TECHNICAL SPECIFICATION

ISO/TS 23471

First edition
2022-11

Experimental designs for evaluation of uncertainty — Use of factorial designs for determining uncertainty functions

Plans d'expériences pour l'évaluation de l'incertitude — Utilisation de plans factoriels pour la détermination des fonctions d'incertitude

Reference number
ISO/TS 23471:2022(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.

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

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

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

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

This document was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

This document has been elaborated in response to the need for standardized single laboratory designs to determine measurement uncertainty (JCGM 100[1]) by means of experiments. It applies in situations where the standard deviation of the observations is not constant but depends on the measurand and where the measurement uncertainty is derived by a top-down approach. This need has been expressed in such areas as consumer protection, food safety, environmental analytics and medical diagnostics.

Uncertainty evaluation usually requires the quantification and subsequent combination of uncertainties arising from random variation and uncertainties associated with corrections. Random variation may arise within a particular experiment under the same conditions, or across a range of conditions. The former kind of variation occurs under repeatability conditions, hence usually being characterised as repeatability standard deviation or repeatability coefficient of variation; precision across a range of conditions is generally termed intermediate precision or reproducibility (ISO 5725 (all parts)[3]).

The most common experimental design for estimating the laboratory variance and the repeatability variance is the ANOVA design described in ISO 5725-2. In this design, an equal number of observations are collected under repeatability conditions for each participating laboratory. Alternative designs for interlaboratory studies, in which other factors are varied in addition to the laboratory factor, are described in ISO 5725-3. Evaluation of uncertainties based on such a study design is discussed in ISO 21748[6]. Similarly, where the observations are not grouped in different laboratories but in groups of different measurement conditions (e.g. different weeks or technicians) within the same laboratory, the between-group variance component can be considered to represent the uncertainty contribution arising from random variation in the measurement condition which the grouping factor represents. For example, if test results are obtained under repeatability conditions once a week, analysis of variance can provide an estimate of the effect of variation between weeks.

While nested designs are among the most common designs for estimating random variation, they are not the only useful class of design. Consider, for example, an experiment conducted by using three instruments, three batches of reagents and three batches of a solid phase extraction (SPE) cartridges, where every possible combination is included in the design for a total 3 × 3 × 3 = 27 runs. As every possible combination has the same number of observations, this design is called balanced, and as factors are not nested within each other, the factors instrument, reagent and SPE cartridge are said to be ‘crossed’. This type of experiment is considered in ISO/TS 17503[5] for the uncertainty evaluation of the mean in two-factor crossed designs. Just as in the case of the nested design, the aim is to extract the variance components corresponding to the three factors. Suitable models are available and are referred to in the statistical literature as random-effects or (if one factor is a fixed effect) mixed-effects models. This approach can be extended to take more than three factors into account. However, if all factor level combinations are included in the design, the corresponding number of measurements can become very high. For example, for five factors, each with three levels, there are already 3^5 = 243 factor level combinations. If it is necessary to include five or more factors in the experiment, the number of levels should be as low as possible (two levels), and it is recommended to implement an orthogonal design, whereby only a selection of factor level combinations is included.

It is assumed in this document that the measured values are non-negative numbers and that all variance components consist of two parts: one part which is proportional to the level of the measurand and another which is constant across levels. Estimation of variance components can be achieved by several methods. For balanced designs, computing expected mean squares from classical analysis of variance is straightforward. Restricted (sometimes also called residual) maximum likelihood estimation (REML) is widely recommended for estimation of variance components and is applicable to both balanced and unbalanced designs.

Experimental designs for evaluation of uncertainty — Use of factorial designs for determining uncertainty functions

1 Scope

This document specifies experimental procedures and statistical analysis for the determination of measurement uncertainty in situations where the following conditions are fulfilled:

Condition 1: The level of the measurand is non-negative, e.g. concentration level of a contaminant in a sample.

Condition 2: Measurement error consists of two independent components: for one of these components the relative standard deviation is constant (that is, the absolute deviation is proportional to the level of the measurand), whereas for the other component the absolute standard deviation is constant (that is, independent of the level of the measurand).

Condition 3: Samples for different levels of the measurand can be made available; if the level of the measurand is the concentration of a chemical substance, samples could be obtained e.g. by fortifying (spiking) blank samples.

Conditions 1 and 2 are met for most applications of instrumental chemical analyses. Condition 3 can be met for chemical analyses if blank samples are available.

This document can also be used to determine precision data for a particular laboratory for different technicians, different environmental conditions, the same or similar test items, with the same level of the measurand, over a certain period of time.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

ISO 3534-2, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

ISO 3534-3, *Statistics — Vocabulary and symbols — Part 3: Design of experiments*

ISO 3534-4, *Statistics — Vocabulary and symbols — Part 4: Survey sampling*

ISO/IEC Guide 98-3, *Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 3534-1, ISO 3534-2, ISO 3534-3, ISO 3534-4, ISO/IEC Guide 98-3 and the following apply. ISO and IEC maintain terminological databases for use in standardization at the following addresses:

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

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

3.1
block
group of settings which are conducted in parallel or in a short time interval, and which are used for the same samples

EXAMPLE Two settings:

Technician 1 + culture medium 2 + temperature 1 + incubator 1

AND

Technician 2 + culture medium 1 + temperature 2 + incubator 2

3.2
factor
qualitative or quantitative parameter that is varied with the intent of assessing its effect on the response variable

Note 1 to entry: A factor may provide an assignable cause for the outcome of an experiment.

Note 2 to entry: The use of factor here is more specific than its generic use as a synonym for predictor variable.

Note 3 to entry: A factor may be associated with the creation of blocks.

3.3
factor level
value or assignment of a factor

EXAMPLE Ordinal-scale levels of a catalyst may be presence and absence. Four levels of a heat treatment may be 100 °C, 120 °C, 140 °C and 160 °C. The nominal-scale variable for a laboratory can have levels A, B and C, corresponding to three facilities.

Note 1 to entry: A synonym is the value of a predictor variable.

Note 2 to entry: The term “level” is normally associated with a quantitative characteristic. However, it also serves as the term describing the version or setting of qualitative characteristics.

