trace elements (e.g. metals, micronutrients) are most often analysed (see Table B.2). While constituting an urban geochemical background, it is recommended to analyse the whole package listed in Table B.2; most of these elements are found with high values due to human activities, but some minerals can appear with high value naturally.

Concerning the analytical methods, a distinction should be drawn between different extraction/preparation methods (see Table B.2), as very few methods determine the total concentration that could be needed, e.g. when calculating element stocks. Besides total concentrations, the (eco-) toxicologically more relevant mobile fractions (see Table B.2) are of increasing interest, e.g. if pathway-related questions are to be examined. An analysis of the parameters in Table B.2 should be carried out in accordance with International Standards given in Table B.2.

5.2.2.2.2 Organic substances

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

It is not recommended to sample and analyse non-persistent substances as they will not be detected by the laboratory because of their behaviour (e.g. volatilization, high degradation). These kinds of substances are normally analysed for detecting sources of contamination.

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 can be extracted from naturally occurring organic materials (e.g. organic matter, decaying vegetation, peat, charcoal), and that non-specific analyses, in particular, can, therefore, give misleading results.

When collecting new data for determining background values, it is recommended that the investigation programme 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, suitable storage of soil samples for subsequent analyses of organic or inorganic substances is of crucial importance. Besides the substances of concern (see Tables B.2 and B.3) and additional soil parameters (see Table B.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 are to be of use for other assessments in future investigations.

NOTE Guidance on the storage of soil samples is provided in ISO 18512.

5.2.2.3 Non-persistent compounds

In some cases, generating background values on non-persistent substances could be of interest. Special care should be taken when volatile, degradable organic substances or transformable, inorganic species are the subject of the study. A detailed description and documentation of sampling and analysis is of particular importance in such cases. Storage or archiving of samples is not recommended because of the behaviour of such organic and inorganic species.

5.2.3 Study area

The definition of the study area (see 5.2.3) can be based on two different principles:

— 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 horizon of interest that is to be studied should also be defined;

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

It is possible, of course, to combine the spatial and typological definition of the study area.

EXAMPLE A combination of spatial and typological definitions of the study area could 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 should be detailed at a level where there cannot be any misinterpretation of 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 should be defined.

As the general objective is to determine background values, a safety zone around that (type of) source can be defined, which thereby excludes parts of the more generally defined study area. There could also be specific zones in which the data are considered separately from the rest of the study area.

Samples to be used for the determination of background values should not be taken from near potential localized sources of contamination (e.g. roads, industrial sites).

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 (e.g. pedogenesis, biogeochemical cycles) as well as by anthropogenic diffuse source input. Two different time scales can be distinguished:

— the period in which the background value can vary significantly 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 could 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 the following (see also ISO 16133):

— the expected enrichment of substances in soils (accumulation, e.g. due to diffuse source input);

— the expected loss of substances in soils (e.g. due to leaching, biodegradation, plant uptake);

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

5.2.5 Scale of sampling

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 could be different. The scale is, therefore, an important technical aspect on which a decision is to be made prior to data collection.

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

The variability of the natural pedo-geochemical concentration and of the background concentration often increases with the size of the study area (population), but decreases with the size of the sample support (scale of sampling). Increasing the sampling support is a method that can reduce the variability of the background concentration. However, increasing the sampling support often makes the sampling more laborious and is only recommended under conditions where the sampling and sampling preparation errors can be minimized. It is recommended to use the same scale of sampling if the natural pedo-geochemical value is used to evaluate soil contamination. If the background values are used to support delineation of contaminated land, it is recommended to use the same scale of sampling. For large scale sampling support, composite sampling (see 5.4.1.6) is often preferred in order to avoid the handling of excessively large amounts of soil.

5.3 Evaluation of existing data

5.3.1 General

When using existing data, specific care should be taken concerning the quality and comparability of data, particularly if the data originate from different sources, by consideration of the measurement uncertainty. Data with appropriate information should 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 bias between the data sets[36]. Harmonization of sample sets is essential if sound and reliable evaluation of the combined data is to be achieved. The harmonization strategy should include:

a) a check on the completeness of the data sets (including estimates of the uncertainty of each measurement);

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

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

NOTE Natural anomalies of the pedo-geochemical background can generate high values that can appear as contaminated samples. It is important to anticipate these cases with an accurate desk study in order to not eliminate these samples.

