BRITISH STANDARD

BS ISO
5725-3:1994
Incorporating Technical Corrigendum No. 1

Implementation of ISO 5725-3:1994

Accuracy (trueness and precision) of measurement methods and results —

Part 3: Intermediate measures of the precision of a standard measurement method

ICS 03.120.30

Committees responsible for this British Standard

The preparation of this British Standard was entrusted to Technical Committee QMS/16, Precision of test methods, upon which the following bodies were represented:

British Gas plc

Chemical Industries Association

Consumers' Association

Department of Trade and Industry (Laboratory of the Government Chemist)

Department of Trade and Industry (National Physical Laboratory)

Institute of Quality Assurance

Ministry of Agriculture, Fisheries and Food

Ministry of Defence

Royal Society of Chemistry

University of London

This British Standard, having been prepared under the direction of the Management Systems Sector Board, was published under the authority of the Standards Board and comes into effect on 15 April 1995

The following BSI references relate to the work on this standard:
Committee reference QMS/16
Draft for comment 91/89301 DC

ISBN 0 580 24029 0

Amendments issued since publication

Amd. No.	Date	Comments
13575 Technical Corrigendum No. 1	15 March 2002	Correction to equations $s^2_{(1)}$ and $s^2_{(2)}$ in **B.2** and the first equation in Table C.2

Contents

National foreword

This British Standard reproduces verbatim ISO 5725-3:1994, including technical corrigendum October 2001, and implements it as the UK national standard.

Parts 1 to 6 of BS ISO 5725 together supersede BS 5497-1:1987 which will be withdrawn upon the publication of BS ISO 5725-5.

This British Standard is published under the direction of the Management Sector Board whose Technical Committee QMS/16 has the responsibility to:

— aid enquirers to understand the text;
— present to the responsible international committee any enquiries on interpretation, or proposals for change, and keep UK interests informed;
— monitor related international and European developments and promulgate them in the UK

NOTE International and European Standards, as well as overseas standards, are available from Customer Services, BSI, 389 Chiswick High Road, London W4 4AL.

A British Standard does not purport to include all the necessary provisions of a contract. Users of British Standards are responsible for their correct application.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

Summary of pages

This document comprises a front cover, an inside front cover, pages i and ii, the ISO title page, pages ii to iv, pages 1 to 29 and a back cover.

The BSI copyright notice displayed in this document indicates when the document was last issued.

INTERNATIONAL STANDARD

ISO 5725-3

First edition
1994-12-15

Technical corrigendum
2001-10-15

Accuracy (trueness and precision) of measurement methods and results —

Part 3:
Intermediate measures of the precision of a standard measurement method

Exactitude (justesse et fidélité) des résultats et méthodes de mesure —

Partie 3: Mesures intermédiaires de la fidélité d'une méthode de mesure normalisée

Reference number
ISO 5725-3:1994(E)

Contents

Descriptors: Measurement, tests, test results, accuracy, reproducibility, statistical analysis.

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.

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.

International Standard ISO 5725-3 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods,* Subcommittee SC 6, *Measurement methods and results*.

ISO 5725 consists of the following parts, under the general title *Accuracy (trueness and precision) of measurement methods and results*:

— *Part 1: General principles and definitions;*

— *Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method;*

— *Part 3: Intermediate measures of the precision of a standard measurement method;*

— *Part 4: Basic methods for the determination of the trueness of a standard measurement method;*

— *Part 5: Alternative methods for the determination of the precision of a standard measurement method;*

— *Part 6: Use in practice of accuracy values.*

Parts 1 to 6 of ISO 5725 together cancel and replace ISO 5725:1986, which has been extended to cover trueness (in addition to precision) and intermediate precision conditions (in addition to repeatability conditions and reproducibility conditions).

Annex A, Annex B and Annex C form an integral part of this part of ISO 5725. Annex D and Annex E are for information only.

Introduction

0.1 ISO 5725 uses two terms "trueness" and "precision" to describe the accuracy of a measurement method. "Trueness" refers to the closeness of agreement between the average value of a large number of test results and the true or accepted reference value. "Precision" refers to the closeness of agreement between test results.

0.2 General consideration of these quantities is given in ISO 5725-1 and so is not repeated here. It is stressed that ISO 5725-1 should be read in conjunction with all other parts of ISO 5725 because the underlying definitions and general principles are given there.

0.3 Many different factors (apart from variations between supposedly identical specimens) may contribute to the variability of results from a measurement method, including:

a) the operator;

b) the equipment used;

c) the calibration of the equipment;

d) the environment (temperature, humidity, air pollution, etc.);

e) the batch of a reagent;

f) the time elapsed between measurements.

The variability between measurements performed by different operators and/or with different equipment will usually be greater than the variability between measurements carried out within a short interval of time by a single operator using the same equipment.

0.4 Two conditions of precision, termed repeatability and reproducibility conditions, have been found necessary and, for many practical cases, useful for describing the variability of a measurement method. Under repeatability conditions, factors a) to f) in **0.3** are considered constants and do not contribute to the variability, while under reproducibility conditions they vary and do contribute to the variability of the test results. Thus repeatability and reproducibility conditions are the two extremes of precision, the first describing the minimum and the second the maximum variability in results. Intermediate conditions between these two extreme conditions of precision are also conceivable, when one or more of factors a) to f) are allowed to vary, and are used in certain specified circumstances.

Precision is normally expressed in terms of standard deviations.

0.5 This part of ISO 5725 focuses on intermediate precision measures of a measurement method. Such measures are called intermediate as their magnitude lies between the two extreme measures of the precision of a measurement method: repeatability and reproducibility standard deviations.

To illustrate the need for such intermediate precision measures, consider the operation of a present-day laboratory connected with a production plant involving, for example, a three-shift working system where measurements are made by different operators on different equipment. Operators and equipment are then some of the factors that contribute to the variability in the test results. These factors need to be taken into account when assessing the precision of the measurement method.

0.6 The intermediate precision measures defined in this part of ISO 5725 are primarily useful when their estimation is part of a procedure that aims at developing, standardizing, or controlling a measurement method within a laboratory. These measures can also be estimated in a specially designed interlaboratory study, but their interpretation and application then requires caution for reasons explained in **1.3** and **9.1**.

0.7 The four factors most likely to influence the precision of a measurement method are the following.

a) **Time:** whether the time interval between successive measurements is short or long.

b) **Calibration:** whether the same equipment is or is not recalibrated between successive groups of measurements.

c) **Operator:** whether the same or different operators carry out the successive measurements.

d) **Equipment:** whether the same or different equipment (or the same or different batches of reagents) is used in the measurements.

0.8 It is, therefore, advantageous to introduce the following *M*-factor-different intermediate precision conditions (M = 1, 2, 3 or 4) to take account of changes in measurement conditions (time, calibration, operator and equipment) within a laboratory.

a) M = 1: only one of the four factors is different;

b) M = 2: two of the four factors are different;

c) M = 3: three of the four factors are different;

d) M = 4: all four factors are different.

Different intermediate precision conditions lead to different intermediate precision standard deviations denoted by $s_{I(\ )}$, where the specific conditions are listed within the parentheses. For example, $s_{I(TO)}$ is the intermediate precision standard deviation with different times (T) and operators (O).

0.9 For measurements under intermediate precision conditions, one or more of the factors listed in **0.7** is or are different. Under repeatability conditions, those factors are assumed to be constant.

The standard deviation of test results obtained under repeatability conditions is generally less than that obtained under the conditions for intermediate precision conditions. Generally in chemical analysis, the standard deviation under intermediate precision conditions may be two or three times as large as that under repeatability conditions. It should not, of course, exceed the reproducibility standard deviation.

As an example, in the determination of copper in copper ore, a collaborative experiment among 35 laboratories revealed that the standard deviation under one-factor-different intermediate precision conditions (operator and equipment the same but time different) was 1,5 times larger than that under repeatability conditions, both for the electrolytic gravimetry and $Na_2S_2O_3$ titration methods.

1 Scope

1.1 This part of ISO 5725 specifies four intermediate precision measures due to changes in observation conditions (time, calibration, operator and equipment) within a laboratory. These intermediate measures can be established by an experiment within a specific laboratory or by an interlaboratory experiment.

Furthermore, this part of ISO 5725

a) discusses the implications of the definitions of intermediate precision measures;

b) presents guidance on the interpretation and application of the estimates of intermediate precision measures in practical situations;

c) does not provide any measure of the errors in estimating intermediate precision measures;

d) does not concern itself with determining the trueness of the measurement method itself, but does discuss the connections between trueness and measurement conditions.

1.2 This part of ISO 5725 is concerned exclusively with measurement methods which yield measurements on a continuous scale and give a single value as the test result, although the single value may be the outcome of a calculation from a set of observations.

1.3 The essence of the determination of these intermediate precision measures is that they measure the ability of the measurement method to repeat test results under the defined conditions.

1.4 The statistical methods developed in this part of ISO 5725 rely on the premise that one can pool information from "similar" measurement conditions to obtain more accurate information on the intermediate precision measures. This premise is a powerful one as long as what is claimed as "similar" is indeed "similar". But it is very difficult for this premise to hold when intermediate precision measures are estimated from an interlaboratory study. For example, controlling the effect of "time" or of "operator" across laboratories in such a way that they are "similar", so that pooling information from different laboratories makes sense, is very difficult. Thus, using results from interlaboratory studies on intermediate precision measures requires caution. Within-laboratory studies also rely on this premise, but in such studies it is more likely to be realistic, because the control and knowledge of the actual effect of a factor is then more within reach of the analyst.

1.5 There exist other techniques besides the ones described in this part of ISO 5725 to estimate and to verify intermediate precision measures within a laboratory, for example, control charts (see ISO 5725-6). This part of ISO 5725 does not claim to describe the only approach to the estimation of intermediate precision measures within a specific laboratory.

NOTE 1 This part of ISO 5725 refers to designs of experiments such as nested designs. Some basic information is given in Annex B and Annex C. Other references in this area are given in Annex E.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO 5725. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO 5725 are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 3534-1:1993*, Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms.*

ISO 5725-1:1994*, Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions.*

ISO 5725-2:1994*, Accuracy (trueness and precision) of measurement methods and results — Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method.*

ISO Guide 33:1989*, Uses of certified reference materials.*

ISO Guide 35:1989*, Certification of reference materials — General and statistical principles*

3 Definitions

For the purposes of this part of ISO 5725, the definitions given in ISO 3534-1 and ISO 5725-1 apply.

The symbols used in ISO 5725 are given in Annex A.

4 General requirement

In order that the measurements are made in the same way, the measurement method shall have been standardized. All measurements forming part of an experiment within a specific laboratory or of an interlaboratory experiment shall be carried out according to that standard.

5 Important factors

5.1 Four factors (time, calibration, operator and equipment) in the measurement conditions within a laboratory are considered to make the main contributions to the variability of measurements (see Table 1).