Note 3 to entry: Responses observed at the various levels of a factor provide information for determining the effect of the factor within the range of levels of the experiment. Extrapolation beyond the range of these levels is usually inappropriate without a firm basis for assuming model relationships. Interpolation within the range may depend on the number of levels and the spacing of these levels. It is usually reasonable to interpolate, although it is possible to have discontinuous or multi-modal relationships that cause abrupt changes within the range of the experiment. The levels may be limited to certain selected fixed values (whether these values are or are not known) or they may represent a purely random selection over the range to be studied.

3.4
in-house repeatability
measurement precision under a set of in-house repeatability conditions of measurement in a particular laboratory

Note 1 to entry: In-house repeatability conditions include the same measurement procedure, same technicians, same measuring system, same operating conditions and same location, and replicate measurements on the same or similar objects over a short period of time in a particular laboratory.

3.5
in-house reproducibility
measurement precision under in-house reproducibility conditions of measurement in a particular laboratory

Note 1 to entry: In-house reproducibility conditions include different technicians, operating conditions, and replicate measurements on the same or similar test items over a certain period of time in a particular laboratory.

3.6
sample material
material from which the samples are made

Note 1 to entry: The choice of sample material can have an effect on the bias of the measurement procedure.

3.7
factor level combination
combination of factor levels

EXAMPLE Technician 1 + culture medium 2 + temperature 1 + etc.

Note 1 to entry: These conditions can be described by the combination of factor levels corresponding to those factors varied within the study.

3.8
test portion
quantity of material, of proper size for measurement of the concentration or other property of interest, removed from the test sample, used for a particular measurement

[SOURCE: ISO 11074:2015[4], 4.3.15, modified — “used for a particular measurement" has been added, and Note 1 to entry and Note 2 to entry have been deleted.]

4 Symbols

A_j	Absolute component of the linear block effect j
$A(j)$	Absolute component of the effect of factor level combination j
a_{ijk}	Absolute component of the repeatability error for measurement result Y_{ijk}
B_j	Relative component of the linear block effect j
$B(j)$	Relative component of the effect of factor level combination j
b_{ijk}	Relative component of the repeatability error for measurement result Y_{ijk}
$df(x)$	Effective degrees of freedom at level x
i	Identifier for a particular level
	Identifier for a particular measurement (see Annex F)
j	Identifier for a particular sample or block or factor level combination
	Identifier for a particular measurement (see Annex F)
$L(\boldsymbol{\theta})$	Log-likelihood for variance components $\boldsymbol{\theta}$
k	Identifier for a particular measurement result at level $i=1,\ldots,m$ in block $j=1,\ldots,n$
	Coverage factor for measurement uncertainty
$\boldsymbol{M}$	Transformation matrix used in REML estimation
m	Number of levels of the measurand
n	Number of samples or blocks or factor level combinations at one level of the measurand
p	Number of measurement results obtained at one level and one sample, block or factor level combination

q	Number of factors
REML	Estimate arising from a restricted maximum likelihood calculation
$s_f(j)$	Factor level of factor $f=1,\ldots,q$ at factor level combination j
$u\left(\hat{x}_{ij}\right)$	Standard measurement uncertainty of the estimated value of the level x
$u(Y)$	Standard measurement uncertainty of the measured value Y for the true level x
$U(Y)$	Expanded measurement uncertainty of the measured value Y for the true level x
$\boldsymbol{V}(\boldsymbol{\theta})$	Covariance matrix used in REML estimation
$\boldsymbol{X}$	Design matrix for REML estimation
x	True value of the level of the measurand
$\hat{x}$	Estimated value of the level of the measurand
x_{ij}	True value of the level of the sample at level $i=1,\ldots,m$ in block $j=1,\ldots,n$.
x_L	Lower limit of measurement uncertainty interval
x_U	Upper limit of measurement uncertainty interval
Y_{ijk}	Measurement result of the test portion (i,j,k) at level $i=1,\ldots,m$ in block $j=1,\ldots,n$ for replicate $k=1,\ldots,p$
Y	Measured value
$\boldsymbol{Y}$	Vector of measured values for REML estimation
Y_corr	Corrected measured value (recovery correction)
α	True value of absolute component of method bias
$\hat{\alpha}$	Estimated value of absolute component of method bias
β	True value of relative component of method bias plus 1 (relative recovery)
	Vector of true value of absolute component of method bias and true value of relative component of method bias plus 1 for REML estimation
$\hat{\beta}$	Estimated value of relative component of method bias plus 1 (relative recovery)
$\boldsymbol{\theta}$	Vector of variance components used in REML estimation
$\mu(x)$	Expected measured value at level x, $\mu(x)=\alpha+\beta x$
σ	True value of a standard deviation
σ^2	True value of a variance
$\hat{\sigma}$	Estimated value of a standard deviation
$\hat{\sigma}^2$	Estimated value of a variance
$\sigma_\mathrm{ri}^2(x)$	True value of a variance function under in-house repeatability conditions at level x
$\sigma_\mathrm{Ri}^2(x)$	True value of a variance function under in-house reproducibility conditions at level x

$\sigma_{\hat{\mu}}^{2}(x)$ True value of the standard variance of the linear function of the expected measured value

5 General principles

5.1 General

Top-down procedures for the determination of measurement uncertainty depend on whether a conventional approach (see 5.2) or a factorial approach (see 5.3) is chosen. In any case, all results are valid only for the laboratory within which the study was conducted.

5.2 Principles of conventional approach

The conventional approach assumes that all measurements can be grouped in blocks. For measurements within a block, in-house repeatability conditions are fulfilled. For measurements between blocks, in-house reproducibility conditions are fulfilled. Thus, between blocks, there is variation with respect e.g. to time of measurement, technician, environmental conditions, batches of reagents or measuring instruments.

5.3 Principles of factorial approach

Compared to the conventional approach where measurement conditions vary randomly from one measurement block to the other, in the factorial approach at least some of these measurement conditions are controlled. For instance, half the measurements are conducted with reagents from batch A, and the other half with reagents from batch B. In contrast, the conventional approach where, apart from the grouping factor (e.g. day 1 and day 2), there is no control of measurement conditions. The factorial approach allows the assessment of the measurement uncertainty of the test results obtained under a variety of typical test conditions in a given laboratory, such as different analysts, different instruments, different lots of reagents, different elapsed assay times and different assay temperatures. The factorial approach aims at establishing reliable precision data and measurement data by simultaneous controlled variation of the selected factors. It allows the evaluation of the combined impact of factorial effects.