5.3.2 Completeness of data sets/minimum requirements

In order to ensure a minimum level of data quality, sufficient and sound information about the data should be provided, for example:

— the date of sampling;

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

— the scale of sampling (area or local);

— the site location (coordinates);

— the sampling depth intervals;

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

— the method of sample preparation (e.g. dying, crushing, sieving);

— the method used to extract and analyse the components (including quality assurance data, estimate of analytical uncertainty and detection limits in accordance with ISO 18400-106);

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

— the urban areas specificities (e.g. excavated materials).

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

The minimum information required about a data set depends, among other things, on the substances of concern, the area, spatial reference to be considered and the approach pursued to achieve an adequate spatial representation of the sampling location.

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

5.3.3 Comparability of data (sampling, nomenclatures, analyses)

Different sampling strategies can have a crucial impact on the comparability of data sets. Problems arise, in particular, through the comparison of horizon versus depth level-related samples and that of composite versus individual samples. In addition, the representative nature of the variate for a sample population with regard to the same scale for an area should be taken into account. Also, an uneven spatial distribution of the sampling points within an area can cause biased estimates of the parameters of the frequency distribution due to an overestimation of some parts of the study area. Following an initial review of the raw data on a map, an appropriate spatial interpolation method should be used. Geographical information system (GIS) software is particularly useful for this task, especially for data sets (see ISO 18400-104:—, 6.1 and Annex H). It is strongly recommended to carefully balance the possible inaccuracies introduced by merging data sets from different campaigns with the advantage of an increasing number of samples and, consequently, an increasing representation of a population.

The extent to which different sample pretreatments and analytical procedures (e.g. extraction, measurements) can be compared and harmonized should be evaluated in each individual case (e.g. against the intended accuracy of the background value). For all the 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 B.2). If the ranges of classified background values as target variables are broader, the demands of data comparability are lower. Nonetheless, the assessor should bear in mind that merging data sets analysed by different analytical procedures invariably requires compromises to be made. By definition, the background values are relatively low, except in natural high pedo-geochemical background zones. It can be important to take into account detection and quantification limits of each device of analysis when judging the comparability of data.

5.3.4 Examination of outliers

The background concentration of substances in soils includes the moderate diffuse input into the soil. Therefore, locally contaminated sites and natural anomalies are excluded from the population of background concentrations. Consequently, data obviously stemming from locally contaminated sites should be identified and eliminated from the respective data set. The detection of data outliers and unusual data behaviour is one of the main tasks in the statistical analysis of geochemical data. To this end, several statistical tests for identifying outliers are applicable, e.g. tests on the distribution of the data, exploratory data analysis (boxplots) and principle component analysis (see also 5.5.1.2.1 and Annex A). In practice, the use of the boxplot for preliminary class selection to display spatial data structure in a map has proven to be a powerful tool for identifying the key geochemical processes behind a data distribution. Different outlier tests are available (e.g. Dixon test, Grubbs test) and should be chosen according to the relevant assumptions for the study[37]. In the case of multidimensional data (samples from different soil horizons), principal component analysis can be important if dimensional reduction is possible on transformed data.

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. On the contrary, if the explanation shows that it is a natural anomaly, the outlier is not eliminated, but the scale of the study should be reconsidered. For example, if there was a few high values that are likely to be due to natural processes, the options are either to increase the spatial scale of the study to confirm this hypothesis or to consider the use of zoning/depth interval classifications (e.g. due to lithology/geology) to better differentiate between different background populations.

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 concentration and the usual concentration 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 could exhibit different compositions. These parameters can also influence the variation of concentrations and should be considered: soil organic matter, soil sorption coefficient, carbonate content, particle size, moisture content, cation exchange capacity, pH.

Land use and distance to contamination sources also influence the concentration 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 one hectare, where one sample of each of the soil layers or horizons of interest is collected.

