TECHNICAL SPECIFICATION

ISO/TS 19138

First edition
2006-12-01

Geographic information — Data quality measures

Information géographique — Mesures de la qualité des données

Reference number
ISO/TS 19138:2006(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword iv
Introduction v
1 Scope 1
2 Conformance 1
3 Normative references 1
4 Terms and definitions 1
5 Relationships to other standards 2
6 Register 3
7 Components of a data quality measure 3
7.1 List of components 3
7.2 Component details 4
7.3 Standardized data quality measures 6
Annex A (normative) Abstract test suite 7
A.1 Test case identifier: Component test 7
A.2 Test case identifier: Name test 7
A.3 Test case identifier: Data quality element and subelement test 7
A.4 Test case identifier: Data quality basic measure test 8
A.5 Test case identifier: Definition test 8
A.6 Test case identifier: Description test 8
A.7 Test case identifier: Parameter test 8
A.8 Test case identifier: Data quality value type test 8
A.9 Test case identifier: Source reference test 9
A.10 Test case identifier: Example test 9
Annex B (normative) Structure of data quality measures 10
B.1 Components defining a data quality measure 10
B.2 Mapping of the components to ISO 19115 and ISO 19135 11
B.3 UML-diagram for data quality measure 11
Annex C (normative) Data quality basic measures 14
C.1 Purpose of data quality basic measures 14
C.2 Counting-related data quality basic measures 14
C.3 Uncertainty-related data quality basic measures 15
Annex D (normative) List of data quality measures 19
D.1 Completeness 19
D.2 Logical consistency 23
D.3 Positional accuracy 37
D.4 Temporal accuracy 58
D.5 Thematic accuracy 61
Bibliography 68

Foreword

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

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In other circumstances, particularly when there is an urgent market requirement for such documents, a technical committee may decide to publish other types of normative document:

— an ISO Publicly Available Specification (ISO/PAS) represents an agreement between technical experts in an ISO working group and is accepted for publication if it is approved by more than 50 % of the members of the parent committee casting a vote;

— an ISO Technical Specification (ISO/TS) represents an agreement between the members of a technical committee and is accepted for publication if it is approved by 2/3 of the members of the committee casting a vote.

An ISO/PAS or ISO/TS is reviewed after three years in order to decide whether it will be confirmed for a further three years, revised to become an International Standard, or withdrawn. If the ISO/PAS or ISO/TS is confirmed, it is reviewed again after a further three years, at which time it must either be transformed into an International Standard or be withdrawn.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/TS 19138 was prepared by Technical Committee ISO/TC 211, *Geographic information/Geomatics*.

Introduction

Knowledge of the quality of geographic data is often crucial for the application of the data, as different users and different applications often have different data quality requirements. A user of geographic data may have multiple datasets from which to choose. Therefore, it is necessary to compare the quality of the datasets to determine which best fulfils the requirements of the user. To facilitate such comparisons, it is essential that the results of the quality reports are expressed in a comparable way and that there is a common understanding of the data quality measures that have been used. These data quality measures provide descriptors of the quality of geographic data through comparison with the universe of discourse. The use of incompatible measures makes data quality comparisons impossible to perform.

Data quality needs to be reported by the producer and evaluated by the user against his or her requirements for different criteria and data quality measures. It is essential that reported quality for a dataset contains the quality measurements that may be of interest to a potential user of the dataset, and that the metrics used to determine the quality are reported and available to the user.

ISO 19113 establishes the principles for the description of geographic data quality and specifies components for reporting quality information. Procedures for the evaluation of geographic data quality are described in ISO 19114.

The objective of this Technical Specification is to guide the producer in choosing the right data quality measures for data quality reporting, and the user in the evaluation of the usefulness of a dataset by standardizing the components and structures of data quality measures and by defining commonly used data quality measures.

Geographic information — Data quality measures

1 Scope

This Technical Specification defines a set of data quality measures. These can be used when reporting data quality for the data quality subelements identified in ISO 19113. Multiple measures are defined for each data quality subelement, and the choice of which to use will depend on the type of data and its intended purpose.