5.2 "Measurements made at the same time" include those conducted in as short a time as feasible in order to minimize changes in conditions, such as environmental conditions, which cannot always be guaranteed constant. "Measurements made at different times", that is those carried out at long intervals of time, may include effects due to changes in environmental conditions.

5.3 "Calibration" does not refer here to any calibration required as an integral part of obtaining a test result by the measurement method. It refers to the calibration process that takes place at regular intervals between groups of measurements within a laboratory.

5.4 In some operations, the "operator" may be, in fact, a team of operators, each of whom performs some specific part of the procedure. In such a case, the team should be regarded as the operator, and any change in membership or in the allotment of duties within the team should be regarded as providing a different "operator".

5.5 "Equipment" is often, in fact, sets of equipment, and any change in any significant component should be regarded as providing different equipment. As to what constitutes a significant component, common sense must prevail. A change of thermometer would be considered a significant component, but using a slightly different vessel to contain a water bath would be considered trivial. A change of a batch of a reagent should be considered a significant component. It can lead to different "equipment" or to a recalibration if such a change is followed by calibration.

5.6 Under repeatability conditions, all four factors are at state 1 of Table 1. For intermediate precision conditions, one or more factors are at state 2 of Table 1, and are specified as "precision conditions with M factor(s) different", where M is the number of factors at state 2. Under reproducibility conditions, results are obtained by different laboratories, so that not only are all four factors at state 2 but also there are additional effects due to the differences between laboratories in management and maintenance of the laboratories, general training levels of operators, and in stability and checking of test results, etc.

5.7 Under intermediate precision conditions with M factor(s) different, it is necessary to specify which factors are at state 2 of Table 1 by means of suffixes, for example:

— time-different intermediate precision standard deviation, $s_{I(T)}$;

— calibration-different intermediate precision standard deviation, $s_{I(C)}$;

— operator-different intermediate precision standard deviation, $s_{I(O)}$;

— [time + operator]-different intermediate precision standard deviation, $S_{I(TO)}$;

— [time + operator + equipment]-different intermediate precision standard deviation, $s_{I(TOE)}$;

— and many others in a similar fashion.

Table 1 — Four important factors and their states

Factor	Measurement conditions within a laboratory	
	State 1 (same)	State 2 (different)
Time	Measurements made at the same time	Measurements made at different times
Calibration	No calibration between measurements	Calibration carried out between measurements
Operator	Same operator	Different operators
Equipment	Same equipment without recalibration	Different equipment

6 Statistical model

6.1 Basic model

For estimating the accuracy (trueness and precision) of a measurement method, it is useful to assume that every test result, y, is the sum of three components:

$$y = m + B + e \quad \ldots (1)$$

where, for the particular material tested,

m is the general mean (expectation);

B is the laboratory component of bias under repeatability conditions;

e is the random error occurring in every measurement under repeatability conditions.

A discussion of each of these components, and of extensions of this basic model, follows.

6.2 General mean, m

6.2.1 The general mean, m, is the overall mean of the test results. The value of m obtained in a collaborative study (see ISO 5725-2) depends solely on the "true value" and the measurement method, and does not depend on the laboratory, equipment, operator or time by or at which any test result has been obtained. The general mean of the particular material measured is called the "level of the test"; for example, specimens of different purities of a chemical or different materials (e.g. different types of steel) will correspond to different levels.

In many situations, the concept of a true value μ holds good, such as the true concentration of a solution which is being titrated. The level m is not usually equal to the true value μ; the difference $(m - \mu)$ is called the "bias of the measurement method".

In some situations, the level of the test is exclusively defined by the measurement method, and the concept of an independent true value does not apply; for example, the Vicker's hardness of steel and the Micum indices of coke belong to this category. However, in general, the bias is denoted by δ($\delta = 0$ where no true value exists), then the general mean m is

$$m = \mu + \delta \quad \ldots (2)$$

NOTE 2 Discussion of the bias term δ and a description of trueness experiments are given in ISO 5725-4.

6.2.2 When examining the difference between test results obtained by the same measurement method, the bias of the measurement method may have no influence and can be ignored, unless it is a function of the level of the test. When comparing test results with a value specified in a contract, or a standard value where the contract or specification refers to the true value μ and not to the level of the test m, or when comparing test results obtained using different measurement methods, the bias of the measurement method must be taken into account.

6.3 Term B

6.3.1 B is a term representing the deviation of a laboratory, for one or more reasons, from m, irrespective of the random error e occurring in every test result. Under repeatability conditions in one laboratory, B is considered constant and is called the "laboratory component of bias".

6.3.2 However, when using a measurement method routinely, it is apparent that embodied within an overall value for B are a large number of effects which are due, for example, to changes in the operator, the equipment used, the calibration of the equipment, and the environment (temperature, humidity, air pollution, etc.). The statistical model [equation (1)] can then be rewritten in the form:

$$y = m + B_0 + B_{(1)} + B_{(2)} + \ldots + e \quad \ldots (3)$$

or

$$y = \mu + \delta + B_0 + B_{(1)} + B_{(2)} + \ldots + e \quad \ldots (4)$$

where B is composed of contributions from variates B_0, $B_{(1)}$, $B_{(2)}$ ... and can account for a number of intermediate precision factors.

In practice, the objectives of a study and considerations of the sensitivity of the measurement method will govern the extent to which this model is used. In many cases, abbreviated forms will suffice.

6.4 Terms B_0, $B_{(1)}$, $B_{(2)}$, etc.

6.4.1 Under repeatability conditions, these terms all remain constant and add to the bias of the test results. Under intermediate precision conditions, B_0 is the fixed effect of the factor(s) that remained the same (state 1 of Table 1), while $B_{(1)}$, $B_{(2)}$, etc. are the random effects of the factor(s) which vary (state 2 of Table 1). These no longer contribute to the bias, but increase the intermediate precision standard deviation so that it becomes larger than the repeatability standard deviation.

6.4.2 The effects due to differences between operators include personal habits in operating measurement methods (e.g. in reading graduations on scales, etc.). Some of these differences should be removable by standardization of the measurement method, particularly in having a clear and accurate description of the techniques involved. Even though there is a bias in the test results obtained by an individual operator, that bias is not always constant (e.g. the magnitude of the bias will change according to his/her mental and/or physical conditions on that day) and the bias cannot be corrected or calibrated exactly. The magnitude of such a bias should be reduced by use of a clear operation manual and training. Under such circumstances, the effect of changing operators can be considered to be of a random nature.

6.4.3 The effects due to differences between equipment include the effects due to different places of installation, particularly in fluctuations of the indicator, etc. Some of the effects due to differences between equipment can be corrected by exact calibration. Differences due to systematic causes between equipment should be corrected by calibration, and such a procedure should be included in the standard method. For example, a change in the batch of a reagent could be treated that way. An accepted reference value is needed for this, for which ISO Guide 33 and ISO Guide 35 shall be consulted. The remaining effect due to equipment which has been calibrated using a reference material is considered a random effect.

6.4.4 The effects due to time may be caused by environmental differences, such as changes in room temperature, humidity, etc. Standardization of environmental conditions should be attempted to minimize these effects.

6.4.5 The effect of skill or fatigue of an operator may be considered to be the interaction of operator and time. The performance of a set of equipment may be different at the time of the start of its use and after using it for many hours: this is an example of interaction of equipment and time. When the population of operators is small in number and the population of sets of equipment even smaller, effects caused by these factors may be evaluated as fixed (not random) effects.

6.4.6 The procedures given in ISO 5725-2 are developed assuming that the distribution of laboratory components of bias is approximately normal, but in practice they work for most distributions provided that these distributions are unimodal. The variance of B is called the "between-laboratory variance", expressed as

$$\mathrm{Var}(B) = \sigma_{\mathrm{L}}^2 \quad \ldots (5)$$

However, it will also include effects of changes of operator, equipment, time and environment. From a precision experiment using different operators, measurement times, environments, etc., in a nested design, intermediate precision variances can be calculated. Var(B) is considered to be composed of independent contributions of laboratory, operator, day of experiment, environment, etc.

$$\mathrm{Var}(B) = \mathrm{Var}(B_0) + \mathrm{Var}(B_{(1)}) + \mathrm{Var}(B_{(2)}) + \ldots \quad \ldots (6)$$

The variances are denoted by

$$\mathrm{Var}(B_0) = \sigma_{(0)}^2$$

$$\mathrm{Var}(B_{(1)}) = \sigma_{(1)}^2$$

$$\mathrm{Var}(B_{(2)}) = \sigma_{(2)}^2, \text{ etc.} \quad \ldots (7)$$

Var(B) is estimated in practical terms as s_{L}^2 and similar intermediate precision estimates may be obtained from suitably designed experiments.

6.5 Error term, e

6.5.1 This term represents a random error occurring in every test result and the procedures given throughout this part of ISO 5725 were developed assuming that the distribution of this error variable is approximately normal, but in practice they work for most distributions provided that they are unimodal.

6.5.2 Within a single laboratory, its variance is called the within-laboratory variance and is expressed as

$$\mathrm{Var}(e) = \sigma_{\mathrm{W}}^2 \quad \ldots (8)$$

6.5.3 It may be expected that σ_{W}^2 will have different values in different laboratories due to differences such as in the skills of the operators, but in this part of ISO 5725 it is assumed that, for a properly standardized measurement method, such differences between laboratories should be small and that it is justifiable to establish a common value of within-laboratory variance for all the laboratories using the measurement method. This common value, which is estimated by the mean of the within-laboratory variances, is called the "repeatability variance" and is designated by

$$\sigma_r^2 = \overline{\mathrm{Var}(e)} \quad \ldots (9)$$

This mean value is taken over all the laboratories taking part in the accuracy experiment which remain after outliers have been excluded.

7 Choice of measurement conditions

7.1 In applying a measurement method, many measurement conditions are conceivable within a laboratory, as follows:

a) repeatability conditions (four factors constant);

b) several intermediate precision conditions with one factor different;

c) several intermediate precision conditions with two factors different;

d) several intermediate precision conditions with three factors different;

e) intermediate precision conditions with four factors different.

In the standard for a measurement method, it is not necessary (or even feasible) to state all possible precision measures, although the repeatability standard deviation should always be specified. As regards intermediate precision measures, common commercial practice should indicate the conditions normally encountered, and it should be sufficient to specify only the one suitable intermediate precision measure, together with the detailed stipulation of the specific measurement conditions associated with it. The measurement condition factor(s) to be changed should be carefully defined; in particular, for time-different intermediate precision, a practical mean time interval between successive measurements should be specified.

7.2 It is assumed that a standardized measurement method will be biased as little as possible, and that the bias inherent in the method itself should have been dealt with by technical means. This part of ISO 5725, therefore, deals only with the bias coming from the measurement conditions.