Two strategies for selecting the sampling sites within the study area are presented: the systematic approach (5.4.1.1.2) and the judgemental approach (5.4.1.1.3). 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 judgemental one. This hybrid approach is recommended when investigating in urban areas: the systematic approach allows optimization of the coverage of the area, while the judgemental approach allows flexibility in the location of sampling points within each cell so as to be able to focus on available soil surface and excluding (potentially) contaminated soil.

5.4.1.1.2 Systematic approach

The sampling locations 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 concentration. In principle, the interval between the sampling locations 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 example, square cells with a 16 km, 5 km or 2,5 km side can be used at the scale of a country, while square cells of a few hundred metres 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 example, 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 nearby 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 concentration of soils is of interest).

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

For urban areas, account should be taken of the possible excavation of soils for urban development. Thus, samples should be taken from both superficial and deep soils that can be excavated during construction or similar activities (typically up to 3,5 m deep).

A comprehensive site and soil description (see Table B.2) should be prepared at the same time as the samples are taken.

NOTE Most of the time, the regular approach is used for the upper layer of the soil.

5.4.1.1.3 Judgemental approach

In the judgemental approach, the soil is stratified according to soil parent material (for inorganic substances), soil type and land use. Potential sampling locations possibly highly contaminated by adjacent point sources are rejected.

The judgemental approach needs detailed information about the area to be investigated. Information (such as on geology, pedology, land use and sources of possible contamination) should be gathered and evaluated in order to elaborate the sampling scheme. ISO 18400-203 gives guidance on such preliminary investigations.

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. of the soil series), as this taxonomy level generally explains most of the variation of the soil properties. Finally, the horizon to be sampled should be chosen. Where there are no obvious horizons, the depth of sampling should be defined.

Within one stratum, the sampling locations 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 and a systematic sampling scheme. To estimate parameters of a frequency distribution of the background concentrations, a random sampling approach is most appropriate. For mapping background concentrations, in general, a (centred) regular grid is more appropriate.

5.4.1.2 Number of samples

Background values cannot be summarized in a central parameter such as the mean. The spatial variability of concentrations in soil should be described as precisely as possible. The extent of the skewness can be calculated and assessed to see whether an outlier is an issue or not. In the case of a normal probability distribution, the number of samples necessary for the estimation of the standard deviation is independent of the standard deviation of the population. It can be determined using Table 1, which shows that a minimum number of 30 samples is necessary to estimate the standard deviation of a normal population with acceptable reliability.

Table 1 — Maximum relative error (e_r) on the estimation of the standard deviation of a normal population

n	e_r %
10	57
20	35
30	27
40	23
50	21
Confidence limit a = 0,05	
n = number of samples	

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.

To obtain a sufficient number to estimate background values, more than 30 samples should be taken in case of the elimination of outliers or unexpected problems.

5.4.1.3 Soil description

The interpretation of background concentrations 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 2. It is important to bear in mind that the reliability of data interpretation strongly depends on a profound knowledge of the study area. Hence, the collection of parameters, indicated in Table 2, should be as comprehensive as possible.

Table 2 — Parameters for site and soil description

Parameter		International Standard
Landform and topography	Topography, landform, land element, position coordinates, slope microtopography	ISO 25177
Land use and vegetation	Land use, human influence, vegetation	ISO 25177
Geology and lithology	Kind of parent material, effective soil depth	ISO 25177
Surface characteristics	Rock outcrops, surface coarse fragments, erosion phenomena, surface sealing, surface cracks, other characteristics	ISO 25177
Soil–water relationship	Surface water balance, rainfall, evapotranspiration, groundwater recharge, presence and depth of water table, site drainage, moisture conditions	ISO 25177
[a] See Reference [38].		

Table 2 *(continued)*

Parameter		International Standard
Soil type/soil profile description	Soil unit in regards of the classification system used	ISO 25177
	Sequence and depth of diagnostic horizons, kind of boundaries	
	Soil colour (matrix, mottling)	
	Organic matter	
	Texture, coarse elements, presence of non-soil material pedo-features	
	Carbonates, field-pH, electrical conductivity	
	Structure, voids, fracturing, inhomogeneties	
	Compactness and consistence	
	Total estimated porosity	
	Roots, worm channels, biological activity[a]	
[a] See Reference [38].		

