INTERNATIONAL STANDARD

ISO 5667-20

First edition
2008-03-15

Water quality — Sampling —

Part 20: **Guidance on the use of sampling data for decision making — Compliance with thresholds and classification systems**

Qualité de l'eau — Échantillonnage —

Partie 20: Lignes directrices relatives à l'utilisation des données d'échantillonnage pour la prise de décision — Conformité avec les limites et systèmes de classification

Reference number
ISO 5667-20:2008(E)

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 5667-20 was prepared by Technical Committee ISO/TC 147, *Water quality*, Subcommittee SC 6, *Sampling (general methods)*.

ISO 5667 consists of the following parts, under the general title *Water quality — Sampling*:

— *Part 1: Guidance on the design of sampling programmes and sampling techniques*

— *Part 3: Guidance on the preservation and handling of water samples*

— *Part 4: Guidance on sampling from lakes, natural and man-made*

— *Part 5: Guidance on sampling of drinking water from treatment works and piped distribution systems*

— *Part 6: Guidance on sampling of rivers and streams*

— *Part 7: Guidance on sampling of water and steam in boiler plants*

— *Part 8: Guidance on the sampling of wet deposition*

— *Part 9: Guidance on sampling from marine waters*

— *Part 10: Guidance on sampling of waste waters*

— *Part 11: Guidance on sampling of groundwaters*

— *Part 12: Guidance on sampling of bottom sediments*

— *Part 13: Guidance on sampling of sludges from sewage and water treatment works*

— *Part 14: Guidance on quality assurance of environmental water sampling and handling*

— *Part 15: Guidance on preservation and handling of sludge and sediment samples*

— *Part 16: Guidance on biotesting of samples*

- *Part 17: Guidance on sampling of bulk suspended solids*
- *Part 18: Guidance on sampling of groundwater at contaminated sites*
- *Part 19: Guidance on sampling of marine sediments*
- *Part 20: Guidance on the use of sampling data for decision making — Compliance with thresholds and classification systems*

The following parts are under preparation:

- *Part 21: Guidance on sampling of drinking water distributed by non-continuous, non-conventional means*
- *Part 22: Guidance on design and installation of groundwater sample points*
- *Part 23: Determination of significant pollutants in surface waters using passive sampling*

Introduction

This part of ISO 5667 concerns the use of information on water quality obtained by taking samples in taking decisions — in measuring success, failure or change, in the context of the inevitable uncertainties associated with sampling. This part of ISO 5667 provides guidance on controlling the risk of such uncertainties leading to non-optimal decisions.

Non-optimal decisions can also stem from the way in which thresholds for discharges and targets for environmental waters are formulated or set out in regulations and permits. This part of ISO 5667 also examines the problems caused when compliance with these thresholds is assessed using data obtained by sampling.

This part of ISO 5667 aims to ensure that future laws, regulations, and guidance assert the requirement to assess and report statistical significance.

NOTE 1 Decisions might result in the commendation or criticism of people, sites, companies, sectors or nations. Decisions can give rise to legal action and/or expensive and time-consuming remedial actions to improve water quality.

Figure 1 shows the links between the following topics:

a) the setting up of thresholds for taking decisions on the need to improve water quality, possibly including criteria to minimize water quality deterioration;

b) the establishment of sampling programmes to satisfy the requirements of these thresholds and the need to assess performance against them;

c) making use of the outcome of sampling programmes to take decisions.