7.3 A change in the factors of the measurement conditions (time, calibration, operator and equipment) from repeatability conditions (i.e. from state 1 to 2 of Table 1) will increase the variability of test results. However, the expectation of the mean of a number of test results will be less biased than under repeatability conditions. The increase in the standard deviation for the intermediate precision conditions may be overcome by not relying on a single test result but by using the mean of several test results as the final quoted result.

7.4 Practical considerations in most laboratories, such as the desired precision (standard deviation) of the final quoted result and the cost of performing the measurements, will govern the number of factors and the choice of the factor(s) whose changes can be studied in the standardization of the measurement method.

8 Within-laboratory study and analysis of intermediate precision measures

8.1 Simplest approach

The simplest method of estimating an intermediate precision standard deviation within one laboratory consists of taking one sample (or, for destructive testing, one set of presumably identical samples) and performing a series of n measurements with a change of factor(s) between each measurement. It is recommended that n should be at least 15. This may not be satisfactory for the laboratory, and this method of estimating intermediate precision measures within a laboratory cannot be regarded as efficient when compared with other procedures. The analysis is simple, however, and it can be useful for studying time-different intermediate precision by making successive measurements on the same sample on successive days, or for studying the effects of calibration between measurements.

A graph of $(y_k - \bar{y})$ versus the measurement number k, where y_k is the k^{th} test result of n replicate test results and $\bar{y}$ is the mean of the n replicate test results, is recommended to identify potential outliers. A more formal test of outliers consists of the application of Grubbs' test as given in subclause **7.3.4** of ISO 5725-2:1994.

The estimate of the intermediate precision standard deviation with M factor(s) different is given by

$$s_{I(\,)} = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(y_k - \bar{y})^2} \qquad \text{... (10)}$$

where symbols denoting the intermediate precision conditions should appear inside the parentheses.

8.2 An alternative method

8.2.1 An alternative method considers t groups of measurements, each comprising n replicate test results. For example, within one laboratory, a set of t materials could each be measured, then the intermediate precision factor(s) could be altered and the t materials remeasured, the procedure being repeated until there are n test results on each of the t materials. Each group of n test results shall be obtained on one identical sample (or set of presumed identical samples in the case of destructive testing), but it is not essential that the materials be identical. It is only required that the t materials all belong to the interval of test levels within which one value of the intermediate precision standard deviation with M factor(s) different can be considered to apply. It is recommended that the value of $t(n-1)$ should be at least 15.

EXAMPLE

One operator performs a single measurement on each of the t materials, then this is repeated by a second operator, and possibly by a third operator, and so on, allowing an estimate of $s_{I(O)}$ to be calculated.

8.2.2 A graph of $(y_{jk} - \bar{y}_j)$ versus the material number j, where y_{jk} is the k^{th} test result on the j^{th} material and $\bar{y}_j$ is the average of the n results on the j^{th} material, is recommended to identify potential outliers. A more formal test of outliers consists of the application of Grubbs' test as given in subclause **7.3.4** of ISO 5725-2:1994 either for each group separately or for all tn test results combined.

The estimate of the intermediate precision standard deviation with M factor(s) different, $s_{I(\ )}$, is then given by

$$s_{I(\ )} = \sqrt{\frac{1}{t(n-1)} \sum_{j=1}^{t} \sum_{k=1}^{n} (y_{jk} - \bar{y}_j)^2} \qquad \text{... (11)}$$

For $n = 2$ (i.e. two test results on each material), the formula simplifies to

$$s_{I(\ )} = \sqrt{\frac{1}{2t} \sum_{j=1}^{t} (y_{j1} - y_{j2})^2} \qquad \text{... (12)}$$

8.3 Effect of the measurement conditions on the final quoted result

8.3.1 The expectation of $\bar{y}$ is different between one combination and another of time, calibration, operator and equipment, even when only one of the four factors changes. This is a limitation on the usefulness of mean values. In chemical analysis or physical testing, $\bar{y}$ is reported as the final quoted result. In trading raw materials, this final quoted result is often used for quality evaluation of the raw materials and affects the price of the product to a considerable extent.

EXAMPLE

In the international trading of coal, the size of the consignment can often exceed 70 000 t, and the ash content is determined finally on a test portion of only 1 g. In a contract stipulating that each difference of 1 % in ash content corresponds to USD 1,5 per tonne of coal, a difference of 1 mg in the weighing of ash by a chemical balance corresponds to 0,1 % in ash content, or USD 0,15 per tonne, which for such a consignment amounts to a difference in proceeds of USD 10 500 (from 0,1 × 1,5 × 70 000).

8.3.2 Consequently, the final quoted result of chemical analysis or physical testing should be sufficiently precise, highly reliable and, especially, universal and reproducible. A final quoted result which can be guaranteed only under conditions of a specific operator, equipment or time may not be good enough for commercial considerations.

9 Interlaboratory study and analysis of intermediate precision measures

9.1 Underlying assumptions

Estimation of intermediate measures of precision from interlaboratory studies relies on the assumption that the effect of a particular factor is the same across all laboratories, so that, for example, changing operators in one laboratory has the same effect as changing operators in another laboratory, or that variation due to time is the same across all laboratories. If this assumption is violated, then the concept of intermediate measures of precision does not make sense, nor do the techniques proposed in the subsequent sections to estimate these intermediate measures of precision. Careful attention to outliers (not necessarily deletion of outliers) must be paid as this will help in detecting departure from the assumptions necessary to pool information from all laboratories. One powerful technique to detect potential outliers is to depict the measurements graphically as a function of the various levels of the factors or the various laboratories included in the study.

9.2 Simplest approach

If material at q levels is sent to p laboratories who each perform measurements on each of the q levels with a change of intermediate precision factor(s) between each of the n measurements, then the analysis is by the same method of calculation as explained in ISO 5725-2, except that an intermediate precision standard deviation is estimated instead of the repeatability standard deviation.

9.3 Nested experiments

A further way of estimating intermediate precision measures is to conduct more sophisticated experiments. These can be fully-nested or staggered-nested experiments (for definitions of these terms, see ISO 3534-3). The advantage of employing a nested experimental design is that it is possible, at one time and in one interlaboratory experiment, to estimate not only repeatability and reproducibility standard deviations but also one or more intermediate precision standard deviations. There are, however, certain caveats which must be considered, as will be explained in **9.8**.

9.4 Fully-nested experiment

A schematic layout of the fully-nested experiment at a particular level of the test is given in Figure 1.

By carrying out the three-factor fully-nested experiment collaboratively in several laboratories, one intermediate precision measure can be obtained at the same time as the repeatability and reproducibility standard deviations, i.e. $\sigma_{(0)}$, $\sigma_{(1)}$ and σ_r can be estimated. Likewise the four-factor fully-nested experiment can be used to obtain two intermediate precision measures, i.e. $\sigma_{(0)}$, $\sigma_{(1)}$, $\sigma_{(2)}$ and σ_r can be estimated.

The subscripts i, j and k suffixed to the data y in Figure 1a) for the three-factor fully-nested experiment represent, for example, a laboratory, a day of experiment and a replication under repeatability conditions, respectively.

The subscripts i, j, k and l suffixed to the data y in Figure 1b) for the four-factor fully-nested experiment represent, for example, a laboratory, a day of experiment, an operator and a replication under repeatability conditions, respectively.

Analysis of the results of an n-factor fully-nested experiment is carried out by the statistical technique "analysis of variance" (ANOVA) separately for each level of the test, and is described in detail in Annex B.

9.5 Staggered-nested experiment

A schematic layout of the staggered-nested experiment at a particular level of the test is given in Figure 2.

The three-factor staggered-nested experiment requires each laboratory i to obtain three test results. Test results y_{i1} and y_{i2} shall be obtained under repeatability conditions, and y_{i3} under intermediate precision conditions with M factor(s) different (M = 1, 2 or 3), for example under time-different intermediate precision conditions (by obtaining y_{i3} on a different day from that on which y_{i1} and y_{i2} were obtained).

In a four-factor staggered-nested experiment, y_{i4} shall be obtained under intermediate precision conditions with one more factor different, for example, under [time + operator]-different intermediate precision conditions by changing the day and the operator.

Again, analysis of the results of an n-factor staggered-nested experiment is carried out by the statistical technique "analysis of variance" (ANOVA) separately for each level of the test, and is described in detail in Annex C.

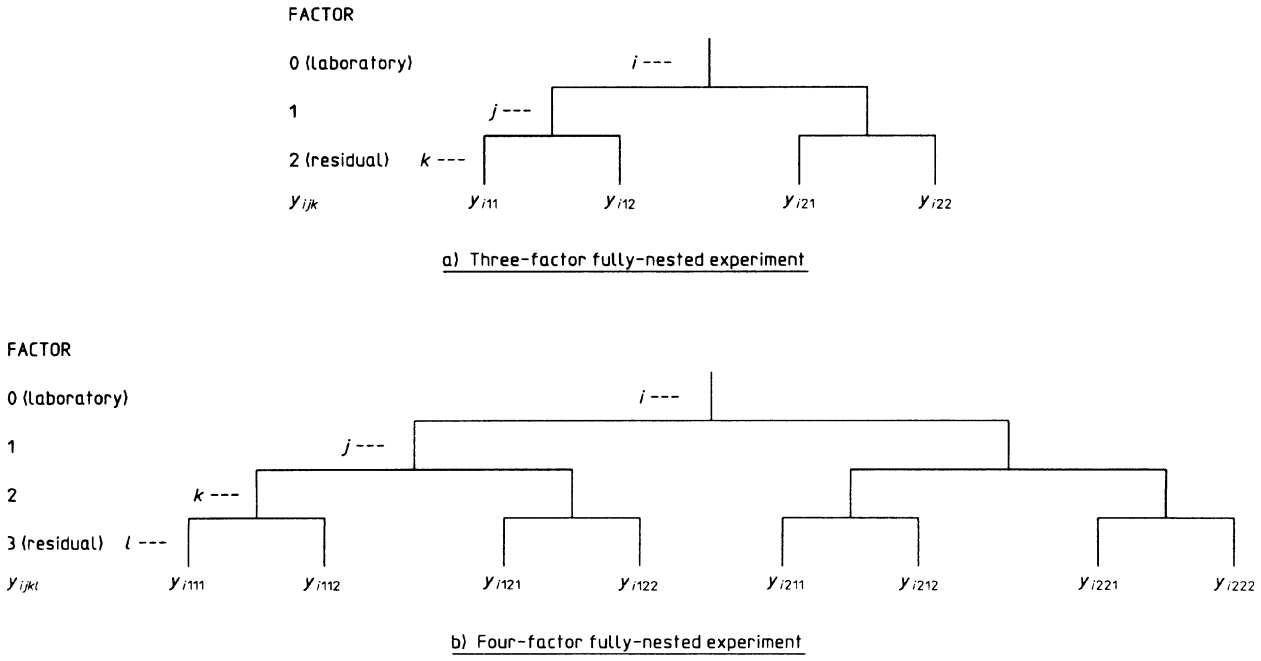

Figure 1 — Schematic layouts for three-factor and four-factor fully-nested experiments

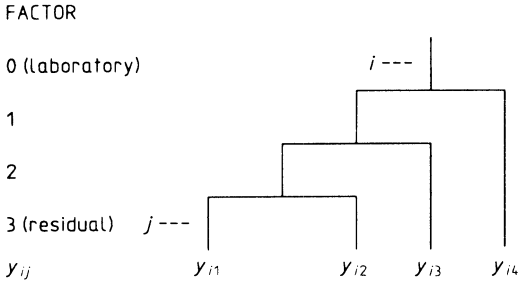

Figure 2 — Schematic layout of a four-factor staggered-nested experiment

9.6 Allocation of factors in a nested experimental design

The allocation of the factors in a nested experimental design is arranged so that the factors affected most by systematic effects should be in the highest ranks (0, 1, ...), and those affected most by random effects should be in the lowest ranks, the lowest factor being considered as a residual variation. For example, in a four-factor design such as illustrated in Figure 1 b) and Figure 2, factor 0 could be the laboratory, factor 1 the operator, factor 2 the day on which the measurement is carried out, and factor 3 the replication. This may not seem important in the case of the fully-nested experiment due to its symmetry.