6 Conventional approach

6.1 General considerations

The conventional approach consists of the following parts:

— selection of levels of the measurand (for example concentration levels) and sample materials;

— design of study and allocation of test portions to different blocks;

— conducting the measurement;

— statistical analyses.

The results (tables and calculations), including discrepant results, if any, shall be given in a study report.

6.2 Selection of samples and levels

The study is typically conducted using m = 4,..., 8 homogeneous sample materials at different levels of the measurand from the same sample type. Samples should differ only with regard to the level of the measurand. The ratio of the maximum level divided by the minimum level should be at least 1,5 but not larger than 50. If the ratio exceeds 4, additional testing of statistical model assumptions (linearity, homoscedasticity) and examination of effective degrees of freedom is required.

6.3 Experimental design

The sample material is available in m different levels of the measurand. For each level, n samples are randomly allocated to n blocks (where $n \geq 8$). Thus, each of the n blocks consists of m samples, each sample with a different level of the measurand. Each sample consists of p test portions, i.e. p represents the number of replicates at each level and each block. There is no limit on the number of replicates p; even without replication ($p = 1$) precision and uncertainty can be computed, although better reliability is obtained with $p \geq 2$ replicates.

Each block contains $m \times p$ test portions, and altogether there are $m \times n \times p$ test portions. The $m \times p$ test portions of each block shall be analysed under in-house repeatability conditions. Different blocks shall be analysed under different conditions, preferably with different equipment, different personnel and in different weeks. Thus, if the blocking (grouping) factor is *week*, a total of n weeks, for instance, are required for measuring $m \times n \times p$ test portions. The blocking factor should be chosen so that temporal autocorrelation can be excluded, i.e. a longer period is generally better than a shorter period.

NOTE 1 Reliability of precision data is highly dependent on the number of blocks (see ISO 5725-1). Relative standard error of uncertainty data can exceed 29 % if there are fewer than $n = 8$ blocks.

6.4 Statistical analysis

6.4.1 Statistical model

Y_{ijk} denotes the measurement result of the test portion (i,j,k) at level $i=1,\ldots,m$ in block $j=1,\ldots,n$ for replicate $k=1,\ldots,p$.

x_{ij} denotes the true value of the level of the corresponding test portion at level $i=1,\ldots,m$ in block $j=1,\ldots,n$.

The statistical model is

$$Y_{ijk} = \alpha + \beta x_{ij} + A_j + B_j x_{ij} + a_{ijk} + b_{ijk} x_{ij} \tag{1}$$

where

α and $\beta - 1$	denote the absolute and the relative component of method bias, respectively;
A_j and B_j	denote the absolute and the relative component of the linear block effect j, respectively;
a_{ijk} and b_{ijk}	denote the absolute and the relative component of the repeatability error for measurement result Y_{ijk}.

NOTE 1 In the model described here, it is assumed that the true value x_{ij} of the measurand is known. In practice, it is an estimated value $\hat{x}_{ij}$ of x_{ij} that is subject to uncertainty. If the uncertainty of this value is small compared to the uncertainty of Y, for computational purposes x_{ij} can be replaced by $\hat{x}_{ij}$. This is the case, for example, when blank material is spiked with a high-purity substance. If spiking is carried out in one step before the material is homogenized and distributed to the individual samples, x_{ij} does not depend on block j. However, if the individual samples are spiked gravimetrically, the concentrations x_{ij} could be different. The same applies if the material cannot be perfectly homogenized and the determination of x_{ij} is carried out by a very accurate reference method.

NOTE 2 The method bias is a linear function of the measurand. It is constant across all blocks. The linear block effect is also a linear function of the measurand. However, it varies from block to block randomly. This random effect results on the one hand from factorial effects, but can also be influenced by sampling, sample preparation and analysis (see References [7] and [8]), because the measurement result can also depend on the composition of the test portion even though the true values of the measurand in the samples may be identical.

It is assumed that α and β are fixed values, whereas

$A_1, \dots, A_n$ are random variables with zero mean and variance σ_A^2;

$B_1, \dots, B_n$ are random variables with zero mean and variance σ_B^2;

$a_{111}, \dots, a_{mnp}$ are random variables with zero mean and variance σ_a^2;

$b_{111}, \dots, b_{mnp}$ are random variables with zero mean and variance σ_b^2.

Finally, it is assumed that all random variables are independent with a distribution closely approximating the normal distribution.

NOTE 1 Temporal autocorrelation between blocks can be examined by calculating the ratio of the mean of squared differences between subsequent block means and the variance of all block means. If this ratio is below 0,66, temporal autocorrelation is considered significant. Information regarding autocorrelation can also be derived from the data of quality control charts.

6.4.2 Calculation of in-house repeatability and in-house reproducibility

In mathematical terms, the model described in 6.4.1, Formula (1) is a mixed linear model. Its unknown parameters σ_A^2, σ_B^2, σ_a^2, σ_b^2 can be estimated by the statistical calculation method REML (Restricted Maximum Likelihood[10][12]).

REML can be obtained by maximizing the likelihood function as described in Annex A. REML is available in many statistical packages. Explicit formulae are not available, so that computer software is necessary.

With these variance components, the variance function under in-house repeatability conditions is

$$\sigma_{\mathrm{ri}}^2(x) = \sigma_a^2 + x^2\sigma_b^2$$

where x denotes the level of the measurand. Similarly, the variance function under in-house reproducibility conditions is

$$\sigma_{\mathrm{Ri}}^2(x) = \sigma_A^2 + x^2\sigma_B^2 + \sigma_a^2 + x^2\sigma_b^2.$$

In addition, estimates for the fixed effects α and β together with their standard errors and an estimate of the standard error $\sigma_{\hat{\mu}}(x)$ for the linear function $\mu(x) = \alpha + \beta x$ are provided by standard statistical software packages.