5.4.1.4 Sampling depth

Background concentrations and values vary with soil depth due to pedogenesis, soil/land use, soil 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. However, it will give a less precise measurement of the background concentration, 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 concentration for those substances that 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 concentration. 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 concentration of this layer or horizon.

Note that, in some cases, contamination can also enter the soil from underground (e.g. by contaminated groundwater).

5.4.1.5 Sampling period

Practical aspects, mainly concerning the access to the soil horizon or layer, should be considered before choosing the sampling period. For example, it can be difficult to sample deep horizons during a wet season, due to the presence of a water table close to the surface. Samples to determine background values should not be taken from a saturated zone. Sampling during the dry season can be rendered difficult by drought, which makes soil hard to penetrate with sampling tools. Access to cultivated plots can be difficult because of growing crops. In this case, it is advisable to sample shortly after harvest and/or immediately after sowing.

When soil parameters are not expected to vary within one year (e.g. metal concentrations), which is the common case for most soil substances, sampling can be carried out at any period of the year. If soil parameters of interest are expected to vary within one year (e.g. soil organic matter), sampling should be carried out at a specified period of time, but can be spread over a number of years.

5.4.1.6 Sampling technique

Sampling should be performed in accordance with ISO 18400-104 and ISO 18400-105. The following recommendations may generally be taken into account. The sampling techniques depend on the depth or horizon that is sampled. If only the surface horizon or layer is sampled, a hand corer or equivalent tool can be used.

If deep horizons or layers are sampled, a powered corer should be used. All sampling tools should be designed and/or used in order to avoid cross contamination between layers or horizons. All of the materials used for sampling, transport, labelling and storage of the samples should not release significant amounts of the elements or substances that are to be determined (see ISO 18400-102, ISO 18400-105 and ISO 18512).

Sampling in a soil trial 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 (for a detailed description of sampling procedures in trial pits, see ISO 18400-102).

Composite samples can be formed by combining several increments from each of the layers or horizons (these could be in the form of spot sample, incremental samples or cores, etc.) taken according to a random or predefined (e.g. regular) pattern[39]. Composite samples are recommended for urban areas because of the very high variability of element concentrations at a local scale, particularly in soils from fill materials. When sampling in a soil trial pit, it is recommended to clear a sufficient surface of the sampled horizon in order to take several cores to be mixed together.

The number of increments in a composite sample should be large enough to ensure representability for the sample volume. A set of 30 increments normally ensures this, but the number can be reduced if it does not introduce an unacceptable sampling error (see ISO 18400-104).

5.4.2 Soil analysis

5.4.2.1 General considerations

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

5.4.2.2 Pretreatment

Sample pretreatment should be carried out in accordance with the International Standard dealing with this step of analysis, as well as with the requirements of the analytical methods. Special care should be taken when volatile, degradable organic substances or transformable inorganic substances are the subject of the study.

NOTE ISO 11464 provides guidance on laboratory pretreatment methods. Pretreatment methods for application in the field are described in ISO 18400-201.

5.4.2.3 Analysis

Analysis should be performed in accordance with the relevant standards for analytical methods, if available (see Tables B.1 to B.3).

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

If it is not possible to reach acceptable detection/quantification limit(s), then the statistical tests should take into account the presence of values under the detection/quantification limit(s)[40].

The analysis should be performed under a quality control procedure (see Clause 6).

5.4.2.4 Storage of the samples

It can 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 retained soil samples from former campaigns are important in order to judge any drift in analytical methodologies or to allow for new techniques. The 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 concentration of non-volatile elements and persistent organic compounds.

NOTE Guidance on the storage of samples is provided in ISO 18512. Guidance on the packaging and preservation of samples on site and their transport to the laboratory is provided in ISO 18400-105.