9.7 Comparison of the nested design with the procedure given in ISO 5725-2

The procedure given in ISO 5725-2, because the analysis is carried out separately for each level of the test (material), is, in fact, a two-factor fully-nested experimental design and produces two standard deviations, the repeatability and reproducibility standard deviations. Factor 0 is the laboratory and factor 1 the replication. If this design were increased by one factor, by having two operators in each laboratory each obtaining two test results under repeatability conditions, then, in addition to the repeatability and reproducibility standard deviations, one could determine the operator-different intermediate precision standard deviation. Alternatively, if each laboratory used only one operator but repeated the experiment on another day, the time-different intermediate precision standard deviation would be determined by this three-factor fully-nested experiment. The addition of a further factor to the experiment, by each laboratory having two operators each carrying out two measurements and the whole experiment being repeated the next day, would allow determination of the repeatability, reproducibility, operator-different, time-different, and [time + operator]-different standard deviations.

9.8 Comparison of fully-nested and staggered-nested experimental designs

An n-factor fully-nested experiment requires 2^{n-1} test results from each laboratory, which can be an excessive requirement on the laboratories. This is the main argument for the staggered-nested experimental design. This design requires less test results to produce the same number of standard deviations, although the analysis is slightly more complex and there is a larger uncertainty in the estimates of the standard deviations due to the smaller number of test results.

Annex A (normative)
Symbols and abbreviations used in ISO 5725

a	Intercept in the relationship $s = a + bm$
A	Factor used to calculate the uncertainty of an estimate
b	Slope in the relationship $s = a + bm$
B	Component in a test result representing the deviation of a laboratory from the general average (laboratory component of bias)
B_0	Component of B representing all factors that do not change in intermediate precision conditions
$B_{(1)}$, $B_{(2)}$, etc.	Components of B representing factors that vary in intermediate precision conditions
c	Intercept in the relationship $\lg s = c + d \lg m$
C, C', C''	Test statistics
C_{crit}, C'_{crit}, C''_{crit}	Critical values for statistical tests
CD_P	Critical difference for probability P
CR_P	Critical range for probability P
d	Slope in the relationship $\lg s = c + d \lg m$
e	Component in a test result representing the random error occurring in every test result
f	Critical range factor
$F_p(v_1, v_2)$	p-quantile of the F-distribution with v_1 and v_2 degrees of freedom
G	Grubbs' test statistic
h	Mandel's between-laboratory consistency test statistic
k	Mandel's within-laboratory consistency test statistic
LCL	Lower control limit (either action limit or warning limit)
m	General mean of the test property; level
M	Number of factors considered in intermediate precision conditions
N	Number of iterations
n	Number of test results obtained in one laboratory at one level (i.e. per cell)
p	Number of laboratories participating in the interlaboratory experiment
P	Probability
q	Number of levels of the test property in the interlaboratory experiment
r	Repeatability limit
R	Reproducibility limit
RM	Reference material
s	Estimate of a standard deviation
$\hat{s}$	Predicted standard deviation
T	Total or sum of some expression
t	Number of test objects or groups
UCL	Upper control limit (either action limit or warning limit)
W	Weighting factor used in calculating a weighted regression
w	Range of a set of test results
x	Datum used for Grubbs' test

y	Test result
$\bar{y}$	Arithmetic mean of test results
$\bar{\bar{y}}$	Grand mean of test results
α	Significance level
β	Type II error probability
γ	Ratio of the reproducibility standard deviation to the repeatability standard deviation (σ_R/σ_r)
Δ	Laboratory bias
$\hat{\Delta}$	Estimate of Δ
δ	Bias of the measurement method
$\hat{\delta}$	Estimate of δ
λ	Detectable difference between two laboratory biases or the biases of two measurement methods
μ	True value or accepted reference value of a test property
v	Number of degrees of freedom
ϱ	Detectable ratio between the repeatability standard deviations of method B and method A
σ	True value of a standard deviation
τ	Component in a test result representing the variation due to time since last calibration
ϕ	Detectable ratio between the square roots of the between-laboratory mean squares of method B and method A
$x_p^2(v)$	p-quantile of the χ^2-distribution with v degrees of freedom

Symbols used as subscripts

C	Calibration-different
E	Equipment-different
i	Identifier for a particular laboratory
$I(\ )$	Identifier for intermediate measures of precision; in brackets, identification of the type of intermediate situation
j	Identifier for a particular level (ISO 5725-2).
	Identifier for a group of tests or for a factor (ISO 5725-3)
k	Identifier for a particular test result in a laboratory i at level j
L	Between-laboratory (interlaboratory)
m	Identifier for detectable bias
M	Between-test-sample
O	Operator-different
P	Probability
r	Repeatability
R	Reproducibility
T	Time-different
W	Within-laboratory (intralaboratory)
1, 2, 3...	For test results, numbering in the order of obtaining them
(1), (2), (3)...	For test results, numbering in the order of increasing magnitude

Annex B (normative)
Analysis of variance for fully-nested experiments

The analysis of variance described in this annex has to be carried out separately for each level of the test included in the interlaboratory experiment. For simplicity, a subscript indicating the level of the test has not been suffixed to the data. It should be noted that the subscript j is used in this part of ISO 5725 for factor 1 (factor 0 being the laboratory), while in the other parts of ISO 5725 it is used for the level of the test.

The methods described in subclause **7.3** of ISO 5725-2:1994 should be applied to check the data for consistency and outliers. With the designs described in this annex, the exact analysis of the data is very complicated when some of the test results from a laboratory are missing. If it is decided that some of the test results from a laboratory are stragglers or outliers and should be excluded from the analysis, then it is recommended that all the data from that laboratory (at the level affected) should be excluded from the analysis.

B.1 Three-factor fully-nested experiment

The data obtained in the experiment are denoted by y_{ijk}, and the mean values and ranges are as follows:

$$\bar{y}_{ij} = \frac{1}{2}(y_{ij1} + y_{ij2}) \qquad \bar{y}_i = \frac{1}{2}(\bar{y}_{i1} + \bar{y}_{i2}) \qquad \bar{\bar{y}} = \frac{1}{p}\sum_i \bar{y}_i$$

$$w_{ij(1)} = |y_{ij1} - y_{ij2}| \qquad w_{i(2)} = |\bar{y}_{i1} - \bar{y}_{i2}|$$

where p is the number of laboratories which have participated in the interlaboratory experiment.

The total sum of squares, SST, can be subdivided as

$$\mathrm{SST} = \sum_i \sum_j \sum_k (y_{ijk} - \bar{\bar{y}})^2 = \mathrm{SS0} + \mathrm{SS1} + \mathrm{SSe}$$

where

$$\mathrm{SS0} = \sum_i \sum_j \sum_k (\bar{y}_i - \bar{\bar{y}})^2 = 4\sum_i (\bar{y}_i - \bar{\bar{y}})^2 = 4\sum_i (\bar{y}_i)^2 - 4p(\bar{\bar{y}})^2$$

$$\mathrm{SS1} = \sum_i \sum_j \sum_k (\bar{y}_{ij} - \bar{y}_i)^2 = 2\sum_i \sum_j (\bar{y}_{ij} - \bar{y}_i)^2 = \sum_i w_{i(2)}^2$$

$$\mathrm{SSe} = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij})^2 = \frac{1}{2}\sum_i \sum_j w_{ij(1)}^2$$

Since the degrees of freedom for the sums of squares SS0, SS1 and SSe are $p - 1$, p and $2p$, respectively, the ANOVA table is composed as shown in Table B.1.

Table B.1 — ANOVA table for a three-factor fully-nested experiment

Source	Sum of squares	Degrees of freedom	Mean square	Expected mean square
0	SS0	$p - 1$	MS0 = SS0/($p - 1$)	$\sigma_r^2 + 2\sigma_{(1)}^2 + 4\sigma_{(0)}^2$
1	SS1	p	MS1 = SS1/p	$\sigma_r^2 + 2\sigma_{(1)}^2$
Residual	SSe	$2p$	MSe = SSe/($2p$)	σ_r^2
Total	SST	$4p - 1$		

The unbiased estimates $s^2_{(0)}$, $s^2_{(1)}$ and s^2_r of $\sigma^2_{(0)}$, $\sigma^2_{(1)}$ and σ^2_r, respectively, can be obtained from the mean squares MS0, MS1 and MSe as

$$s^2_{(0)} = \frac{1}{4}(\text{MS0} - \text{MS1})$$

$$s^2_{(1)} = \frac{1}{2}(\text{MS1} - \text{MSe})$$

$$s^2_r = \text{MSe}$$

The estimates of the repeatability variance, one-factor-different intermediate precision variance and reproducibility variance are, respectively, as follows:

$$s^2_r$$

$$s^2_{I(1)} = s^2_r + s^2_{(1)}$$

$$s^2_R = s^2_r + s^2_{(1)} + s^2_{(0)}$$

B.2 Four-factor fully-nested experiment

The data obtained in the experiment are denoted by y_{ijkl}, and mean values and ranges are as follows:

$$\bar{y}_{ijk} = \frac{1}{2}(y_{ijk1} + y_{ijk2}) \qquad w_{ijk(1)} = |y_{ijk1} - y_{ijk2}|$$

$$\bar{y}_{ij} = \frac{1}{2}(\bar{y}_{ij1} + \bar{y}_{ij2}) \qquad w_{ij(2)} = |\bar{y}_{ij1} - \bar{y}_{ij2}|$$

$$\bar{y}_i = \frac{1}{2}(\bar{y}_{i1} + \bar{y}_{i2}) \qquad w_{i(3)} = |\bar{y}_{i1} - \bar{y}_{i2}|$$

$$\bar{\bar{y}} = \frac{1}{p}\sum_i \bar{y}_i$$

where p is the number of laboratories which have participated in the interlaboratory experiment.

