ISO/FDIS 19983:~~2023~~2024(E)

__ISO TC 45/SC 2

~~2023-12~~ 2024-03-15

Secretariat: JISC

Rubber — Determination of precision of test methods

Caoutchouc — Détermination de la fidélité des méthodes d'essai

~~Edited DIS - MUST BE USED FOR FINAL DRAFT~~

Contents

Foreword

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

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO document 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).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents.www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

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

For an explanation of 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 www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 45, *Rubber and rubber products*, Subcommittee SC 2, *Testing and analysis*.

This third edition cancels and replaces the second edition (ISO 19983:2022), which has been technically revised.

The main changes compared to the previous edition are as follows:

— A new Annex F has been addedmodified to provide an example of outlier treatment for Method A.;

— A— previous Annex F has been changed to Annex G;

— a comparison of Method A and Method B is shown in 6.7.3 and 6.7.4.

— Annex F at DIS stage was changed to Annex G

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

The procedures used for several years ~~by ISO/TC 45/SC 2~~ for estimating precision of test methods by means of interlaboratory tests (ISO/TR 9272[1]) were closely related to ASTM D4483. ISO/TR 9272 was found to have serious flaws which users were using work-arounds to counteract. ~~It was the desire of the ISO TC 45/SC2/WG4 members that~~This document has therefore been developed to replace ISO/TR 9272 ~~be replaced with a new standard that included~~and includes using ISO 5725 ~~(all parts)~~-1 and ISO 5725-2 with specific choices and variations of procedures to suit the particular requirements of rubbers.

This document provides two methods for determining the precision values of a test method:

— Method A based on ISO ~~5725 (all parts)~~ 5725-1, ISO 5725-2 and ISO 5725-3 to calculate repeatability, day-to-day repeatability, and reproducibility;

— Method B based on ASTM D4483 to calculate day-to-day repeatability and reproducibility.

[1] Cancelled and replaced by ISO 19983:2017.

Rubber — Determination of precision of test methods

1 Scope

This document provides guidelines and specifies requirements for estimating the precision of rubber test methods by means of interlaboratory test programmes based on the procedures given in:

— Method A using ISO 5725~~ (all parts);~~-1, ISO 5725-2 and ISO 5725-3;

— Method B using ASTM D4483.

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 5725-1:2023, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*

ISO 5725-2~~ : 2019.~~, *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 5725-3~~:2023~~, *Accuracy (trueness and precision) of measurement methods and results — Part 3: Intermediate measures of the precision of a standard measurement method*

~~ISO 5725-4:2020, *Accuracy (trueness and precision) of measurement methods and results — Part 4: Basic methods for the determination of the trueness of a standard measurement method*~~

~~ISO 5725-5:1998, *Accuracy (trueness and precision) of measurement methods and results — Part 5: Alternative methods for the determination of the precision of a standard measurement method*~~

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

ASTM D4483, *Standard Practice for Determining Precision for Test Method Standards in the Rubber and Carbon Black Industries*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 3534-1, ISO 3534-2, ISO 5725 ~~(all parts),~~-1, ISO 5725-2, ISO 5725-3, ASTM D4483, and the following apply.

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

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

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

3.1
day-to-day repeatability
precision under the conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment

Note 1 to entry: The time interval between repeated tests is normally between one and seven days.

3.2
type 1 precision
precision determined directly on a target material

Note 1 to entry: Prepared test pieces or test portions of the target material (class of elements) drawn from a homogeneous source are tested, with no processing or other operations required prior to testing.

3.3
type 2 precision
precision determined indirectly for a target material

Note 1 to entry: The target material is usually combined with a number of homogeneous ancillary materials to form a composite material and testing is conducted on samples of this and the property response of the target material is determined.

3.4
pooled standard deviation
square root of the average variance of a set of selected individual variances

Note 1 to entry: The pooled standard deviation, as well as the average variance, is intended as an overall or general descriptor of some set of variances and their standard deviations.

4 Symbols

D_{ij}	day-to-day effect, the day-to-day variance component of which is σ_D^2
h values	Mandel's between-laboratory consistency test statistic
k values	Mandel's within-laboratory consistency test statistic
L_i	between-laboratory effect, the between-laboratory variance component of which is σ_L^2
M_{ijk}	repeatability effect, the repeatability variance component of which is σ_M^2
n	number of measurements
p	number of laboratories