The relative standard error of the in-house reproducibility standard deviation $\sigma_{\mathrm{Ri}}(x)$ should be less than 0,30. It is very much dependent on the number of blocks n. An approximation and upper limit for this relative standard error is 0,32 for $n = 6$, 0,27 for $n = 8$ and 0,24 for $n = 10$.

If there is a significant bias, i.e. if α and β are significantly different from 0 and 1, respectively, the measured value Y should be replaced by

$$Y_{\text{corr}} = \frac{Y - \hat{\alpha}}{\hat{\beta}}$$

where $\hat{\alpha}$ and $\hat{\beta}$ represents the estimated values of the unknown parameters α and β.

NOTE 1 For the calculation of variance components REML does not require any replicates. However, with replicates, calculation of variance components generally becomes more reliable and can be examined more easily with regard to non-linearity or heteroscedasticity.

NOTE 2 It is possible that under certain conditions there are random non-linearities in the measured values within blocks. These random non-linearities will inflate both the repeatability standard deviation and the in-house reproducibility standard deviation.

NOTE 3 It is also possible that under certain conditions there are systematic non-linearities in the measurement results across blocks. These systematic non-linearities will inflate the in-house reproducibility standard deviation, but not the repeatability standard deviation.

NOTE 4 If the levels x_{ij} of the measurand are not known and are thus replaced by the respective arithmetic means of the measurement values, method bias cannot be determined.

NOTE 5 Within each block, measurement values are correlated. Correlation between measurement values at level x is determined by $1 - \sigma_{\text{ri}}^2(x) / \sigma_{\text{Ri}}^2(x)$.

NOTE 6 The upper limit of the relative standard error of the in-house reproducibility standard deviation is a good approximation for the case when $\sigma_{\text{ri}}^2(x)$ is small compared to $\sigma_{\text{Ri}}^2(x)$. By means of Monte Carlo simulation, a more accurate calculation is possible.

6.4.3 Calculation of measurement uncertainty

6.4.3.1 Conventional measurement uncertainty

Compute the standard measurement uncertainty of the measured value Y for the true level x:

$$u(Y) = \sqrt{\hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2 + \hat{\sigma}_{\hat{\mu}}^2(x)}$$

and the expanded measurement uncertainty

$$U(Y) = k \cdot u(Y) = k\sqrt{\hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2 + \hat{\sigma}_{\hat{\mu}}^2(x)}$$

where $\hat{\sigma}_A^2, \hat{\sigma}_B^2, \hat{\sigma}_a^2, \hat{\sigma}_b^2$ denote the estimates of the unknown variance components $\sigma_A^2, \sigma_B^2, \sigma_a^2, \sigma_b^2$ and $\hat{\sigma}_{\hat{\mu}}^2(x)$ represents the estimate of the standard variance of the linear function of expected measured value, $\alpha + \beta x$.

If it is known that bias is negligible, the standard error $\hat{\sigma}_{\hat{\mu}}(x)$ can be ignored, i.e. $\sigma_{\hat{\mu}}^2(x) = 0$.

If there is significant bias in Y, compute the standard measurement uncertainty of the corrected measured value $Y_{\text{corr}} = \dfrac{Y-\hat{\alpha}}{\hat{\beta}}$:

$$u\left(Y_{\text{corr}}\right) = \sqrt{\frac{\hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2}{\hat{\beta}^2} + \hat{\sigma}_{\hat{\mu},\text{corr}}^2\left(x\right)}$$

and the expanded measurement uncertainty

$$U\left(Y_{\text{corr}}\right) = k \cdot u\left(Y_{\text{corr}}\right) = k\sqrt{\frac{\hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2}{\hat{\beta}^2} + \hat{\sigma}_{\hat{\mu},\text{corr}}^2\left(x\right)}$$

where $\sigma_{\hat{\mu},\text{corr}}\left(x\right)$ denotes the standard error of the corrected estimate $\hat{\mu}_{\text{corr}}\left(x\right)$.

NOTE In the model described in 6.4.1, Formula (1), it is assumed that the true value of the measurand, x_{ij}, is known. In practice, it is an estimated value $\hat{x}_{ij}$ with a standard uncertainty $u\left(\hat{x}_{ij}\right)$. With this additional standard uncertainty, the standard uncertainty of the measured value Y and the corrected measured value Y_{corr} for the true level x can be computed $u\left(Y\right) = \sqrt{\hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2 + \hat{\sigma}_{\hat{\mu}}^2\left(x\right) + u^2\left(\hat{x}_{ij}\right)}$ and $u\left(Y_{\text{corr}}\right) = \sqrt{\dfrac{\hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2}{\hat{\beta}^2} + \hat{\sigma}_{\hat{\mu},\text{corr}}^2\left(x\right) + u^2\left(\hat{x}_{ij}\right)}$, respectively.

7 Factorial approach

7.1 General considerations

The factorial single-laboratory validation approach allows assessing the in-house reproducibility of the test results obtained under a variety of normal test conditions in a given laboratory, such as different analysts, different instruments, different lots of reagents, different elapsed assay times, different assay temperatures, etc. The factorial approach aims at establishing reliable precision data by simultaneous controlled variation of factors that are typically encountered in the routine application of the method and which cannot be fixed. It allows evaluating the combined impact of factorial effects. This approach differs from a robustness study, in which the focus is the detection of specific significant method parameters and the optimization of the performance of the method.

Compared to the conventional approach, a properly designed factorial approach requires a smaller number of samples and a smaller number of tests for the same level of reliability in the computation of validation parameters.

The factorial approach consists of the following parts:

— selection of levels of the measurand (concentration levels) and sample materials;

— selection of factors and factor levels;

— design of study and allocation of test portions to different factor level combinations;

— conducting the measurement;

— statistical analyses.

The results (tables and calculations), including discrepant results, should be given in a study report.

7.2 Selection of sample materials, factors and factor levels

The study is typically conducted at $m=4,\ldots,8$ different levels of the measurand. The ratio of maximum level divided by minimum level should be at least 1,5 but not larger than 50. If the ratio exceeds 4, additional testing of statistical model assumptions (linearity, homoscedasticity) and effective degrees of freedom is required.

Relevant method factors that are difficult to control under routine conditions shall be selected and varied systematically, e.g., sample material from different batches (matrix mismatch effects), conditions during measurement (e.g. temperature and humidity), sample preparation, sample storage (duration and temperature), laboratory personnel and laboratory equipment personnel.