The data quality measures are structured so that they can be maintained in a register established in conformance with ISO 19135.

This Technical Specification does not attempt to describe every possible data quality measure, only a set of commonly used ones.

2 Conformance

Any set of data quality measures claiming conformance with this Technical Specification shall pass all of the conditions specified in the abstract test suite (Annex A).

3 Normative references

The following referenced documents are indispensable for the application 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/TS 19103:2005, *Geographic information — Conceptual schema language*

ISO 19113:2002, *Geographic information — Quality principles*

ISO 19115:2003, *Geographic information — Metadata*

ISO 19135:2005, *Geographic information — Procedures for item registration*

4 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

4.1
correctness
correspondence with the universe of discourse

4.2
data quality basic measure
generic data quality measure used as a basis for the creation of specific data quality measures

NOTE Data quality basic measures are abstract data types. They cannot be used directly when reporting data quality.

4.3
data quality scope
extent or characteristic(s) of the data for which quality information is reported

[ISO 19113]

NOTE A data quality scope for a dataset can comprise a dataset series to which the dataset belongs, the dataset itself, or a smaller grouping of data located physically within the dataset sharing common characteristics. Common characteristics can be an identified feature type, feature attribute, or feature relationship; data collection criteria; original source; or a specified geographic or temporal extent.

4.4
error
discrepancy with the universe of discourse

4.5
measurand
particular quantity subject to measurement

[International Vocabulary of Basic and General Terms in Metrology (VIM)]

4.6
universe of discourse
view of the real or hypothetical world that includes everything of interest

[ISO 19101]

5 Relationships to other standards

ISO 19113 describes relevant data quality elements and their corresponding data quality subelements and it indicates how quality should be reported. ISO 19114 describes procedures for the evaluation of quantitative quality. ISO 19115 contains elements and classes for data quality reporting within the UML models and data dictionaries.

ISO 19113 specifies a set of descriptors for a data quality subelement, for use in recording data quality. One of these descriptors is the data quality measure. A data quality measure is described by the components listed in 7.1.

Table 1 provides a list of data quality elements and data quality subelements as defined in ISO 19113.

Table 1 — Data quality elements and data quality subelements with definitions (ISO 19113)

Data quality element	Data quality subelement	Definition
completeness	commission	excess data present in a dataset
	omission	data absent from a dataset
logical consistency	conceptual consistency	adherence to rules of the conceptual schema
	domain consistency	adherence of values to the value domains
	format consistency	degree to which data is stored in accordance with the physical structure of the dataset
	topological consistency	correctness of the explicitly encoded topological characteristics of a dataset
positional accuracy	absolute or external accuracy	closeness of reported coordinate values to values accepted as or being true
	relative or internal accuracy	closeness of the relative positions of features in a dataset to their respective relative positions accepted as or being true
	gridded data position accuracy	closeness of gridded data position values to values accepted as or being true
temporal accuracy	accuracy of a time measurement	correctness of the temporal references of an item (reporting of error in time measurement)
	temporal consistency	correctness of ordered events or sequences, if reported
	temporal validity	validity of data with respect to time
thematic accuracy	classification correctness	comparison of the classes assigned to features or their attributes to a universe of discourse (e.g. ground truth or reference dataset)
	non-quantitative attribute correctness	correctness of non-quantitative attribute
	quantitative attribute accuracy	accuracy of quantitative attributes

6 Register

A register of data quality measures shall contain a set of data quality measures, described using the components listed in 7.1. The registration procedures shall be performed according to ISO 19135.

Annex D of this Technical Specification contains the list of standardized data quality measures. A register shall contain these data quality measures and may also contain additional data quality measures submitted through the procedures defined within ISO 19135. The registration process also allows retiring data quality measures.

7 Components of a data quality measure

7.1 List of components

Each data quality measure shall be described using the following technical components:

— name (7.2.1)

— alias (7.2.2)

- data quality element (7.2.3)
- data quality subelement (7.2.4)
- data quality basic measure (7.2.5)
- definition (7.2.6)
- description (7.2.7)
- parameter (7.2.8)
- data quality value type (7.2.9)
- data quality value structure (7.2.10)
- source reference (7.2.11)
- example (7.2.12)
- identifier (7.2.13)

7.2 Component details

7.2.1 Name

Name refers to the name of the data quality measure.