q	number of days
r	repeatability
r_{DA}	day-to-day repeatability as determined from method A calculations
r_{DB}	day-to-day repeatability as determined from method B calculations
R	reproducibility
(r)	relative repeatability
(r_{DA})	relative day-to-day repeatability as determined from method A calculations
(r_{DB})	relative day-to-day repeatability as determined from method B calculations
(R)	relative reproducibility
$s_M{}^2$	repeatability variance
$s_{rD}{}^2$	day-to-day repeatability variance as determined from method A calculations
$s_R{}^2$	reproducibility variance
$s_D{}^2$	day-to-day variance as determined from method B calculations
$s_L{}^2$	between-laboratory variance
s	standard deviation of data
s_r	repeatability standard deviation
s_{rD}	day-to-day repeatability standard deviation as determined from method A calculations
s_R	reproducibility standard deviation
S_D	day-to-day repeatability standard deviation as determined from method B calculations
SS_T	total sum of squares
SS_L	between-laboratory sum of squares
SS_D	day-to-day sum of squares
SS_M	repeatability sum of squares
T	total sum of data
V_L	between-laboratory mean square
V_D	day-to-day mean square
V_M	repeatability mean square
y_{ijk}	data *i*, *j*, *k*: each data of laboratory, day, repeat
~~$\bar{y}$~~ $\bar{y}$	mean value of data
~~$\bar{\bar{y}}$~~ $\bar{\bar{y}}$	mean value of ~~$\bar{y}$~~ $\bar{y}$
ϕ_T	total degree of freedom
ϕ_L	between-laboratory degree of freedom
ϕ_D	day-to-day degree of freedom
ϕ_M	repeatability degree of freedom
μ	population mean

σ_M^2	repeatability variance component
σ_D^2	day-to-day variance component
σ_L^2	between-laboratory variance component

NOTE The symbols r_{DB} and (r_{DB}) in this document are the same as r and (r), respectively, in ASTM D4483.

5 Interlaboratory test programme

To evaluate precision for test method standards by means of interlaboratory test programmes (ITPs), use either one of the two methods:

a) Method A, where three precisions, namely the repeatability, the day-to-day repeatability and the reproducibility, are calculated in accordance with ISO 5725-3;

b) Method B, where two precisions, namely the day-to-day repeatability and the reproducibility, are calculated in accordance with ASTM D4483.

NOTE If two or more results are available from within-a-day repeated tests, ~~method~~Method A is applicable to evaluate the variance of measurement errors.

6 Procedures

6.1 Application

A standard measurement method is taken to mean an established international test method for rubber.

A determination of the precision of a test method is normally conducted with a selected group of materials typical of those used with that method, and by a group of volunteer laboratories that have experience of the method.

Caution is necessary in applying precision results for a particular test method to product testing for commercial product accepted procedures. For this purpose, the precision estimates should be obtained from special programmes that are specific to the product in question and carried out by the interested laboratories.

6.2 Repeatability conditions

Repeatability conditions are where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time.

NOTE "Short interval of time" indicates that tests are repeated within a day, when the time needed to complete a test allows repeating the test within the same day.

"Identical test items" is interpreted as nominally identical, i.e. no intentional differences.

For rubbers, repeatability can be dependent on the magnitude or level of the measured property and is usually reported for each of several materials having particular property levels.

6.3 Day-to-day repeatability conditions

Day-to-day repeatability conditions are where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment.

The “intervals of time” between repeated measurements of test results may be selected by the consensus of a particular testing community. For ~~ISO/TC 45 and~~ the international rubber manufacturing industry, the time interval between repeat tests is of the order of one to seven days, but most commonly seven days. For special tests (long ageing periods), however, replicate tests can require a longer time span.

NOTE The “repeatability” traditionally used ~~by ISO/TC 45/SC 2~~ is equivalent to the day-to-day repeatability defined in this document.

6.4 Reproducibility conditions

Reproducibility conditions are where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment.

"Identical test items" is interpreted as nominally identical, i.e. no intentional differences.

~~For ISO/TC 45, different~~Different equipment means apparatus that ~~might~~can have different manufacturers but complies with the requirements of the test standard in question, including calibration.

For rubbers, reproducibility ~~might~~can possibly be dependent on the magnitude or level of the measured property and is usually reported for each of several materials having particular property levels.

6.5 Testing elements

The element that is tested is either a test piece or a test sample as defined in the test method standard. The test method standard will also define the number of test elements to be tested to obtain a result for the property.

6.6 Planning

Select either type 1 precision or type 2 precision as defined in 3.2 and 3.3.

It is possible that a type 1 precision programme can be conducted on test pieces or portions that require some minimum processing or other simple operations prior to actual testing.

Unless circumstances dictate otherwise, using type 1 precision is preferred.

For type 1 precision, the test pieces or test samples need to be produced from the same lot of material by the same procedures and then stored and conditioned in the same manner, in order to be nominally identical. This is best achieved by test pieces being prepared in one laboratory and distributed to the others with instructions for conditioning.