In addition, a block factor is selected, i.e. measurements must be performed e.g. in different blocks and different weeks. Each factor is varied on two or more levels. Typically, $q=2,\ldots,8$ factors are selected.

Factors and factor levels shall reflect the typical variation within the single laboratory under routine conditions.

NOTE 1 If factors have significant and persistent effects, the measurement method can be corrected for this effect. Otherwise the measurement method cannot be considered “optimal” (as it allows further optimization). This TS considers factors with smaller, non-significant effects in the range of the repeatability. The term non-significant effect only means that the effect is not significant in the sense of a statistical test. Factors exhibiting such effects are often held constant in the conduct of experimental studies, but in routine analyses their combined effect can have a significant impact on the measurements and thus on the measurement uncertainty.

7.3 Experimental design

At each level of the measurand, $n \times p$ test portions are allocated to n selected factor level combinations of the q factors. Each factor level combination $j=1,\ldots,n$ can be represented by a vector of factor levels, $\left(s_1(j),\ldots,s_q(j)\right)$, and shall be conducted in p replications ($p \geq 1$), and at each level of the measurand, the same factor level combinations shall be conducted. All $m \times p$ measurements belonging to the same factor level combination shall be conducted under repeatability conditions, i.e. within one series or within one day.

Factor level combinations should be selected in such a way as to ensure factorial effects can be estimated independently from each other, e.g. by using balanced or orthogonal designs. A balanced design consists of all possible combinations of factor levels: e.g. with three factors each with three levels there are 27 factor level combinations. Each of these 27 factor level combinations represents one block.

If a balanced factorial design is not possible, an orthogonal design can be conducted. Annex B describes an orthogonal design with $n=8$ factor level combinations without replications ($p=1$) for $q=7$ factors with 2 levels. Annex C describes an orthogonal design with $n=12$ factor level combinations for 3 factors with 2 factor levels and 1 factor with 3 factor levels. Annex D describes an orthogonal design with $n=9$ factor level combinations for $q=4$ factors with 3 factor levels.

7.4 Statistical analysis

7.4.1 Statistical model

Y_{ijk} denotes the measurement result of the test portion at level $i=1,\ldots,m$ for factor level combination $j=1,\ldots,n$ and replicate $k=1,\ldots,p$. All measurements of factor level combination j are obtained in the same block j, i.e. within one measurement series or within one day.

x_{ij} denotes the true value of the level of the corresponding test portion at level $i=1,\ldots,m$ in factor level combination $j=1,\ldots,n$.

The statistical model is

$$Y_{ijk} = \alpha + \beta x_{ij} + A(j) + B(j)x_{ij} + A_j + B_j x_{ij} + a_{ijk} + b_{ijk} x_{ij} \quad (2)$$

where

α and $\beta - 1$	denote the absolute and the relative component of method bias;
$A(j)$ and $B(j)$	denote the absolute and the relative component of the effect of factor level combination j, respectively;
A_j and B_j	denote the absolute and the relative component of the block effect of block j;
a_{ijk} and b_{ijk}	denote the absolute and the relative component of the repeatability error for measurement result Y_{ijk} .

Furthermore, it is assumed that α and β are fixed (non-random) values.

NOTE 1 In the model described here, it is assumed that the true value of the measurand, x_{ij}, is known. In practice, it is an estimated value $\hat{x}_{ij}$ that is subject to uncertainty. If the uncertainty of this value is small compared to the uncertainty of Y, x_{ij} can be replaced by $\hat{x}_{ij}$.

NOTE 2 The method bias is a linear function that depends linearly on the measurand. It is constant across all blocks. The block effect is also a linear function that depends linearly on the measurand. However, it varies from block to block randomly.

The absolute and the relative component of the effect of factor level combination j can be written

$A(j) = A_1(s_1(j)) + \dots + A_q(s_q(j))$ and

$B(j) = B_1(s_1(j)) + \dots + B_q(s_q(j))$.

$A_1(s), \dots, A_q(s)$ are random variables with zero mean and variances $\sigma^2_{A,1}, \dots, \sigma^2_{A,q}$, respectively, and

$B_1(s), \dots, B_q(s)$ are random variables with zero mean and variances $\sigma^2_{B,1}, \dots, \sigma^2_{B,q}$, respectively.

$A_1, \dots, A_n$ are random variables with zero mean and variance σ^2_A, and

$B_1, \dots, B_n$ are random variables with zero mean and variance σ^2_B.

$a_{111}, \dots, a_{mnp}$ are random variables with zero mean and variance σ^2_a, and

$b_{111}, \dots, b_{mnp}$ are random variables with zero mean and variance σ^2_b,

Finally, it is assumed that all random variables are independent with a distribution closely approximating the normal distribution.

7.4.2 Calculation of in-house repeatability and in-house reproducibility

In mathematical terms, the model described in 7.4.1 is a mixed linear model. Its unknown variance components $\sigma^2_{A,1}, \dots, \sigma^2_{A,q}, \sigma^2_{B,1}, \dots, \sigma^2_{B,q}, \sigma^2_A, \sigma^2_B, \sigma^2_a, \sigma^2_b$ shall be estimated by the statistical calculation method REML (Restricted Maximum Likelihood). REML is available in many statistical packages. Explicit formulae are not available, so that computer software is necessary.

With these variance components, the variance function under in-house repeatability conditions can be computed

$$\sigma_{\mathrm{ri}}^2(x)=\sigma_a^2+x^2\sigma_b^2$$

where x denotes the level of the measurand. Similarly, the variance function under in-house reproducibility conditions can be computed:

$$\sigma_{\mathrm{Ri}}^2(x)=\sigma_{A,1}^2+\ldots+\sigma_{A,q}^2+x^2\left(\sigma_{B,1}^2+\ldots+\sigma_{B,q}^2\right)+\sigma_A^2+x^2\sigma_B^2+\sigma_a^2+x^2\sigma_b^2.$$

In addition, estimates for the fixed effects α and β together with their standard errors and an estimate of the standard error $\sigma_{\hat{\mu}}(x)$ for the linear function $\mu(x)=\alpha+\beta x$ of the expected measured value are provided by statistical software packages.