If the data quality measure already has a commonly used name, this name should be used. If no name exists, a name shall be chosen that reflects the nature of the measure.

NOTE The component name is specified in the base standard for registers, ISO 19135.

7.2.2 Alias

Alias refers to other recognized name for the same data quality measure. It may be a different commonly used name, or an abbreviation or a short name.

More than one alias may be provided.

7.2.3 Data quality element

Data quality element refers to the name of the data quality element to which this data quality measure applies.

NOTE A list of data quality elements is provided in Table 1.

7.2.4 Data quality subelement

Data quality subelement refers to the name of the data quality subelement to which this data quality measure applies.

NOTE A list of data quality subelements is provided in Table 1.

7.2.5 Data quality basic measure

Each data quality basic measure is described by its name, definition and value type. Data quality basic measures are identified by their names.

A variety of data quality measures are based on counting of erroneous items. There are also several data quality measures dealing with the uncertainty of numerical values. In order to avoid repetition, all possible methods of constructing counting-related data quality measures as well as general statistical measures for one- and two-dimensional random variables shall be defined in terms of data quality basic measures.

The data quality basic measures are defined in Annex C.

If a data quality measure is based on one of the set of data quality basic measures, the name of the data quality basic measure shall be provided in the field data quality basic measure. If the data quality measure is not based on a data quality basic measure, it shall be indicated in this field that a data quality basic measure is not applicable. The data quality basic measures shall also be used as appropriate for creating new data quality measures, for instance for reporting unclosed surface patches or other application-dependent data quality measures.

7.2.6 Definition

Definition states the fundamental concept of the data quality measure.

If the data quality measure is derived from a data quality basic measure, the definition is based on the data quality basic measure definition and specialized for this data quality measure.

NOTE The component definition is specified in the base standard for registers, ISO 19135.

7.2.7 Description

Description refers to the description of the data quality measure including methods of calculation, with all formulae and/or illustrations needed to establish the result of applying the measure.

If the data quality measure uses the concept of errors, it shall be stated how an item shall be classified as incorrect.

NOTE The component description is specified in the base standard for registers, ISO 19135.

7.2.8 Parameter

Parameter refers to an auxiliary variable used by the data quality measure. It shall include name, definition and description.

More than one parameter may be provided.

7.2.9 Data quality value type

Data quality value type refers to the value type for reporting a data quality result.

A data quality value type shall be provided for a data quality result. The data types defined in ISO/TS 19103 shall be used when appropriate.

Table 2 — Examples of data quality value types

Boolean
Real
Integer
Ratio (numerator of type integer : denominator of type integer)
Percentage
Measure(s) [value(s) + unit(s)]

7.2.10 Data quality value structure

Data quality value structure gives the structure for reporting a complex data quality result.

A data quality result may consist of multiple values. In this case, the data quality result shall be structured using the data quality value structures as given in Table 3. The structure may consist of homogeneous or heterogeneous data quality value types. The possible data quality value types are given in 7.2.9.

Table 3 — Data quality value structures

Bag
Set
Sequence
Table
Matrix
Coverage

NOTE The values within a structure can be multiple. For example, the covariance matrix as given in Table D.32 is reported as matrix of measure, where the matrix elements may have different units of measure. A list may consist of different data quality value types.

7.2.11 Source reference

Source reference gives the citation of the source of the data quality measure.

When a data quality measure for which additional information is provided in an external source is added to the list of standardized data quality measures, a reference to that source may be provided here.

NOTE The component source reference is specified in the base standard for registers, ISO 19135.

7.2.12 Example

Example may provide examples of applying the data quality measure or the result obtained for the data quality measure.

More than one example may be provided.

7.2.13 Identifier

Identifier consists of an integer number that uniquely identifies a data quality measure.

If data quality measures are administered in a register, then identifiers may only be assigned by the register manager.