For type 2 precision, the properties of the composite material are directly related to the quality of properties of the target material. As an example, to determine the quality of a grade of SBR, a sample of the rubber plus curatives, fillers, antioxidants, etc. are mixed and cured. The precision of the resulting test pieces is determined and reflects sample preparation and the properties response of the target SBR.

The estimation of precision for rubber test methods is normally conducted using a balanced uniform level design with three or more materials sent to each participating laboratory with tests conducted to yield an independent test result by the same technician on each of two test days.

NOTE A balanced uniform level design is a plan for an interlaboratory test programme for precision, where all laboratories test all the materials selected for the programme and each laboratory conducts the same number of repeated tests, *n*, on each material.

The test method, materials, participating laboratories, test equipment and time interval for test in a laboratory are addressed in 6.1 to 6.6. Other aspects of planning shall be addressed in accordance with ISO 5725-1:~~1994~~2023, Clause 6.

6.7 Methodology

6.7.1 Method A

Method A determines the repeatability variance component (measurement error component) σ_M^2, the day-to-day variance component σ_D^2 and the between-laboratory variance component σ_L^2, by calculating the expected mean square in accordance with a suitable ANOVA table in ISO 5725-3, fully-nested experiments.

Then, the day-to-day repeatability variance s_{rD}^2 and the reproducibility variance s_R^2 are given by Formulae (1) and (2):

$$s_{rD}^2 = \sigma_M^2 + \sigma_D^2 \quad s_{rD}^2 = \sigma_M^2 + \sigma_D^2 \qquad (1)$$

Field Code Changed

$$s_R^2 = \sigma_M^2 + \sigma_D^2 + \sigma_L^2 \quad s_R^2 = \sigma_M^2 + \sigma_D^2 + \sigma_L^2 \qquad (2)$$

Field Code Changed

The repeatability, r, the day-to-day repeatability, r_{DA}, and the reproducibility, R, are given by Formulae (3), (4), and (5), respectively:

$$r = 2{,}83\left(s_M^2\right)^{\frac{1}{2}} = 2{,}83\left(\sigma_M^2\right)^{\frac{1}{2}} = 2{,}83 s_M = 2{,}83\sigma_M$$

$$r = 2{,}83\left(s_M^2\right)^{\frac{1}{2}} = 2{,}83\left(\sigma_M^2\right)^{\frac{1}{2}} = 2{,}83 s_M = 2{,}83\sigma_M \qquad (3)$$

Field Code Changed

$$r_{DA} = 2{,}83\left(s_{rD}^2\right)^{\frac{1}{2}} = 2{,}83 s_{rD} \quad r_{DA} = 2{,}83\left(s_{rD}^2\right)^{\frac{1}{2}} = 2{,}83 s_{rD} \qquad (4)$$

Field Code Changed

$$R = 2{,}83\left(s_R^2\right)^{\frac{1}{2}} = 2{,}83 s_R \quad R = 2{,}83\left(s_R^2\right)^{\frac{1}{2}} = 2{,}83 s_R \qquad (5)$$

Field Code Changed

Calculations for method A shall be in accordance with Annex A. An example is given in D.3.

For rubber tests, it is usually possible to have two or more repeated tests within one day.

6.7.2 Method B

Method B determines the day-to-day variance (between-day variance), s_D^2, the between-laboratory variance s_L^2 and the reproducibility variance s_R^2 (which is equal to $s_L^2 + s_D^2$), according to the calculation procedures in ASTM D4483.

The day-to-day repeatability, r_D, and the reproducibility, R, are given by Formulae (6) and (7):

$$r_{DB} = 2{,}83\left(s_D^2\right)^{\frac{1}{2}} = 2{,}83 s_D \quad r_{DB} = 2{,}83\left(s_D^2\right)^{\frac{1}{2}} = 2{,}83 s_D \qquad (6)$$

Field Code Changed

$$R = 2{,}83\left(s_R^2\right)^{\frac{1}{2}} = 2{,}83 s_R \quad R = 2{,}83\left(s_R^2\right)^{\frac{1}{2}} = 2{,}83 s_R \qquad (7)$$

Field Code Changed

Calculations for methodMethod B shall be in accordance with Annex B for 2two test results or ASTM D4483 for more than 2two test results. An example is given in D.4.

When there are two or more data from repeated tests (individual determinations) within the same day, estimate the median values or the mean values, as appropriate, and apply them for the method B procedures.

6.7.3 Method A ~~vs~~versus Method B — Day-to-day repeatability value

Refer to the NOTE ~~D.2~~ in ~~Annex~~ D.4 for the reason why r_{DA} is large when r_{DA} and r_{DB} are calculated with the same data set.

When calculating with Method B, even if there is repeated data within a day, the average value or median value is used for calculation without considering the variation.

Therefore, the intra-day variation r cannot be evaluated in Method B.

If intra-day repeat data exist, Method A should be used.