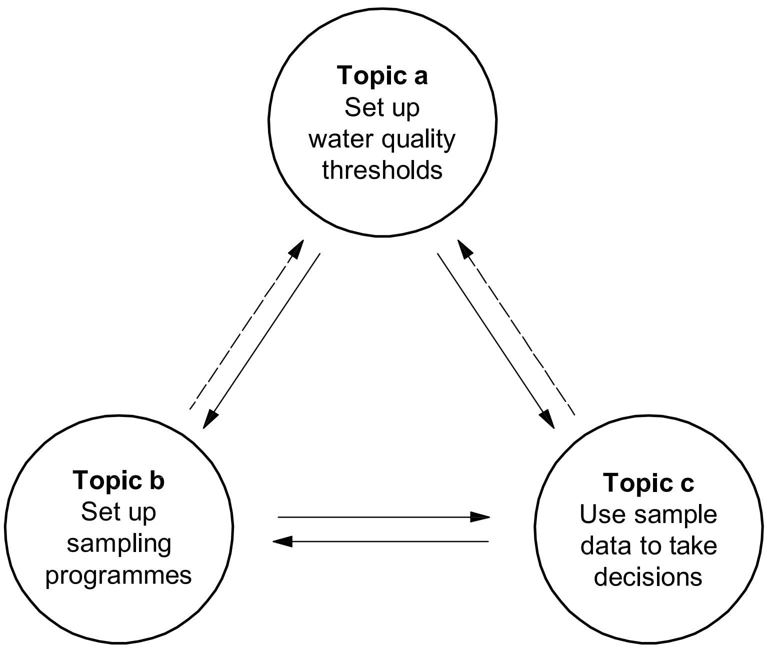


Figure 1 — Links between topics associated with sampling and taking decisions

This part of ISO 5667 deals with topic c). Topics a) and b) are huge and wide ranging in their own right, and their detailed treatment lies outside the scope of this part of ISO 5667. Nevertheless, this part of ISO 5667 does make recommendations for the expression of targets and thresholds for water quality [topic a)], which are important when using sample data to take decisions. This part of ISO 5667 also gives advice on what is required for sampling programmes [topic b)] in order that they be compatible with the way thresholds are defined, and so as to place no unnecessary difficulties and errors in the process of taking decisions.

Other areas which lie outside the scope of this part of ISO 5667 are: the detailed mechanics of taking and handling samples; assurance that samples are representative over time of the body of water being sampled; and performance of chemical analyses on samples. These are all covered in other documents. Nonetheless, if poorly obtained results from these areas can add substantially to overall sampling uncertainties and impose extra difficulties in taking decisions. This part of ISO 5667 describes some of these extra difficulties.

This part of ISO 5667 does not cover the full range of statistical techniques that may be applied and the circumstances in which they should be used. The main purpose is to establish the principle that uncertainty from sampling and analysis (and errors generally) should always be assessed and taken into account as part of the process of taking decisions. If this is not done, incorrect decisions can result, for example, on where action is needed, and the scale of that action.

NOTE 2 Some statistical techniques are used as illustrative examples. These are techniques that have seen routine use in some regulatory regimes that take proper account of statistical uncertainties. They are suitable for use in situations that resemble the worked examples discussed.

It is not the purpose of this part of ISO 5667 to direct the development of regulatory conditions. This part of ISO 5667 provides principles and tools to support management, including regulation. It is recognised that regulatory thresholds are developed using a range of strategies that incorporate technical, social and legal considerations. It is also recognised that tools other than statistical data analysis are likely to be used in interpreting and applying thresholds.

Water quality — Sampling —

Part 20:
Guidance on the use of sampling data for decision making — Compliance with thresholds and classification systems

1 Scope

This part of ISO 5667 establishes principles, basic requirements, and illustrative methods for dealing with the use of sample data for decision making based on the assessment of the confidence that water quality:

a) meets targets and complies with thresholds;

b) has changed; and/or

c) lies in a particular grade in a classification system.

This part of ISO 5667 also specifies methods for preliminary examination of the sensitivity of decisions to error and uncertainty, although it does not cover the full range of statistical techniques.

This part of ISO 5667 provides general advice on decision making related to constraint formulation for expression of thresholds and targets and the form and scale of sampling programmes.

NOTE 1 In the water industry, “standard” is commonly used to indicate the value or limit of a parameter of interest. However, in this part of ISO 5667, the term “threshold” is used to avoid confusion with published national, regional, and International Standards.

NOTE 2 This document is framed in terms of sampling and measurement of chemical concentrations, in particular those subject to strong day-to-day temporal variations. The principles apply, however, to any item estimated by sampling which is subject to random error, including microbiological and biological data, and data subject to strong spatial variations.