The total sum of squares, SST, can be subdivided as follows:

$$\text{SST} = \sum_i \sum_j \sum_k \sum_l (y_{ijkl} - \bar{\bar{y}})^2 = \text{SS0} + \text{SS1} + \text{SS2} + \text{SSe}$$

where

$$\text{SS0} = \sum_i \sum_j \sum_k \sum_l (\bar{y}_i - \bar{\bar{y}})^2 = 8\sum_i (\bar{y}_i)^2 - 8p(\bar{\bar{y}})^2$$

$$\text{SS1} = \sum_i \sum_j \sum_k \sum_l (\bar{y}_{ij} - \bar{y}_i)^2 = 4\sum_i \sum_j (\bar{y}_{ij} - \bar{y}_i)^2 = 2\sum_i w^2_{i(3)}$$

$$\mathrm{SS2} = \sum_i \sum_j \sum_k \sum_l (\bar{y}_{ijk} - \bar{y}_{ij})^2 = 2\sum_i \sum_j \sum_k (\bar{y}_{ijk} - \bar{y}_{ij})^2 = \sum_i \sum_j w_{ij(2)}^2$$

$$\mathrm{SSe} = \sum_i \sum_j \sum_k \sum_l (y_{ijkl} - \bar{y}_{ijk})^2 = \frac{1}{2}\sum_i \sum_j \sum_k w_{ijk(1)}^2$$

Since the degrees of freedom for the sums of squares SS0, SS1, SS2 and SSe are $p - 1$, p, $2p$ and $4p$, respectively, the ANOVA table is composed as shown in Table B.2.

Table B.2 — ANOVA table for a four-factor fully-nested experiment

Source	Sum of squares	Degrees of freedom	Mean square	Expected mean square
0	SS0	$p - 1$	MS0 = SS0/($p - 1$)	$\sigma_r^2 + 2\sigma_{(2)}^2 + 4\sigma_{(1)}^2 + 8\sigma_{(0)}^2$
1	SS1	p	MS1 = SS1/p	$\sigma_r^2 + 2\sigma_{(2)}^2 + 4\sigma_{(1)}^2$
2	SS2	$2p$	MS2 = SS2/($2p$)	$\sigma_r^2 + 2\sigma_{(2)}^2$
Residual	SSe	$4p$	MSe = SSe/($4p$)	σ_r^2
Total	SST	$8p - 1$		

The unbiased estimates $s_{(0)}^2$, $s_{(1)}^2$, $s_{(2)}^2$ and s_r^2 of $\sigma_{(0)}^2$, $\sigma_{(1)}^2$, $\sigma_{(1)}^2$ and σ_r^2, respectively, can be obtained from the mean squares MS0, MS1, MS2 and MSe as follows:

$$s_{(0)}^2 = \frac{1}{8}(\mathrm{MS0} - \mathrm{MS1})$$

$$s_{(1)}^2 = \frac{1}{4}(\mathrm{MS1} - \mathrm{MS2})$$

$$s_{(2)}^2 = \frac{1}{2}(\mathrm{MS2} - \mathrm{MSe})$$

$$s_r^2 = \mathrm{MSe}$$

The estimates of the repeatability variance, one-factor-different intermediate precision variance, two-factors different intermediate precision variance and reproducibility variance are, respectively, as follows:

$$s_r^2$$

$$s_{I(1)}^2 = s_r^2 + s_{(2)}^2$$

$$s_{I(2)}^2 = s_r^2 + s_{(2)}^2 + s_{(1)}^2$$

$$s_R^2 = s_r^2 + s_{(2)}^2 + s_{(1)}^2 + s_{(0)}^2$$

Annex C (normative)
Analysis of variance for staggered-nested experiments

The analysis of variance described in this annex has to be carried out separately for each level of the test included in the interlaboratory experiment. For simplicity, a subscript indicating the level of the test has not been suffixed to the data. It should be noted that the subscript j is used in this part of ISO 5725 for the replications within the laboratory, while in the other parts of ISO 5725 it is used for the level of the test.

The methods described in subclause **7.3** of ISO 5725-2:1994 should be applied to check the data for consistency and outliers. With the designs described in this annex, the exact analysis of the data is very complicated when some of the test results from a laboratory are missing. If it is decided that some of the test results from a laboratory are stragglers or outliers and should be excluded from the analysis, then it is recommended that all the data from that laboratory (at the level affected) should be excluded from the analysis.

C.1 Three-factor staggered-nested experiment

The data obtained in the experiment within laboratory i are denoted by y_{ij} ($j = 1, 2, 3$), and the mean values and ranges are as follows:

$$\bar{y}_{i(1)} = \frac{1}{2}(y_{i1} + y_{i2}) \qquad w_{i(1)} = |y_{i1} - y_{i2}|$$

$$\bar{y}_{i(2)} = \frac{1}{3}(y_{i1} + y_{i2} + y_{i3}) \qquad w_{i(2)} = |\bar{y}_{i(1)} - y_{i3}|$$

$$\bar{\bar{y}} = \frac{1}{p}\sum_i \bar{y}_{i(2)}$$

where p is the number of laboratories which have participated in the interlaboratory experiment.

The total sum of squares, SST, can be subdivided as follows:

$$\mathrm{SST} = \sum_i \sum_j (y_{ij} - \bar{\bar{y}})^2 = \mathrm{SS0} + \mathrm{SS1} + \mathrm{SSe}$$

where

$$\mathrm{SS0} = 3\sum_i (\bar{y}_{i(2)})^2 - 3p(\bar{\bar{y}})^2$$

$$\mathrm{SS1} = \frac{2}{3}\sum_i w_{i(2)}^2$$

$$\mathrm{SSe} = \frac{1}{2}\sum_i w_{i(1)}^2$$

Since the degrees of freedom for the sums of squares SS0, SS1 and SSe are $p - 1$, p and p, respectively, the ANOVA table is composed as shown in Table C.1.

Table C.1 — ANOVA table for a three-factor staggered-nested experiment

Source	Sum of squares	Degrees of freedom	Mean square	Expected mean square
0	SS0	$p - 1$	SS0/($p - 1$)	$\sigma_r^2 + \frac{5}{3}\ \sigma_{(1)}^2 + 3\sigma_{(0)}^2$
1	SS1	p	SS1/p	$\sigma_r^2 + \frac{4}{3}\ \sigma_{(1)}^2$
Residual	SSe	p	SSe/p	σ_r^2
Total	SST	$3p - 1$		

The unbiased estimates $s_{(0)}^2$, $s_{(1)}^2$ and s_r^2 of $\sigma_{(0)}^2$, $\sigma_{(1)}^2$ and σ_r^2, respectively, can be obtained from the mean squares MS0, MS1 and MSe as follows:

$$s_{(0)}^2 = \frac{1}{3}\ \mathrm{MS0} - \frac{5}{12}\ \mathrm{MS1} + \frac{1}{12}\ \mathrm{MSe}$$

$$s_{(1)}^2 = \frac{3}{4}\ \mathrm{MS1} - \frac{3}{4}\ \mathrm{MSe}$$

$$s_r^2 = \mathrm{MSe}$$

The estimates of the repeatability variance, one-factor-different intermediate precision variance and reproducibility variance are, respectively, as follows:

$$s_r^2$$

$$s_{I(1)}^2 = s_r^2 + s_{(1)}^2$$

$$s_R^2 = s_r^2 + s_{(1)}^2 + s_{(0)}^2$$

C.2 Four-factor staggered-nested experiment

The data obtained in the experiment within laboratory i are denoted by y_{ij} (j = 1, 2, 3, 4), and the mean values and ranges are as follows:

$$\bar{y}_{i(1)} = \frac{1}{2}\ (y_{i1} + y_{i2}) \qquad w_{i(1)} = |y_{i1} - y_{i2}|$$

$$\bar{y}_{i(2)} = \frac{1}{3}\ (y_{i1} + y_{i2} + y_{i3}) \qquad w_{i(2)} = |\bar{y}_{i(1)} - y_{i3}|$$

$$\bar{y}_{i(3)} = \frac{1}{4}\ (y_{i1} + y_{i2} + y_{i3} + y_{i4}) \qquad w_{i(3)} = |\bar{y}_{i(2)} - y_{i4}|$$

$$\bar{\bar{y}} = \frac{1}{p} \sum_i \bar{y}_{i(3)}$$

where p is the number of laboratories which have participated in the interlaboratory experiment.

The ANOVA table is composed as shown in Table C.2.

Table C.2 — ANOVA table for a four-factor staggered-nested experiment

Source	Sum of squares	Degrees of freedom	Mean square	Expected mean square
0	$4\sum_i (\bar{y}_{i(3)})^2 - 4p(\bar{\bar{y}})^2$	$p-1$	SS0/($p-1$)	$\sigma_r^2 + \frac{3}{2}\sigma_{(2)}^2 + \frac{5}{2}\sigma_{(1)}^2 + 4\sigma_{(0)}^2$
1	$\frac{3}{4}\sum_i w_{i(3)}^2$	p	SS1/p	$\sigma_r^2 + \frac{7}{6}\sigma_{(2)}^2 + \frac{3}{2}\sigma_{(1)}^2$
2	$\frac{2}{3}\sum_i w_{i(2)}^2$	p	SS2/p	$\sigma_r^2 + \frac{4}{3}\sigma_{(2)}^2$
Residual	$\frac{1}{2}\sum_i w_{i(1)}^2$	p	SSe/p	σ_r^2
Total	$\sum_i\sum_j (y_{ij} - \bar{\bar{y}})^2$	$4p-1$		

C.3 Five-factor staggered-nested experiment

The data obtained in the experiment within laboratory i are denoted by y_{ij} (j = 1, 2, 3, 4, 5), and the mean values and ranges are as follows:

$$\bar{y}_{i(1)} = \frac{1}{2}(y_{i1} + y_{i2}) \qquad w_{i(1)} = |y_{i1} - y_{i2}|$$

$$\bar{y}_{i(2)} = \frac{1}{3}(y_{i1} + y_{i2} + y_{i3}) \qquad w_{i(2)} = |\bar{y}_{i(1)} - y_{i3}|$$

$$\bar{y}_{i(3)} = \frac{1}{4}(y_{i1} + y_{i2} + y_{i3} + y_{i4}) \qquad w_{i(3)} = |\bar{y}_{i(2)} - y_{i4}|$$

$$\bar{y}_{i(4)} = \frac{1}{5}(y_{i1} + y_{i2} + y_{i3} + y_{i4} + y_{i5}) \qquad w_{i(4)} = |\bar{y}_{i(3)} - y_{i5}|$$

$$\bar{\bar{y}} = \frac{1}{p}\sum_i \bar{y}_{i(4)}$$

where p is the number of laboratories which have participated in the interlaboratory experiment.