The relative standard error of the in-house reproducibility standard deviation should be less than 0,3. This criterion is violated when one or more of the factors $1,\ldots,q$ have a very large effect on the measurement values with a variance which is clearly exceeding the in-house repeatability standard deviation. In such a case, the measurement method cannot be considered under statistical control and needs further optimization.

If the relative standard error of the in-house reproducibility standard deviation $\sigma_{\mathrm{Ri}}^2(x)$ is not available, it can be approximated by Monte Carlo simulation[14].

NOTE 1 For the calculation of variance components REML does not require any replicates. However, with replicates, calculations of variance components become more reliable and can be examined more easily with regard to non-linearity or heteroscedasticity.

NOTE 2 It is possible that under certain conditions there are random deviations from the linear relationship between measurand and expected measured value for each factor level combination *j*. These non-linearities will inflate both the repeatability standard deviation and the in-house reproducibility standard deviation.

NOTE 3 It is also possible that under certain conditions there may be systematic non-linearities in the measurement results across factor level combinations. These non-linearities will inflate the in-house reproducibility standard deviation but not the repeatability standard deviation.

NOTE 4 It is possible that there are interaction effects which are not explicitly considered in the model. By means of orthogonal design, these interaction effects are confounded with main effects and hence are taken into consideration implicitly.

NOTE 5 It is possible that there are interaction effects which are not explicitly considered in the model. By means of orthogonal design, these interaction effects are confounded with main effects and hence are taken into consideration implicitly.

NOTE 6 Some of the factors mentioned in 7.4.1 can have a systematic effect, e.g. storage duration. If routine samples are analysed, the storage time of the samples is usually a random quantity with a mean value m and a standard deviation s. If the effect of the storage time on the measured values is linear, its contribution to the total variance could be determined from the said standard deviation, if a sufficient number of measured values are available. Computationally, an orthogonal factorial approach and a REML estimate give (almost) the same result if the experimental design is orthogonal and if the two factor levels for the storage time in the experimental design are $m-s/\sqrt{2}$ and $m+s/\sqrt{2}$. The advantage of the method is that significantly fewer measurements are required.

NOTE 7 If the levels x_{ij} of the measurand are not known and thus replaced by the respective arithmetic mean of the measured values, method bias cannot be determined.

7.4.3 Calculation of measurement uncertainty

Compute the standard measurement uncertainty of the measured value Y for the true level x:

$$u(Y) = \sqrt{\hat{\sigma}_{A,1}^2 + \ldots + \hat{\sigma}_{A,q}^2 + x^2\left(\hat{\sigma}_{B,1}^2 + \ldots + \hat{\sigma}_{B,q}^2\right) + \hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2 + \hat{\sigma}_{\hat{\mu}}^2(x)}$$

and the expanded measurement uncertainty

$$U(Y) = k \cdot u(Y) = k\sqrt{\hat{\sigma}_{A,1}^2 + \ldots + \hat{\sigma}_{A,q}^2 + x^2\left(\hat{\sigma}_{B,1}^2 + \ldots + \hat{\sigma}_{B,q}^2\right) + \hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2 + \hat{\sigma}_{\hat{\mu}}^2(x)}$$

where $\sigma_{\hat{\mu}}^2(x)$ represents the standard error of the bias correction function.

If it is known that bias is negligible, the standard error $\sigma_{\hat{\mu}}(x)$ can be ignored, i.e. $\sigma_{\hat{\mu}}^2(x) = 0$.

If there is significant bias in Y, compute the standard measurement uncertainty of the corrected measured value $Y_{\text{corr}} = \dfrac{Y - \hat{\alpha}}{\hat{\beta}}$:

$$u(Y_{\text{corr}}) = \frac{1}{\hat{\beta}}\sqrt{\hat{\sigma}_{A,1}^2 + \ldots + \hat{\sigma}_{A,q}^2 + x^2\left(\hat{\sigma}_{B,1}^2 + \ldots + \hat{\sigma}_{B,q}^2\right) + \hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2 + \hat{\sigma}_{\hat{\mu},\text{corr}}^2(x)},$$

and the expanded measurement uncertainty

$$U(Y_{\text{corr}}) = k \cdot u(Y_{\text{corr}}) = \frac{k}{\hat{\beta}}\sqrt{\hat{\sigma}_{A,1}^2 + \ldots + \hat{\sigma}_{A,q}^2 + x^2\left(\hat{\sigma}_{B,1}^2 + \ldots + \hat{\sigma}_{B,q}^2\right) + \hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2 + \hat{\sigma}_{\hat{\mu},\text{corr}}^2(x)}$$

where $\sigma_{\hat{\mu},\text{corr}}(x)$ denotes the standard error of the corrected estimate $\hat{\mu}_{\text{corr}}(x)$.

An example is provided in Annex F. A method for the calculation of an asymmetric measurement uncertainty interval is described in Annex E.

NOTE In the model described here, it is assumed that the true value x_{ij} of the measurand is known. In practice, it is an estimated value $\hat{x}_{ij}$ with a standard uncertainty $u(\hat{x}_{ij})$. With this additional standard uncertainty, the standard uncertainty of the measured value Y and the corrected measured value Y_{corr} for the true level x can be computed:

$$u(Y) = \sqrt{\hat{\sigma}_{A,1}^2 + \ldots + \hat{\sigma}_{A,q}^2 + x^2\left(\hat{\sigma}_{B,1}^2 + \ldots + \hat{\sigma}_{B,q}^2\right) + \hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2 + \hat{\sigma}_{\hat{\mu}}^2(x) + u^2(\hat{x}_{ij})} \text{ and}$$

$$u(Y_{\text{corr}}) = \sqrt{\frac{\hat{\sigma}_{A,1}^2 + \ldots + \hat{\sigma}_{A,q}^2 + x^2\left(\hat{\sigma}_{B,1}^2 + \ldots + \hat{\sigma}_{B,q}^2\right) + \hat{\sigma}_A^2 + x^2\hat{\sigma}_B^2 + \hat{\sigma}_a^2 + x^2\hat{\sigma}_b^2}{\hat{\beta}^2} + \sigma_{\hat{\mu},\text{corr}}^2(x) + u^2(\hat{x}_{ij})},$$

respectively.