5.5 Data processing and presentation

5.5.1 Statistical evaluation of data

5.5.1.1 General considerations

The pedo-geochemical or background concentration 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:

— select the data corresponding to pedo-geochemical or background concentration from the whole data set (e.g. test for outliers)[40];

— 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 different homogeneous strata using the results from sampling. To determine areas for representative background values, the raw data should be reviewed on a map to check for patterns or previously unknown additional populations with different background values. This is useful check in the 2D plane and is also appropriate for 3D. 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 initially defined strata are considered homogeneous and are not contaminated by close point sources.

The following subclauses propose examples for selecting sufficiently representative data and for characterizing pedo-geochemical and background concentrations. The examples are indicative and are applicable to both starting positions (e.g. evaluating existing data or collecting new data). Other techniques can be found in the specialized literature.

If the concentrations are below the analytical quantification limit, a decision should be taken about how to take them into account in any statistical analysis. A sensitive evaluation of the influence of the analytical quantification limit should be done in order to make a decision. The choice of the method for dealing with the quantification limit should be justified[41].

5.5.1.2 Selection of representative data

5.5.1.2.1 Systematic approach

A 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. The numerical divisions for the frequency graph should be consistent with the range and precision of the data recorded for each measurand.

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 should be used with great care, after checking the reliability of this assumption.

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

The use of a geographical information system (GIS) is typically used to compare the data with background mapping data (e.g. geology, area infrastructure), and to help to discriminate representative population for background values.

An efficient way of selecting homogeneous statistical samples for pedo-geochemical or background concentration is to gather values measured on samples supposed to belong to the same horizon or stratum. This can be done by selecting and gathering the data according to stratification by land use, distance to point sources, soil parent material, and soil and/or horizon characteristics. It 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 different homogeneous strata using the results from sampling, and also if sufficient data are available to represent each stratum. For example, samples can be gathered according to soil type, texture class, pH, organic matter content or total iron content. Then, the selection of the data of each a posteriori stratum can be carried out as described before.

5.5.1.2.2 Judgemental approach

The judgemental approach provides a statistical sample for each investigated stratum. As these strata are self-evidently 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 example, because of sampling in the wrong stratum, the existence of unexpected contamination, analytical mistake.

It is, therefore, necessary to check the homogeneity of each statistical sample representing each stratum. This can be done using histogram, exploratory data analysis or principal component analysis, as described for the systematic approach (see also Annex A). A simple way to detect anomalous concentrations is to plot the substance concentration against a soil characteristic known to affect the distribution of the considered substance (e.g. clay content, cation exchange capacity, total iron content). 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 the soil being studied.

NOTE Z score can also be used if the sample size if 30 or above.

5.5.2 Data presentation and reporting

Background values can be presented in various manners, for example, in tables, boxplots or histograms. A useful tool is the Tukey's whisker, which allows the calculation of an indicator able to detect a suspicious trace element amount while taking into account the local context[41].

Background values can be represented on a map of the investigated area. 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. A sampling report in accordance with ISO 18400-107 should be prepared on completion of the field work. A detailed report containing the analytical results and the conclusions drawn from the investigation should be prepared to complete the investigation (some relevant guidance can be found in ISO 18400-203). An example of the structure and the main contents of such a report is given in ISO 18400-107.

The reports presenting the results of data analyses and compilation of background values will often be scrutinized by regulators and other interested parties, including the general public. It is important, in order to be able to identify any "drift" in results caused by changes in analytical methodologies, that such reports are of a high technical standard, but that they also take account of the diverse and often non-technical readership. Use should, therefore, be made of tabular summaries, and 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 document provides guidance on the types of data that are required for compiling background values and indicates for which parameters or procedures International Standards are available. The assessor should 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 should be evaluated. The data should be sufficient in terms of:

— type;

— quantity;

— 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 document 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 an International Standard (including requirements for sampling), is no greater than the confidence there is in the representativeness of the data.