2 Summary of key points

Water quality is often assessed by the results of chemical analysis of a number of samples taken over a period of time.

Uncertainty is introduced by the action of random chance in taking samples. It can be present in any set of measurements of water quality taken over a period of time. The values for chemical analysis of these samples depend on the quality of the particular small volumes of water that are extracted or measured. If water quality varies in space or time, a second set of samples taken over the same period will have different values because these samples are made up of different small volumes of water taken at different times. Each set of samples allows an estimate of the true water quality. These estimates will differ: they will have a different mean and span a different range. They have the potential, if taken at face value, to suggest different conclusions about compliance with thresholds and targets.

Sampling uncertainty (or sampling error) is the term often given to this effect. Sampling uncertainty includes uncertainties and errors associated with chemical analysis, and occurs even in the case of trivial errors in chemical analysis and if there are no mistakes in the methods by which samples are taken and handled.

Sampling uncertainty is reduced if more samples are taken, but the scale of the uncertainty is often unappreciated.

In this part of ISO 5667, "overall uncertainty" includes these chance sampling effects and all the other sources of variation in a set of samples. This variability reflects the underlying signals generated by natural or perhaps unnatural processes; it includes the effects of errors in chemical analysis and the handling of samples. It might contain systematic variations from trends and diurnal, weekly, and seasonal cycles. In this context, the more appropriate term is "overall uncertainty", "overall error" or "total assay error" (ISO/IEC Guide 99:1993[5]).

Overall uncertainty should be quantified, at least approximately, and taken into account in all cases where water quality varies and sampling is used to estimate information used in decision making. This includes assessing compliance with thresholds (see Clause 5), deciding whether water quality has changed (see Clause 7), and putting waters into grades in classification systems (see Clause 8). This part of ISO 5667 recommends that:

a) thresholds for which compliance is assessed by sampling should be defined or used so that the overall uncertainty can be estimated and dealt with appropriately (see 5.2);

b) thresholds defined as absolute limits should be treated as percentiles when assessing compliance using sampling (see 5.3);

c) thresholds defined as limits to be met by a percentage of samples should be defined or used as the corresponding percentiles (see 5.4);

d) the degree of confidence should be estimated when assessing compliance with thresholds (see Clause 4); and,

e) the degree of confidence in changes or differences should be estimated when aiming to demonstrate change or no change (see 8.2).

3 Types of error and variation

3.1 General

In many procedures by which sample data are used to take decisions, there is a set of results taken over a period of time (e.g. a year). This information might be used to make such judgements as whether:

a) water quality in a river failed to meet required thresholds;

b) a treatment works performed better this year than last;

c) water quality in a lake needs improvement;

d) one company has better effluent discharge compliance than another; or

e) most of the risk of environmental impact is from a particular type of effluent discharge.

There are unlikely to be many significant changes in water quality from second to second throughout a year, but variations from day to day are common. These can be due to diurnal cycles, the play of random errors and bias from the laboratory, the weather, step changes, day-to-day and hour-by-hour variations (perhaps in the natural processes in water or caused by discharges and abstractions and changes in these), seasonal and economic cycles, and several underlying and overlapping long-term trends and cycles.

NOTE 1 Sometimes several or most of the data are reported by a laboratory as being less than a specified limit of detection. Such data are called censored data. Depending on the types of decisions that depend on the data, special statistical techniques are available for estimating the values of summary statistics and their uncertainties.

In addition, the total set of samples shall be representative of the average quality of the masses of water from which they were taken, e.g. over a period of time under review. In estimating an annual mean, it is not acceptable for all samples to be taken in April, for example. These requirements should be set up in the design of the sampling programme.

NOTE 2 Guidance on all these aspects is given in more detail in ISO 5667-1[1].

3.2 Analytical error

Analytical errors are those introduced by the process of chemical analysis and reflect that these measurements are not error free. It might be that the result for a single sample can be specified to within a specific range, e.g. ±15 %.