The ANOVA table is composed as shown in Table C.3.

Table C.3 — ANOVA table for a five-factor staggered-nested experiment

Source	Sum of squares	Degrees of freedom	Mean square	Expected mean square
0	$5\sum_i (\bar{y}_{i(4)})^2 - 5p(\bar{\bar{y}})^2$	$p-1$	SS0/($p-1$)	$\sigma_r^2 + \frac{7}{5}\sigma_{(3)}^2 + \frac{11}{5}\sigma_{(2)}^2 + \frac{17}{5}\sigma_{(1)}^2 + 5\sigma_{(0)}^2$
1	$\frac{4}{5}\sum_i w_{i(4)}^2$	p	SS1/p	$\sigma_r^2 + \frac{11}{10}\sigma_{(3)}^2 + \frac{13}{10}\sigma_{(2)}^2 + \frac{8}{5}\sigma_{(1)}^2$
2	$\frac{3}{4}\sum_i w_{i(3)}^2$	p	SS2/p	$\sigma_r^2 + \frac{7}{6}\sigma_{(3)}^2 + \frac{3}{2}\sigma_{(2)}^2$
3	$\frac{2}{3}\sum_i w_{i(2)}^2$	p	SS3/p	$\sigma_r^2 + \frac{4}{3}\sigma_{(3)}^2$
Residual	$\frac{1}{2}\sum_i w_{i(1)}^2$	p	SSe/p	σ_r^2
Total	$\sum_i\sum_j (y_{ij} - \bar{\bar{y}})^2$	$5p-1$		

C.4 Six-factor staggered-nested experiment

The data obtained in the experiment within laboratory i are denoted by y_{ij} (j = 1, 2, 3, 4, 5, 6), and the mean values and ranges are as follows:

$$\bar{y}_{i(1)} = \frac{1}{2}(y_{i1} + y_{i2}) \qquad w_{i(1)} = |y_{i1} - y_{i2}|$$

$$\bar{y}_{i(2)} = \frac{1}{3}(y_{i1} + y_{i2} + y_{i3}) \qquad w_{i(2)} = |\bar{y}_{i(1)} - y_{i3}|$$

$$\bar{y}_{i(3)} = \frac{1}{4}(y_{i1} + y_{i2} + y_{i3} + y_{i4}) \qquad w_{i(3)} = |\bar{y}_{i(2)} - y_{i4}|$$

$$\bar{y}_{i(4)} = \frac{1}{5}(y_{i1} + y_{i2} + y_{i3} + y_{i4} + y_{i5}) \qquad w_{i(4)} = |\bar{y}_{i(3)} - y_{i5}|$$

$$\bar{y}_{i(5)} = \frac{1}{6}(y_{i1} + y_{i2} + y_{i3} + y_{i4} + y_{i5} + y_{i6}) \qquad w_{i(5)} = |\bar{y}_{i(4)} - y_{i6}|$$

$$\bar{\bar{y}} = \frac{1}{p}\sum_i \bar{y}_{i(5)}$$

where p is the number of laboratories which have participated in the interlaboratory experiment.

The ANOVA table is composed as shown in Table C.4.

Table C.4 — ANOVA table for a six-factor staggered-nested experiment

Source	Sum of squares	Degrees of freedom	Mean square	Expected mean square
0	$6\sum_i(\bar{y}_{i(5)})^2 - 6p(\bar{\bar{y}})^2$	$p-1$	SS0/($p-1$)	$\sigma_r^2 + \frac{4}{3}\sigma_{(4)}^2 + 2\sigma_{(3)}^2 + 3\sigma_{(2)}^2 + \frac{13}{3}\sigma_{(1)}^2 + 6\sigma_{(0)}^2$
1	$\frac{5}{6}\sum_i w_{i(5)}^2$	p	SS1/p	$\sigma_r^2 + \frac{16}{15}\sigma_{(4)}^2 + \frac{6}{5}\sigma_{(3)}^2 + \frac{7}{5}\sigma_{(2)}^2 + \frac{5}{3}\sigma_{(1)}^2$
2	$\frac{4}{5}\sum_i w_{i(4)}^2$	p	SS2/p	$\sigma_r^2 + \frac{11}{10}\sigma_{(4)}^2 + \frac{13}{10}\sigma_{(3)}^2 + \frac{8}{5}\sigma_{(2)}^2$
3	$\frac{3}{4}\sum_i w_{i(3)}^2$	p	SS3/p	$\sigma_r^2 + \frac{7}{6}\sigma_{(4)}^2 + \frac{3}{2}\sigma_{(3)}^2$
4	$\frac{2}{3}\sum_i w_{i(2)}^2$	p	SS4/p	$\sigma_r^2 + \frac{4}{3}\sigma_{(4)}^2$
Residual	$\frac{1}{2}\sum_i w_{i(1)}^2$	p	SSe/p	σ_r^2
Total	$\sum_i\sum_j (y_{ij} - \bar{\bar{y}})^2$	$6p-1$		

Annex D (informative)
Examples of the statistical analysis of intermediate precision experiments

D.1 Example 1 — Obtaining the [time + operator]-different intermediate precision standard deviation, $s_{I(TO)}$, within a specific laboratory at a particular level of the test

D.1.1 *Background*

a) **Measurement method:** Determination of the carbon content in steel by vacuum emission spectrometry with test results expressed as percentages by mass.

b) **Source:** Routine report of a steel works in November 1984.

c) **Experimental design:** In a specific laboratory, a randomly selected sample out of the materials analysed on one day was analysed again on the next day by a different analyst. In a month 29 pairs of such data were obtained (see Table D.1).

D.1.2 ***Analysis***

The data y_{j1}, y_{j2} and $w_j = |y_{j1} - y_{j2}|$ are shown in Table D.1. The analysis follows the procedure given in **8.2**. A plot of the data [deviations from the mean of the measurements on both days $(y_{jk} - \bar{y}_j)$ versus the sample number j] is shown in Figure D.1. This plot and the application of Cochran's test indicate that the ranges for samples with numbers 20 and 24 are outliers. There is a large discrepancy between the daily measurements on these samples which is most likely due to errors in recording the data. The values for these two samples were removed from the computation of the [time + operator]-different intermediate precision standard deviation, $s_{I(TO)}$, which is calculated according to equation (12) as

$$s_{I(TO)} = \sqrt{\frac{1}{2 \times 27} \sum_{j=1}^{27} w_j^2} = 2{,}87 \times 10^{-3}$$

Table D.1 — Original data — Carbon content, % (m/m)

Sample No. j	**First day** y_{j1}	**Next day** y_{j2}	**Range** w_j
1	0,130	0,127	0,003
2	0,140	0,132	0,008
3	0,078	0,080	0,002
4	0,110	0,113	0,003
5	0,126	0,128	0,002
6	0,036	0,032	0,004
7	0,050	0,047	0,003
8	0,143	0,140	0,003
9	0,091	0,089	0,002
10	0,040	0,030	0,010
11	0,110	0,113	0,003
12	0,142	0,145	0,003
13	0,143	0,143	0,007
14	0,169	0,165	0,004
15	0,169	0,173	0,004
16	0,149	0,144	0,005
17	0,044	0,044	0,000
18	0,127	0,122	0,005
19	0,050	0,048	0,002
20	0,042	0,146	0,104
21	0,150	0,145	0,005
22	0,135	0,133	0,002
23	0,044	0,045	0,001
24	0,100	0,161	0,061
25	0,132	0,131	0,001
26	0,047	0,045	0,002
27	0,168	0,165	0,003
28	0,092	0,088	0,004
29	0,041	0,043	0,002

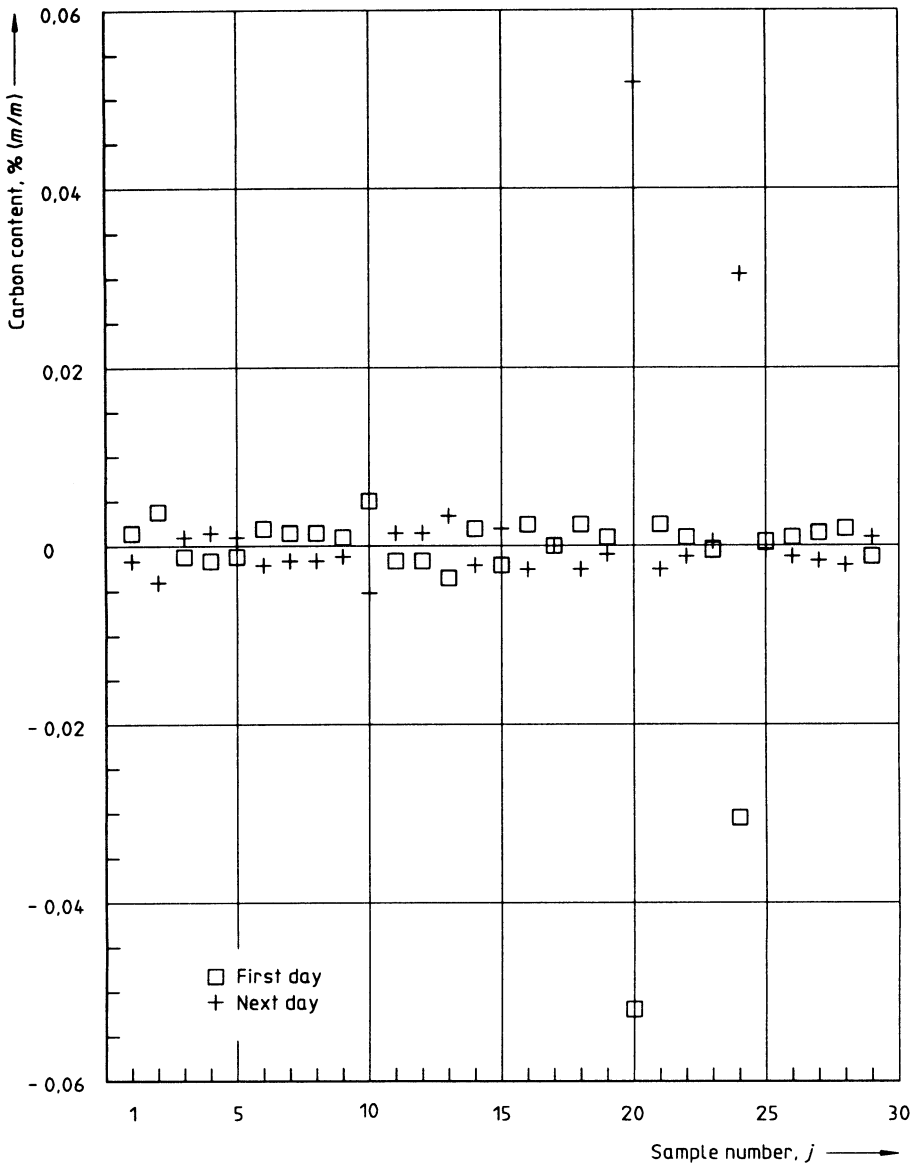

Figure D.1 — Carbon content in steel — Deviations from the mean of the measurements on both days versus the sample number

D.2 Example 2 — Obtaining the time-different intermediate precision standard deviation by interlaboratory experiment

D.2.1 *Background*

a) **Measurement method:** Determination of the vanadium content in steel by the atomic absorption spectrometric method described in the instructions for the experiment. Test results are expressed as percentages by mass.

b) **Source:** ISO/TC 17, *Steel*/SC 1, *Methods of determination of chemical composition.* Experiment carried out in May 1985.

c) **Experimental design:** A three-factor staggered-nested experiment was carried out with 20 laboratories each reporting two test results obtained under repeatability conditions on one day, followed by one further test result on the next day at each of the six levels included in the experiment. All measurements in any laboratory were carried out by one operator, using the same measurement equipment.