The assessor should bear in mind the disproportionate costs and time delays that can result if it is necessary to carry out an additional sampling exercise. For example, if 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 can be ensured by:

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

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

— using laboratories that apply methods accredited to ISO/IEC 17025;

— using laboratories accredited to a quality management system, e.g. ISO 9001;

— 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 in the same proficiency test in order to control precision and accuracy, and thus quality and comparability of the results;

— using a 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 could be present for individual samples, for several samples, or for one or several of the parameters analysed.

Annex A
(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 and 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 pretreatment, 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 contamination;

— a difference in soil type.

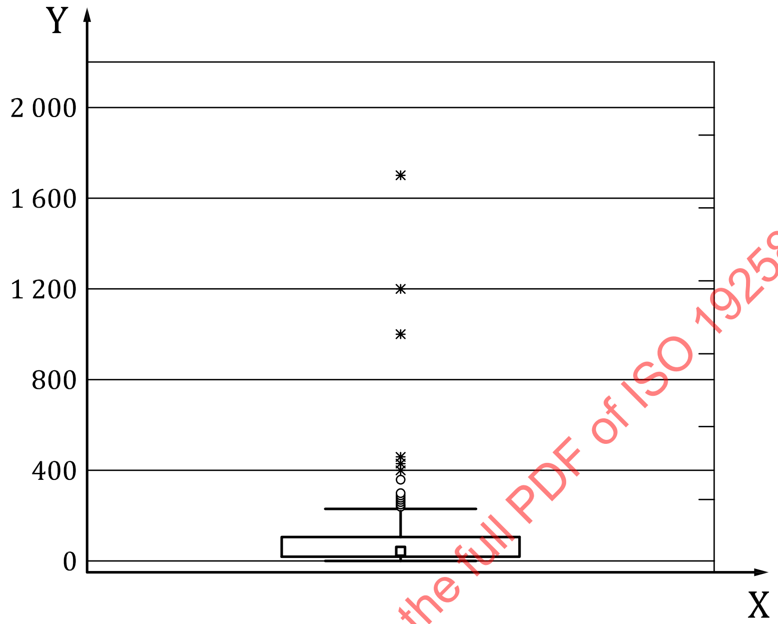
In all three cases, the outlier should not be included in the (part of the) database describing the background soil quality, but should be kept in the overall database, as these high values participate to describe the local context.

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 falls outside the expected width of that distribution, which makes the value an outlier.

In order to detect outliers in a simple and visualized 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. In particular, the value used for positioning the whiskers (in this case 1,5) is based on an assumption of the underlying distribution of the measurements.

Figure A.1 shows a box and whisker plot for lead concentrations found in the study area. The outliers in Figure A.1 are divided into two groups, the “outliers” and “extremes”, where the extremes represent measurements over two times the inter-quartile range. Other factors for the distinction between outliers and extremes (such as a value of 3) are also used; sometimes the 10th and 90th percentiles are used to determine the outliers defined as the values respectively below and above those percentiles.

As shown by Figure A.1, a number of outliers are detected. However, whether these are truly outliers should be investigated.



Key

X distribution of the population

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

non-outlier

non-outlier min

75 %

25 %

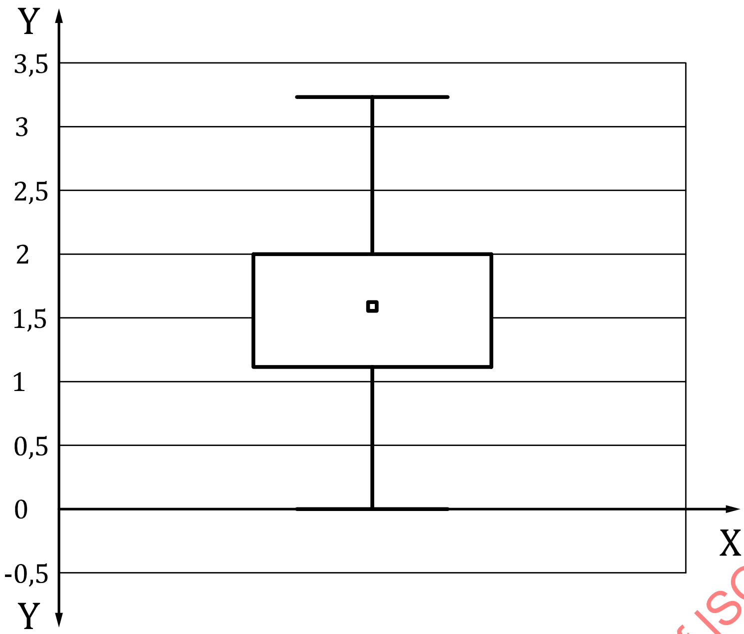
median

outliers

extremes

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

Figure A.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 A.3.