Annex A
(informative)

REML estimation

Given an observed $n \times 1$ response vector Y and $n \times 2$ predictor matrix, $\boldsymbol{X}$, the variance components model postulates that $\boldsymbol{Y} \sim N[\boldsymbol{X}\beta, \boldsymbol{V}(\boldsymbol{\theta})]$, where $\boldsymbol{V}(\boldsymbol{\theta}) = \sum_{i=1}^{m} \boldsymbol{\theta}_i \boldsymbol{V}_i$, and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)^T$ is the vector of variance components $\sigma_A^2, \sigma_B^2, \sigma_a^2, \sigma_b^2$ in case of the conventional approach or $\sigma_{A,1}^2, \ldots, \sigma_{A,q}^2, \sigma_{B,1}^2, \ldots, \sigma_{B,q}^2, \sigma_A^2, \sigma_B^2, \sigma_a^2, \sigma_b^2$ in case of the factorial approach. $\boldsymbol{V}_1, \ldots, \boldsymbol{V}_m$ are m fixed symmetric matrices that are determined by the design of the study.

The restricted (or residual, or reduced) maximum likelihood (REML) approach is a particular form of maximum likelihood estimation which does not base estimates on a maximum likelihood fit of all the information, but instead uses a likelihood function calculated from the transformed data that do not include the fixed-effects parameters. A transformation matrix $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}\left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T$ is created based on the design matrix, $\boldsymbol{X}$, which consists of 2 columns, the constant 1 and the concentration level x of the respective observation. 2 redundant rows of $\boldsymbol{M}$ (that leave the matrix rank unchanged) are removed. After pre-multiplying $\boldsymbol{Y}$ by $\boldsymbol{M}$, the parameters of fixed effects α and β are removed from the model. Thus, only the parameters of random effects are estimated.

REML estimates the variance components $\boldsymbol{\theta}$ by maximizing the log-likelihood of the redefined measurement vector $\boldsymbol{MY}$, which is assumed to be normally distributed with mean 0 and covariance matrix $\boldsymbol{MV}(\,\cdot\,)\boldsymbol{M}^T$ where $\boldsymbol{V}(\boldsymbol{\theta})$ denotes the covariance matrix of the measurement vector $\boldsymbol{Y}$. The log-likelihood of $\boldsymbol{MY}$ can be written

$$L(\boldsymbol{\theta}) = -\frac{1}{2} \ln \det\left(\boldsymbol{MV}(\boldsymbol{\theta})\boldsymbol{M}^T\right) - \frac{1}{2} \boldsymbol{Y}^T \boldsymbol{M}^T \left(\boldsymbol{MV}(\boldsymbol{\theta})\boldsymbol{M}^T\right)^{-1} \boldsymbol{MY}$$

REML estimation is available in a number of general-purpose statistical software packages.

Annex B
(informative)

Orthogonal design with 8 factor level combinations for 7 factors with 2 factor levels

B.1 Sample materials and factors

The study is conducted at $m = 4$ different levels of the measurand. At each level, 8 test portions are selected from the sample material. Test portions are denoted 1-8 for level 1, 9-16 for level 2, 17-24 for level 3 and 25-32 for level 4.

7 factors each with 2 levels are selected that represent normal use of the method (e.g. 2 technicians, 2 storage temperatures for test portions, 2 measurement instruments, 2 batches of standard reagents, etc.). Factors and factor levels are expected to reflect the normal variation within the single laboratory (see 7.2).

Each factor level combination is tested in a separate block, i.e. in different days or weeks. The design can be used with ($p > 1$) and without ($p = 1$) replicates. Several examples are presented in References [9] and [11].

B.2 Experimental design

Table B.1 represents the 8 factor level combinations which are conducted separately in 8 blocks. In each block 4 test portions at 4 different levels of the measurand are measured under repeatability conditions.

Table B.1 — Experimental design for 7 factors each at 2 factor levels

Factor level combination j	Test portions	Factors							
		1	2	3	4	5	6	7	Block (week)
1	1,9,17,25	1	1	1	2	2	2	1	1
2	2,10,18,26	1	1	2	2	1	1	2	2
3	3,11,19,27	1	2	1	1	2	1	2	3
4	4,12,20,28	1	2	2	1	1	2	1	4
5	5,13,21,29	2	1	1	1	1	2	2	5
6	6,14,22,30	2	1	2	1	2	1	1	6
7	7,15,23,31	2	2	1	2	1	1	1	7
8	8,16,24,32	2	2	2	2	2	2	2	8

B.3 Statistical analysis

If the design is used without ($p = 1$) replicates, the factor “Block” is confounded with factors 1-7 so that it is not possible to determine the variance components of the block effect separately. Its effect will inflate the effects of factors 1-7, so that the computed variance of $A(j)+B(j)x_{ij}$ already includes the block effect (day or week effect). In this case, the analysis is based on the statistical model

$$Y_{ijk} = \alpha + \beta x_{ij} + A(j) + B(j)x_{ij} + a_{ijk} + b_{ijk}x_{ij},$$

where Y_{ijk} denotes the measurement result of the test portion at level $i=1,\ldots,4$ for factor level combination $j=1,\ldots,9$ and $k=p=1$.

If the design is used with (p > 1) replicates, the block effect can be determined separately from the other factors. In this case, the analysis is based on the statistical model shown in 7.4.1, Formula (2):

$$Y_{ijk} = \alpha + \beta x_{ij} + A(j) + B(j)x_{ij} + A_j + B_j x_{ij} + a_{ijk} + b_{ijk}x_{ij}$$

where Y_{ijk} denotes the measurement result of the test portion at level $i=1,\ldots,4$ for factor level combination $j=1,\ldots,8$.

Annex C
(informative)

Orthogonal design with 12 factor level combinations for 3 factors with 2 factor levels and 1 factor with 3 factor levels

C.1 Sample materials and factors

The study is conducted at $m=4$ different levels of the measurand. At each level, 12 test portions are selected from the sample material. Test portions are denoted 1 to 12 for level 1, 13 to 24 for level 2, 25 to 36 for level 3 and 37 to 48 for level 4.