NOTE 1 The actual value of the analytical error depends on the capabilities of the equipment and the laboratory that has been used to perform the analysis. The discussion in this part of ISO 5667 focuses on random error, but there is always a risk of non-random error, e.g. when there is a change of instrument or methods, when the sample matrix varies greatly from the calibration materials, and for results just above the detection limit (ISO/IEC Guide 99:1993[5]).

NOTE 2 The results of chemical analysis are nowadays reported with uncertainty values in accordance with ISO/IEC 17025[6].

When a mean is calculated from n samples, the effect on the uncertainty in the estimate of the mean of random errors in chemical analysis tends to average down according to $\sqrt{n}$. For example, if the analytical error associated with a single sample were ±15 %, then the error in the estimate of the mean of a set of chemical analyses would tend to reduce to something like ±4 % for 12 samples or to ±2,5 % for 36 samples.

In using samples to take decisions, this kind of error from chemical analysis augments but is often smaller than other contributions to the overall uncertainty, especially that associated with chance in the taking of a limited number of samples. Chemical analysis error comes through as an addition to that associated with chance in the taking of a limited number of samples, but it might be a small addition. {Nevertheless, some studies need to separate sampling variance from local environmental heterogeneity (see Reference [7]).}

NOTE 3 It is not commonly understood that data fully within the statistical control of a laboratory might be unsuitable for particular interpretations because of errors associated with taking a small number of samples.

NOTE 4 This observation on the relative importance of analytical error applies generally to the types of issues considered in this part of ISO 5667, but it follows from estimating the analytical error in such cases, and comparing it with other errors. The analytical error should always be estimated. Similar points can be made about making sure samples are representative, and about checking changes to methods of sampling.

When the sample results are used to estimate the value of other summary statistics such as percentiles (e.g. the 95-percentile, which is the value exceeded for 5 % of the time), the picture is similar to that for the mean, i.e. the errors are inversely proportional to $\sqrt{n}$, but are larger than for the mean.

3.3 Overall uncertainty

Uncertainty occurs because of variations in the quality of the water being sampled, and the ability of the sampling process to accurately reflect these variations. In a set of samples taken over a period of time, the results are affected by the operation of the laws of chance in the way the particular samples came to be collected. This produces uncertainty even if:

— analytical errors are close to zero1);

— the sampling programme guarantees samples that are truly representative in time and space;

— there are no mistakes in handling the samples and recording the results of analysis.

1) Nearly always this is a hypothetical possibility. Many trace elements are measured near their detection limits and have analytical uncertainty of about ±100 %. Many organic chemicals can have recoveries of ±50 %.

In using sampling, the main source of uncertainty is usually associated with the number of samples taken. In the types of decision on activities listed in items a) to f) below, this source of uncertainty is usually a bigger issue than, for example, that associated with errors of chemical analysis. Overall uncertainty should be assessed and used to quantify uncertainty in cases where water quality varies and decisions are taken as a consequence of the following types of activities:

a) using sampling to measure and report on water quality;

b) using samples to estimate summary statistics, e.g. the monthly mean, the annual percentile or the annual maximum;

c) making statements about whether this year's summary statistics are higher or lower than last year's (see ISO/IEC 17025[6] for a wider view of the issue of looking for change);

d) establishing whether water quality exceeds a threshold;

e) using summary statistics to place water quality in a particular class within a classification system; or

f) assessing whether a change in class has occurred.

In all these situations, the aim is to assess whether the change or the status is statistically significant and to require that future laws, regulations and guidance assert the requirement to assess and report statistical significance.

4 Activities

4.1 Estimation of summary statistics2)

An estimate of a summary statistic depends on the values of water quality in the small volumes of water that happen to be captured by sampling and whether these values are measured accurately. The estimate, due to the overall uncertainty, is almost certain to differ from the true value of the summary statistic — the value that would be obtained if it were possible to achieve continuous error-free monitoring over the entire period for which the summary statistic applies.