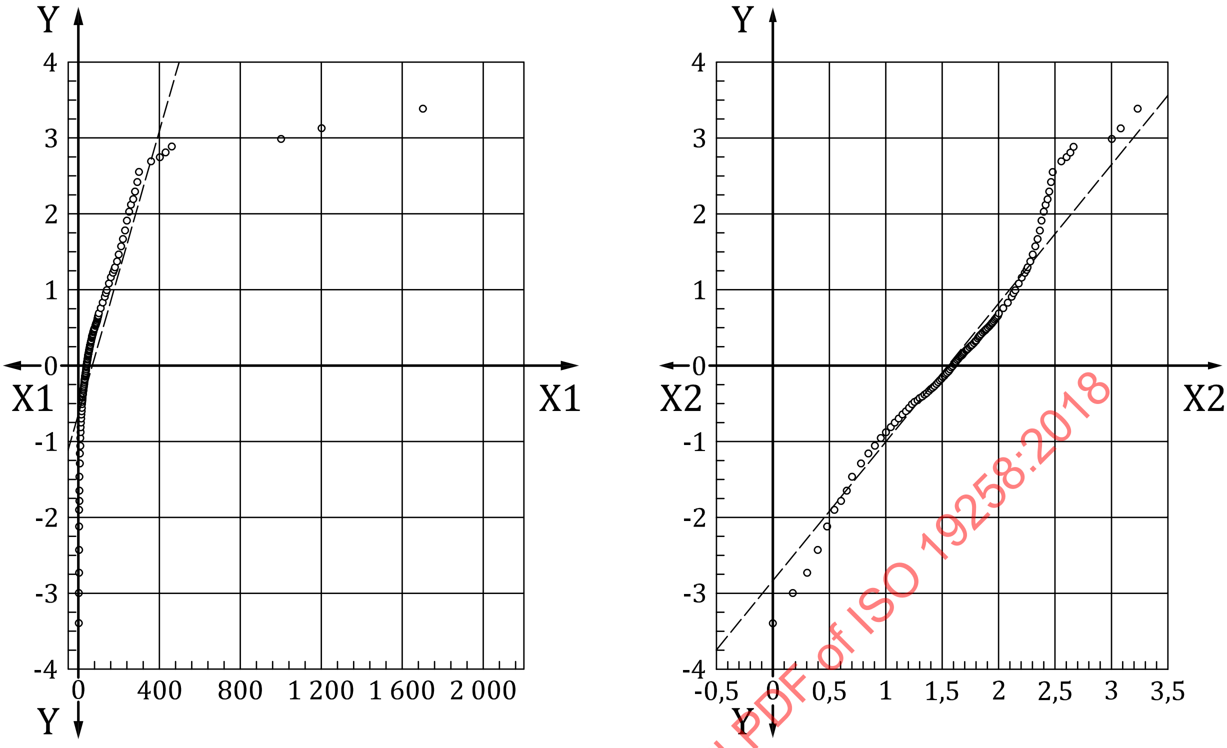


Key

X distribution of the population

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

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



Key

Y expected normal value

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

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

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

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 A.3) reasonably well. This does not imply that no explanation for the outliers should be sought, but it can provide part of the explanation as to why the measurements 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 while the data are high, they 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 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 for explaining and understanding the available measurements.

Another approach to selecting the values corresponding to pedo-geochemical or background concentration 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 simple to interpret (visually). It is essential that the tested data meet the conditions for applying the test (e.g. normality of the data). If these conditions are not met, measurements can be falsely detected as outliers.