3 factors each with 2 levels and 1 factor with 3 levels are selected that represent normal use of the method (e.g. 2 technicians, 2 storage temperatures for test portions, 2 batches of standard reagents, 3 sample batches, ...). Factors and factor levels are expected to reflect the normal variation within the single laboratory (see 7.2).

Each factor level combination is tested in a separate block, i.e. in different days or weeks.

C.2 Experimental design

Table C.1 represents the 12 factor level combinations which are conducted separately in 8 blocks. In each block 4 test portions at 4 different levels of the measurand are measured under repeatability conditions. The design can be used with ($p>1$) and without ($p=1$) replicates.

Table C.1 — Experimental design for 3 factors each at 2 factor levels and 1 factor with 3 factor levels

Factor level combination j	Test portions	Factors				
		1	2	3	4	Block (week)
1	1,13,25,37	1	1	1	1	1
2	2,14,26,38	1	2	2	1	2
3	3,15,27,39	2	2	2	2	3
4	4,16,28,40	2	1	1	2	4
5	5,17,29,41	1	1	1	3	5
6	6,18,30,42	1	2	2	3	6
7	7,19,31,43	2	1	2	1	7
8	8,20,32,44	2	2	1	1	8
9	9,21,33,45	1	2	1	2	9
10	10,22,34,46	1	1	2	2	10
11	11,23,35,47	2	1	2	3	11
12	12,24,36,48	2	2	1	3	12

C.3 Statistical analysis

The analysis is based on the statistical model shown in 7.4.1, Formula (2):

$$Y_{ijk} = \alpha + \beta x_{ij} + A(j) + B(j) x_{ij} + A_j + B_j x_{ij} + a_{ijk} + b_{ijk} x_{ij}$$

where Y_{ijk} denotes the measurement result of the test portion at level $i = 1, \ldots, 4$ for factor level combination $j = 1, \ldots, 12$.

Annex D
(informative)

Orthogonal design with 9 factor level combinations for 4 factors with 3 factor levels

D.1 Sample materials and factors

The study is conducted at $m=4$ different levels of the measurand. At each level, 9 test portions are selected from the sample material. Test portions are denoted 1 to 9 for level 1, 10 to 18 for level 2, 19 to 27 for level 3 and 28 to 36 for level 4.

4 factors each with 3 levels are selected that represent normal use of the method (e.g. 3 technicians, 3 instruments, 3 reagent batches, 3 sample batches, ...). Factors and factor levels are expected to reflect the normal variation within the single laboratory (see 7.2).

Each factor level combination is tested in a separate block, i.e. in different days or weeks.

D.2 Experimental design

Table D.1 represents the 9 factor level combinations which are conducted separately in 9 blocks. In each block 4 test portions at 4 different levels of the measurand are measured under repeatability conditions. The design can be used with ($p>1$) and without ($p=1$) replicates.

Table D.1 — Experimental design for 4 factors each at 3 factor levels

Factor level combination j	Test portions	Factors				
		1	2	3	4	Block (week)
1	1,10,19,28	1	1	1	1	1
2	2,11,20,29	1	2	2	2	2
3	3,12,21,30	1	3	3	3	3
4	4,13,22,31	2	1	2	3	4
5	5,14,23,32	2	2	3	1	5
6	6,15,24,33	2	3	1	2	6
7	7,16,25,34	3	1	3	2	7
8	8,17,26,35	3	2	1	3	8
9	9,18,27,36	3	3	2	1	9

D.3 Statistical analysis

If the design is used without (p = 1) replicates, the factor “Block” is confounded with factors 1 to 4 so that it is not possible to determine the variance components of the block effect separately. Its effect will inflate the effects of factors 1 to 4, so that the computed variance of $A(j)+B(j)x_{ij}$ already includes the block effect (day or week effect). In this case, the analysis is based on the statistical model

$$Y_{ijk} = \alpha + \beta x_{ij} + A(j) + B(j)x_{ij} + a_{ijk} + b_{ijk}x_{ij}$$

where Y_{ijk} denotes the measurement result of the test portion at level $i = 1,\ldots,4$ for factor level combination $j = 1,\ldots,9$ and $k = p = 1$.

If the design is used with (p>1) replicates, the block effect can be determined separately from the other factors. In this case, the analysis is based on the statistical model shown in 7.4.1, Formula (2):

$$Y_{ijk} = \alpha + \beta x_{ij} + A(j) + B(j)x_{ij} + A_j + B_j x_{ij} + a_{ijk} + b_{ijk}x_{ij}$$

where Y_{ijk} denotes the measurement result of the test portion at level $i = 1,\ldots,4$ for factor level combination $j = 1,\ldots,9$.

Annex E
(informative)

Asymmetric measurement uncertainty interval

If the relative in-house reproducibility standard deviation $\sigma_{Ri}(x)/\mu(x)$ is larger than 0,25, the coverage interval $[Y-U(Y),Y+U(Y)]$ or $[Y_{corr}-U(Y_{corr}),Y_{corr}+U(Y_{corr})]$ is not suitable in all cases to describe the range of possible values. For example, in some cases the lower limit of this interval is very close to zero or even negative, which is not appropriate for non-negative measurands. An alternative method is the Bayesian uncertainty, as described for example in JCGM 106:2012, 6.3[2].

A simpler alternative is to use the following method, which is characterised by the fact that mathematically, the measurement uncertainty function U is not a function of the measured value Y, but a function of the level x to be determined. Accordingly, all those values x belong to the measurement uncertainty interval $[x_L,x_U]$ for which the measured value Y lies within the interval $[x-U(x),x+U(x)]$.

x_L and x_U can be computed iteratively using the implicit relationships

$$x_L+U(x_L)=Y\ ,\ x_U-U(x_U)=Y\ .$$

If in the interval between $Y-U(Y)$ and $Y+U(Y)$ the measurement uncertainty does not change by more than 10 %, x_L and x_U can be approximated by $Y-U(Y)$ and $Y+U(Y)$, respectively.

Explanations are given in Reference [13].

NOTE If the true level of the measurand is equal to x_L, measurement values at or above x are very unlikely, as $x_L+U(x_L)=x$. If the true level of the measurand is equal to x_U, measurement values at or below x are very unlikely, as $x_U-U(x_U)=x$.