D.2.2 *Analysis*

The data for all six levels are given in Table D.2.

The analysis of variance is presented for only one of the levels, i.e. level 1.

A plot of the data (test results for day 1 and day 2 versus laboratory number i) is shown in Figure D.2. This plot indicates that laboratory 20 is an outlier laboratory. There is a large discrepancy between the test result for day 2 and the mean value for day 1, which is very large compared with the test results for the other laboratories. This laboratory was removed from the calculations of the precision measures.

In accordance with **C.1** of Annex C, $w_{i(1)}$, $w_{i(2)}$ and $\bar{y}_{i(2)}$ were computed and the results are shown in Table D.3.

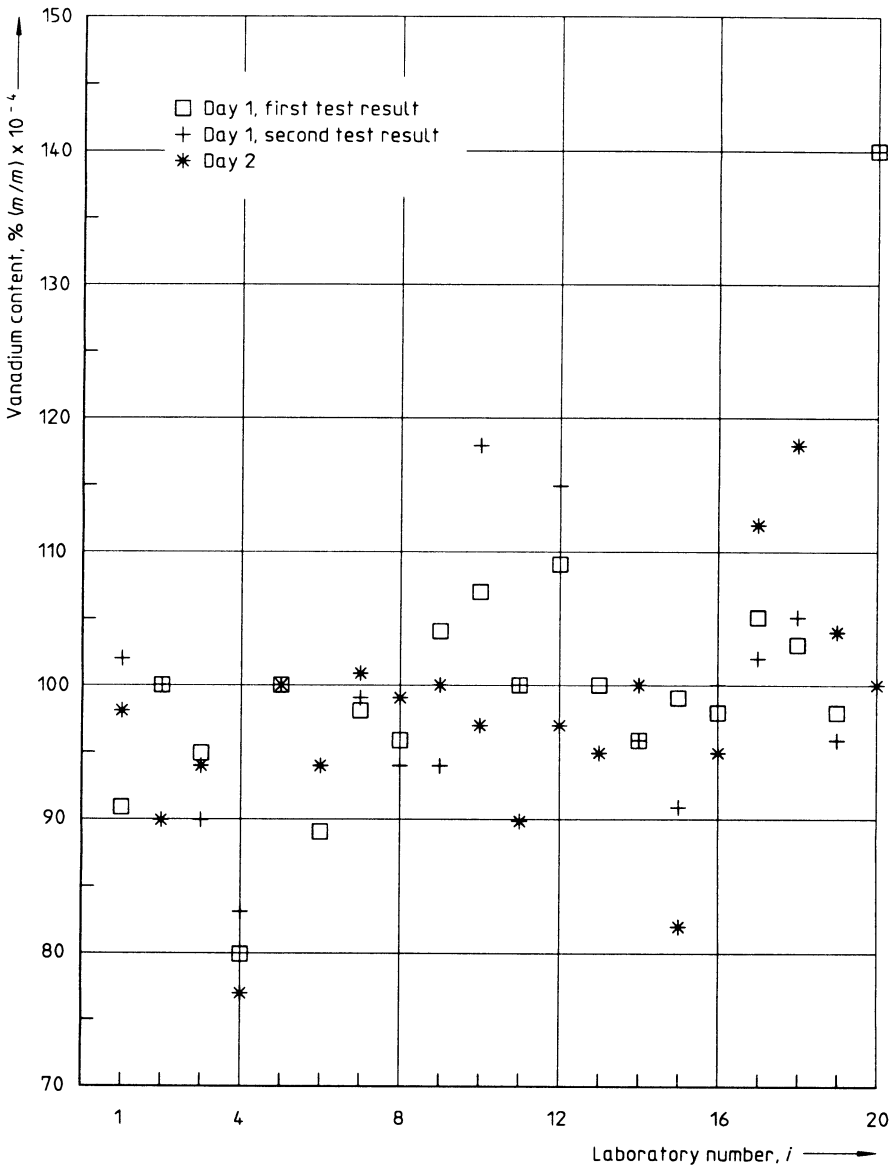

Figure D.2 — Vanadium content in steel — Test results for day 1 and 2 at level 1 versus laboratory number

Table D.2 — Original data — Vanadium content, % (*m*/*m*)

Laboratory No.	**Level 1** (0,01 %)			**Level 2** (0,04 %)			**Level 3** (0,1 %)			**Level 4** (0,2 %)			**Level 5** (0,5 %)			**Level 6** (0,75 %)		
	Day 1		**Day 2**	**Day 1**		**Day 2**	**Day 1**		**Day 2**	**Day 1**		**Day 2**	**Day 1**		**Day 2**	**Day 1**		**Day 2**
i	y_{i1}	y_{i2}	y_{i3}	y_{i1}	y_{i2}	y_{i3}	y_{i1}	y_{i2}	y_{i3}	y_{i1}	y_{i2}	y_{i3}	y_{i1}	y_{i2}	y_{i3}	y_{i1}	y_{i2}	y_{i3}
1	0,009 1	0,010 2	0,009 8	0,038 2	0,038 8	0,038 5	0,101	0,103	0,102	0,214	0,211	0,210	0,514	0,510	0,513	0,755	0,753	0,751
2	0,010 0	0,010 0	0,009 0	0,041 0	0,041 0	0,039 0	0,111	0,111	0,108	0,220	0,220	0,215	0,520	0,540	0,540	0,800	0,755	0,750
3	0,009 5	0,009 0	0,009 4	0,039 0	0,038 0	0,037 0	0,108	0,110	0,107	0,213	0,215	0,215	0,500	0,514	0,504	0,738	0,730	0,724
4	0,008 0	0,008 3	0,007 7	0,037 4	0,036 1	0,038 2	0,109	0,106	0,104	0,214	0,222	0,201	0,519	0,518	0,518	0,744	0,742	0,732
5	0,010 0	0,010 0	0,010 0	0,035 0	0,037 0	0,037 0	0,103	0,103	0,110	0,210	0,210	0,205	0,495	0,500	0,512	0,743	0,753	0,750
6	0,008 9	0,009 4	0,009 4	0,036 8	0,036 8	0,037 7	0,106	0,106	0,108	0,232	0,240	0,221	0,526	0,532	0,513	0,733	0,740	0,746
7	0,009 8	0,009 9	0,010 1	0,037 6	0,038 0	0,038 4	0,107	0,105	0,108	0,215	0,215	0,216	0,521	0,519	0,526	0,754	0,756	0,756
8	0,009 6	0,009 4	0,009 9	0,037 9	0,036 6	0,037 9	0,108	0,107	0,108	0,193	0,195	0,210	0,507	0,493	0,511	0,732	0,729	0,732
9	0,010 4	0,009 4	0,010 0	0,036 5	0,037 0	0,036 7	0,104	0,106	0,105	0,211	0,205	0,213	0,509	0,515	0,515	0,734	0,738	0,747
10	0,010 7	0,011 8	0,009 7	0,037 0	0,037 5	0,038 0	0,105	0,110	0,105	0,210	0,220	0,225	0,520	0,520	0,525	0,760	0,760	0,765
11	0,010 0	0,010 0	0,009 0	0,038 0	0,038 0	0,037 5	0,102	0,102	0,102	0,213	0,211	0,214	0,513	0,516	0,514	0,746	0,748	0,746
12	0,010 9	0,011 5	0,009 7	0,039 0	0,039 0	0,039 0	0,101	0,108	0,105	0,208	0,215	0,210	0,509	0,528	0,510	0,758	0,748	0,750
13	0,010 0	0,009 5	0,009 5	0,037 5	0,037 5	0,037 5	0,103	0,104	0,108	0,212	0,222	0,215	0,510	0,520	0,505	0,735	0,755	0,750
14	0,009 6	0,009 6	0,010 0	0,037 4	0,037 4	0,038 9	0,104	0,106	0,110	0,218	0,218	0,212	0,520	0,528	0,522	0,740	0,735	0,742
15	0,009 9	0,009 1	0,008 2	0,038 1	0,037 5	0,039 2	0,109	0,106	0,107	0,214	0,210	0,211	0,510	0,510	0,515	0,749	0,729	0,744
16	0,009 8	0,010 0	0,009 5	0,037 3	0,037 7	0,039 7	0,105	0,105	0,104	0,215	0,212	0,218	0,519	0,517	0,531	0,754	0,751	0,759
17	0,010 5	0,010 2	0,011 2	0,038 9	0,038 2	0,037 3	0,107	0,108	0,104	0,214	0,210	0,209	0,517	0,515	0,514	0,735	0,728	0,741
18	0,010 3	0,010 5	0,011 8	0,038 2	0,038 0	0,037 4	0,103	0,104	0,103	0,224	0,218	0,217	0,515	0,514	0,517	0,788	0,798	0,787
19	0,009 8	0,009 6	0,010 4	0,038 3	0,037 5	0,036 6	0,110	0,109	0,104	0,217	0,215	0,215	0,530	0,525	0,520	0,755	0,745	0,740
20	0,014 0	0,014 0	0,010 0	0,037 0	0,040 8	0,036 9	0,104	0,106	0,107	0,214	0,214	0,203	0,518	0,518	0,481	0,730	0,737	0,658

Table D.3 — Values of $w_{i(1)}$, $w_{i(2)}$ and $\bar{y}_{i(2)}$

Laboratory No. i	$w_{i(1)}$	$w_{i(2)}$	$\bar{y}_{i(2)}$
1	0,001 1	0,000 15	0,009 700
2	0,000 0	0,001 00	0,009 667
3	0,000 5	0,000 15	0,009 300
4	0,000 3	0,000 45	0,008 000
5	0,000 0	0,000 00	0,010 000
6	0,000 5	0,000 25	0,009 233
7	0,000 1	0,000 25	0,009 933
8	0,000 2	0,000 40	0,009 633
9	0,001 0	0,000 10	0,009 933
10	0,001 1	0,001 55	0,010 733
11	0,000 0	0,001 00	0,009 667
12	0,000 6	0,001 50	0,010 700
13	0,000 5	0,000 25	0,009 667
14	0,000 0	0,000 40	0,009 733
15	0,000 8	0,001 30	0,009 067
16	0,000 2	0,000 40	0,009 767
17	0,000 3	0,000 85	0,010 633
18	0,000 2	0,001 40	0,010 867
19	0,000 2	0,000 70	0,009 933

The sums of squares of $w_{i(1)}$, $w_{i(2)}$ and $\bar{y}_{i(2)}$ and the mean value $\bar{\bar{y}}$ are computed as

$$\sum_i w_{i(1)}^2 = 5{,}52 \times 10^{-6}$$

$$\sum_i w_{i(2)}^2 = 12{,}44 \times 10^{-6}$$

$$\sum_i (\bar{y}_{i(2)})^2 = 1\ 832{,}16 \times 10^{-6}$$

$$\bar{\bar{y}} = \frac{1}{19} \sum_i \bar{y}_{i(2)} = 0{,}009\ 798\ 25$$

From these values, the sums of squares SS0, SS1 and SSe are obtained, and the ANOVA table is prepared as shown in Table D.4.