6.7.4 Method A ~~vs~~versus Method B — Number of replicates

~~The~~In the accuracy ~~standard of TC 45~~standards (ISO/TR 9272~~,~~ and ISO 19983:~~1997) limited to two repetitions per day~~2017[2]), days could only be repeated twice.

~~This~~The reason is ~~because~~that the accuracy calculation of s_D^2 in Method B can only repeat in a day = 2 times. See Formula D.17 in Annex D.

On the other hand, in the accuracy calculation of Method A, based on a suitable ANOVA table according to ISO 5725-3, the number of repetitions of measurement and the number of repetitions of day are not limited to two.

~~Replaced~~ Additional values with n = 3 to 5 were added to the n column of Table C.2 ~~in Annex C~~ (Mandel's test table~~) so that the test can be performed even if the repetition of days is 2 to 5. from Mandel's table published in ISO 5725-2.~~).

6.8 Detection of outliers

For detecting outliers, this document adopts two measures called Mandel's h and k statistics. The h statistic is a parameter used to review the difference between averages, while the k statistics is a parameter used to review the difference between variances. This treatment is applied separately for h and k for each material. It may be noted that, as well as describing the variability of the measurement method, these help in laboratory evaluation. The calculation of h and k statistic values and the determination of their critical values at 5 % significance level for replicates (n = 2~~~~ to 5) shall be in accordance with Annex C.

For some test methods, a test result is defined as some statistical parameter, such as mean or median, calculated from the individual measurements. For ~~method A~~Method B, Mandel's h and k statistics are calculated using average and standard deviation values of repeated individual test measurement data usually within the same day. These individual measurements are used to calculate a test result. For ~~method B~~Method A, Mandel's h and k statistics are calculated using average and standard deviation values of test result data from performing the test method on ~~2~5~~approximately two to five different days.

6.9 Treatment of outliers

As with outlier detection, the treatment of outliers is performed separately for the h and k statistics for each material. There are several techniques that can be used for outlier treatment, such as data deletion, data replacement, parameter replacement, parameter deletion and retesting. It is also possible to not treat the outliers and keep them in the precision calculations. There are several methods for deriving the

[2] Cancelled and replaced by ISO 19983:2022.

replacement values. Once an outlier treatment method is selected, that method is applied to all outliers at a given significance level.

Each treatment option has advantages and disadvantages. Each laboratory/material is evaluated by the *h* and *k* statistics. A given laboratory/material may be an outlier for *h*, *k*, or both. Most laboratory/materials will not be outliers. A small percentage of laboratory/materials will be outliers for *h* or *k* but not for both. A rare few laboratories will be outliers for all, or nearly all, materials. When a laboratory is an outlier for all materials, all of that laboratory's data should be deleted, the critical *h* and *k* statistics recalculated for a lower number of laboratories, and the *h* and *k* detection of outliers repeated at the new critical *h* and *k* values. When a laboratory is an outlier for nearly all the materials, the analyst should consider removing all that laboratory's data and proceeding as above.

a) Data deletion. Most laboratories that have an outlier value only have it for either *h* or *k* but not both. Deleting a laboratory's data that has an outlier for *h* or *k* also removes the non-outlier data from the precision calculations. For example, a laboratory has an outlier for *h* but not *k*. Deleting the laboratory's data removes both the mean and standard deviation values, so no *h* or *k* values are calculated. However, only the mean was an outlier so the standard deviation value could have been used to calculate the precision. Parameter replacement or parameter deletion for the mean outlier would have kept the non-outlier standard deviation value in the precision calculations. A similar condition occurs with *k* outliers.

b) Parameter replacement. There are several techniques that can be used to calculate a replacement value for the average or standard deviation parameters for outlier laboratories. Parameter replacement may alter the data distribution and impact the precision values. Some examples of techniques to use to obtain parameter replacement values are:

   1) Mean:

      i) average of all non-outlier mean values;

      ii) ascending order trend (AOT) of mean values;

      iii) mean ±2 standard deviations;

   2) Standard Deviation:

      i) average of all non-outlier variances to calculate the replacement standard deviation value;

      ii) ascending order trend (AOT) of standard deviation values;

      iii) mean ±2 standard deviations is ~~NOT~~not an option for standard deviation replacement values because standard deviations and variances do not have a normal distribution.

c) Data replacement. When there are only two replicates, the parameter replacement value and the data range can be used to calculate replacement data values for any outlier. This requires more effort than parameter replacement and has the same possibility of altering the data distribution and precision values.

d) Parameter deletion. Parameter deletion requires the least effort. However, this may not be an option for small data sets because it reduces the number of laboratories data that is used in the precision calculations.

e) Retesting. Retesting introduces additional elements of variation into the precision parameters. During the elapsed time between the initial testing and the retest, laboratory conditions, such as