Uncertainty can be managed by calculating confidence limits. Confidence limits define the range within which the true value of the estimate of the summary statistic is expected to lie. In the example in Table 1, the estimate of the mean from eight samples is 101 mg/l and there is a pair of 95 % confidence limits, 46 mg/l and 156 mg/l. There is 95 % confidence that the true mean exceeds the lower 95 % confidence limit of 46 mg/l and 95 % confidence that the true mean is less than the upper 95 % confidence limit of 156 mg/l. Overall there is 90 % confidence that the true value of the mean falls in the range between 46 mg/l and 156 mg/l3).

This range in the estimate of the mean represents large errors but these errors are seldom estimated or used to help take decisions based on the data. Also this discussion is for normally distributed random error. Such assumptions should be stated. Random error might not be normally distributed; it could be non-random and subject to mistakes and blunders. As a rule, the effect of these will be to increase the scale of the error. Errors should always be estimated even if this is done by making an assumption that they follow a normal distribution.

NOTE 1 The mean is used in this example because this summary statistic is commonly required by legislation. In other cases, there may grounds and opportunity to use other statistics like the median, e.g. to explain differences between large and small samples. The median is useful for data sets affected by outliers, and confidence limits can be calculated for the median.

---

2) Some documents use the concept of "sampling target". The sampling target could be the annual water quality, and a mean value over 1 year, or a 95-percentile over 1 year, is what is estimated.

3) This range is sometimes called the 90 % confidence interval, calculated from: $\overline{X} \pm t\sigma_{\overline{X}}$, where $\overline{X}$ is the mean; t is derived from the t-distribution with $n - 1$ degrees of freedom, used instead of the normal standard deviation for low rates of sampling (ISO/IEC Guide 99:1993[5]); and $\sigma_{\overline{X}}$ is the "standard error" derived from the standard deviation divided by the square root of the number of samples, $\sigma/\sqrt{n}$.

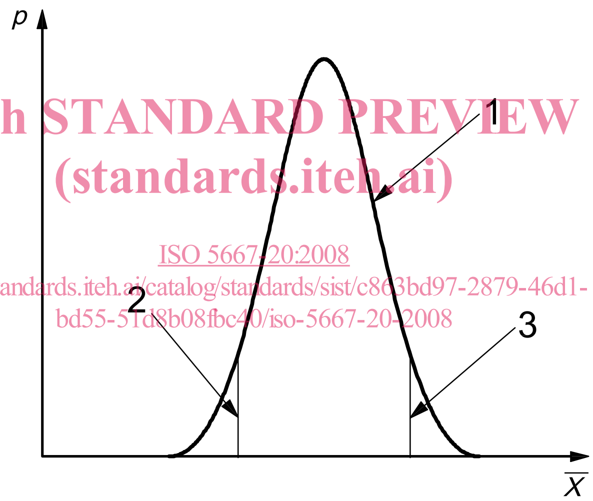
NOTE 2 Assumptions should always be stated. In this case, a normal distribution of errors is assumed.

NOTE 3 There are differences between large and small samples, i.e. use of the t-statistic versus the standard normal deviate, z.

Figure 2 illustrates the range of uncertainty. This range is an estimate of the distribution of errors in the estimate of the mean. The confidence limits are shown as points that mark 5 % and 95 % of the area of this distribution.

Table 1 — Example of confidence limits for the mean

Parameter	Value
Estimate of the mean	101 mg/l
Standard deviation	82 mg/l
No. samples	8
Lower confidence limit	46 mg/l
Upper confidence limit	156 mg/l



Key

p probability

$\overline{X}$ value of the mean

1 distribution of errors in the estimate of the mean

2 lower confidence limit

3 upper confidence limit

Figure 2 — Confidence limits on the mean

The gap between the confidence limits widens as the sampling rate is decreased (and would vanish for a continuous error-free monitor). It is also larger for estimates from the same number of samples, of more extreme summary statistics such as the 95-percentile and 99-percentile. For a typical water pollutant, the confidence limits for a mean estimated from 12 samples are ±30 %. For an estimate of the 95-percentile, this range is −20 % to +80 %.