The unbiased estimates of variances between laboratories $s_{(0)}^2$, between days within a laboratory $s_{(1)}^2$, and the estimated repeatability variance s_r^2 are obtained as

$$s_{(0)}^2 = 0,278 \times 10^{-6}$$

$$s_{(1)}^2 = 0,218 \times 10^{-6}$$

$$s_r^2 = 0,145 \times 10^{-6}$$

The reproducibility standard deviation s_R, time-different intermediate precision standard deviation $s_{I(T)}$, and repeatability standard deviation s_r are obtained as

$$s_R = \sqrt{s_r^2 + s_{(1)}^2 + s_{(0)}^2} = 0{,}801 \times 10^{-3}$$

$$s_{I(T)} = \sqrt{s_r^2 + s_{(1)}^2} = 0{,}603 \times 10^{-3}$$

$$s_r = \sqrt{s_r^2} = 0{,}381 \times 10^{-3}$$

The values of these standard deviations for the six levels of vanadium content are summarized in Table D.5 and shown in Figure D.3.

Table D.4 — ANOVA table — Vanadium content

Source	Sum of squares	Degrees of freedom	Mean square	Expected mean square
0 (laboratory)	$24{,}16 \times 10^{-6}$	18	$1{,}342 \times 10^{-6}$	$\sigma_r^2 + \frac{5}{3}\sigma_{(1)}^2 + 3\sigma_{(0)}^2$
1 (day)	$8{,}29 \times 10^{-6}$	19	$0{,}436 \times 10^{-6}$	$\sigma_r^2 + \frac{4}{3}\sigma_{(1)}^2$
Residual	$2{,}76 \times 10^{-6}$	19	$0{,}145 \times 10^{-6}$	σ_r^2
Total	$35{,}21 \times 10^{-6}$	56		

Table D.5 — Values of s_r, $s_{I(T)}$ and s_R for six levels of vanadium content in steel

Level	Outlier laboratories No.	Average (%)	s_r (%)	$s_{I(T)}$ (%)	s_R (%)
1	20	0,009 8	$0{,}381 \times 10^{-3}$	$0{,}603 \times 10^{-3}$	$0{,}801 \times 10^{-3}$
2	2	0,037 8	$0{,}820 \times 10^{-3}$	$0{,}902 \times 10^{-3}$	$0{,}954 \times 10^{-3}$
3	—	0,105 9	$1{,}739 \times 10^{-3}$	$2{,}305 \times 10^{-3}$	$2{,}650 \times 10^{-3}$
4	6 and 8	0,213 8	$3{,}524 \times 10^{-3}$	$4{,}710 \times 10^{-3}$	$4{,}826 \times 10^{-3}$
5	20	0,516 4	$6{,}237 \times 10^{-3}$	$6{,}436 \times 10^{-3}$	$9{,}412 \times 10^{-3}$
6	20	0,748 4	$9{,}545 \times 10^{-3}$	$9{,}545 \times 10^{-3}$	$15{,}962 \times 10^{-3}$

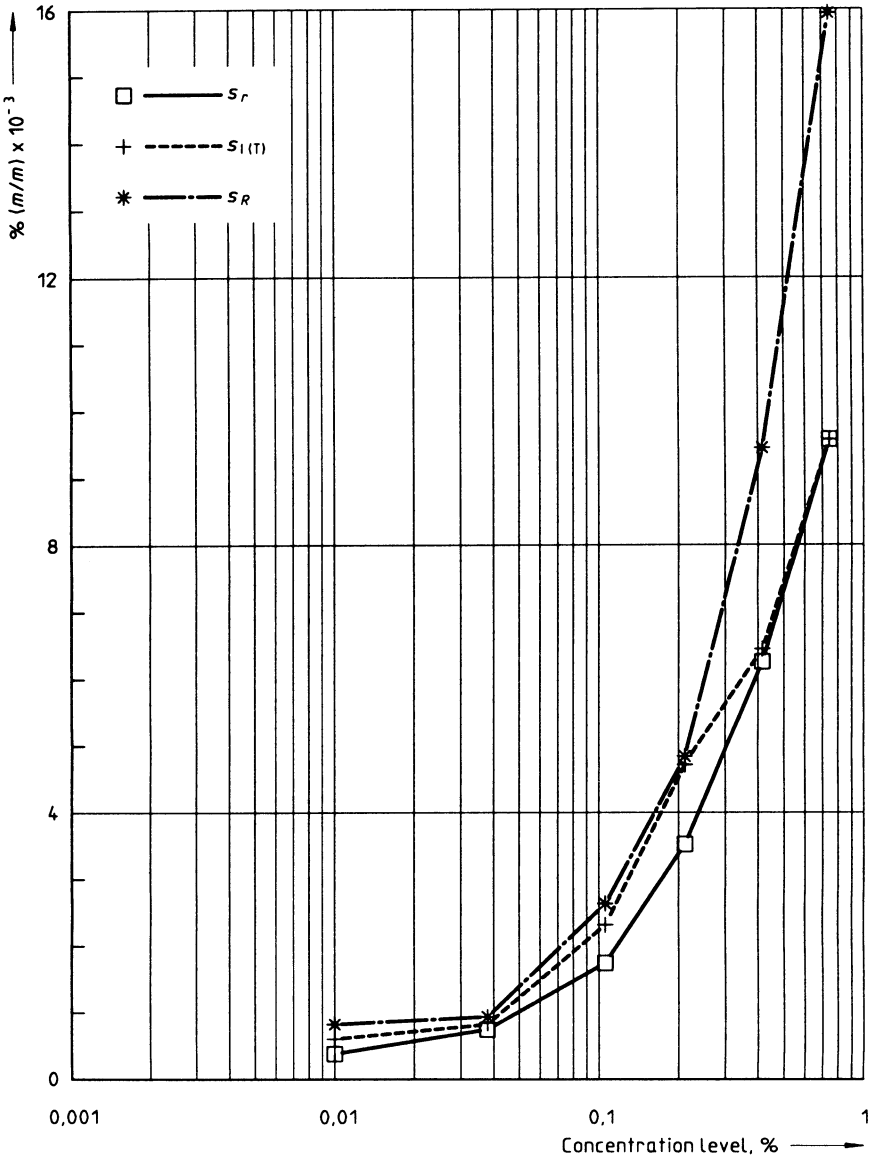

Figure D.3 — Vanadium content in steel — Repeatability standard deviation s_r, time-different intermediate precision standard deviation $s_{I(T)}$ and reproducibility standard deviation s_R as functions of the concentration level

Annex E (informative)
Bibliography

[1] ISO 3534-2:1993, *Statistics — Vocabulary and symbols — Part 2: Statistical quality control.*

[2] ISO 3534-3:1985, *Statistics — Vocabulary and symbols — Part 3: Design of experiments.*

[3] ISO 5725-4:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 4: Basic methods for the determination of the trueness of a standard measurement method.*

[4] ISO 5725-5:—, *Accuracy (trueness and precision) of measurement methods and results — Part 5: Alternative methods for the determination of the precision of a standard measurement method*[1)]

[5] ISO 5725-6:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 6: Use in practice of accuracy values.*

[6] Winer, B.J. *Statistical principles in experimental design.* McGraw-Hill, 1962.

[7] Snedecor, G.W. and Cochran, W.G. *Statistical methods.* Iowa University Press, 1967.

[1)] To be published.

BSI — British Standards Institution

BSI is the independent national body responsible for preparing British Standards. It presents the UK view on standards in Europe and at the international level. It is incorporated by Royal Charter.

Revisions

British Standards are updated by amendment or revision. Users of British Standards should make sure that they possess the latest amendments or editions.

It is the constant aim of BSI to improve the quality of our products and services. We would be grateful if anyone finding an inaccuracy or ambiguity while using this British Standard would inform the Secretary of the technical committee responsible, the identity of which can be found on the inside front cover. Tel: +44 (0)20 8996 9000. Fax: +44 (0)20 8996 7400.

BSI offers members an individual updating service called PLUS which ensures that subscribers automatically receive the latest editions of standards.

Buying standards

Orders for all BSI, international and foreign standards publications should be addressed to Customer Services. Tel: +44 (0)20 8996 9001. Fax: +44 (0)20 8996 7001. Email: orders@bsi-global.com. Standards are also available from the BSI website at http://www.bsi-global.com.

In response to orders for international standards, it is BSI policy to supply the BSI implementation of those that have been published as British Standards, unless otherwise requested.

Information on standards

BSI provides a wide range of information on national, European and international standards through its Library and its Technical Help to Exporters Service. Various BSI electronic information services are also available which give details on all its products and services. Contact the Information Centre. Tel: +44 (0)20 8996 7111. Fax: +44 (0)20 8996 7048. Email: info@bsi-global.com.

Subscribing members of BSI are kept up to date with standards developments and receive substantial discounts on the purchase price of standards. For details of these and other benefits contact Membership Administration. Tel: +44 (0)20 8996 7002. Fax: +44 (0)20 8996 7001. Email: membership@bsi-global.com.

Information regarding online access to British Standards via British Standards Online can be found at http://www.bsi-global.com/bsonline.

Further information about BSI is available on the BSI website at http://www.bsi-global.com.

Copyright

BSI
389 Chiswick High Road
London
W4 4AL