NOTE The component identifier is specified in the base standard for registers, ISO 19135.

7.2.14 Obligation of the above-listed components

Some of the components are mandatory, others are conditional or optional. Table B.1 provides further information on the obligation of each technical component.

7.3 Standardized data quality measures

In order to make data quality related metadata and data quality reports comparable, standardized data quality measures shall be used in evaluating and reporting data quality, where appropriate. Annex D gives a list of commonly used data quality measures with all required components for data quality measures as specified in this Technical Specification.

Annex A
(normative)

Abstract test suite

A.1 Test case identifier: Component test

a) Test purpose: to determine conformance by ensuring that all necessary components of a data quality measure are provided.

b) Test method: examine the entry for the data quality measure and verify that the components have been provided as required by Table B.1.

c) Reference: 7.2 and Annex B.

d) Test type: Capability.

A.2 Test case identifier: Name test

a) Test purpose: to determine if a distinct name for the data quality measure is used.

b) Test method: determine if the name for the data quality measure is distinct from other measures with different concepts, and if the name is not in conflict with other data quality basic measures, their definitions and descriptions.

c) Reference: 7.2.1.

d) Test type: Capability.

A.3 Test case identifier: Data quality element and subelement test

a) Test purpose: to determine

— if data quality element and subelement are assigned;

— if they are taken from the list of data quality elements and subelements in ISO 19113 or if they are an additional data quality element and subelement created in conformance with the rules of ISO 19113;

— if the data quality measure is relevant for the given data quality element and subelement.

b) Test method: check if proper values are assigned to the data quality element and subelement components and if the data quality measure has bearing on these.

c) Reference: 7.2.3 and 7.2.4.

d) Test type: Capability.

A.4 Test case identifier: Data quality basic measure test

a) Test purpose: to determine if a data quality measure is properly derived from a data quality basic measure.

b) Test method: check if an appropriate data quality basic measure for the data quality measure exists and, if it does, that the data quality measure is utilizing this data quality basic measure in conformance with this Technical Specification.

c) Reference: 7.2.5.

d) Test type: Capability.

A.5 Test case identifier: Definition test

a) Test purpose: to determine if a fitting, correct and complete definition is provided.

b) Test method: check that the given definition contains no ambiguities and that it is in conformance with characteristics of a definition as stated in ISO 19135:2005, 7.3.1.

c) Reference: 7.2.6 and ISO 19135:2005, 7.3.1.

d) Test type: Capability.

A.6 Test case identifier: Description test

a) Test purpose: to determine if an exhaustive description is provided.

b) Test method: check if the description contains a comprehensive explanation with all required formulae to facilitate the application of the data quality measure.

c) Reference: 7.2.7.

d) Test type: Capability.

A.7 Test case identifier: Parameter test

a) Test purpose: to determine if required parameters are provided.

b) Test method: check if all parameters occurring in the description are provided in the parameter component.

c) Reference: 7.2.8.

d) Test type: Capability.

A.8 Test case identifier: Data quality value type test

a) Test purpose: to determine if a proper data quality value type is provided.

b) Test method: check if the provided data quality value type is included in the list in Table 3.

c) Reference: 7.2.9.

d) Test type: Capability.

A.9 Test case identifier: Source reference test

a) Test purpose: to determine if a proper source reference is provided.

b) Test method: check if the cited reference source exists and if it reflects the concept of the provided data quality measure.

c) Reference: 7.2.11.

d) Test type: Capability.

A.10 Test case identifier: Example test

a) Test purpose: to determine if the example, if provided, is a valid example for the data quality measure.

b) Test method: check if the example is free of errors and if it is representative of the usage of the data quality measure.

c) Reference: 7.2.12.

d) Test type: Capability.

Annex B
(normative)

Structure of data quality measures

B.1 Components defining a data quality measure

Table B.1 shall be used for the technical specification of every data quality measure. The descriptor for obligation/condition may have the following values: M (mandatory), C (conditional), or O (optional).

Table B.1 — Components defining a data quality measure

Line	Component	Description	Obligation/condition
1	Name	Name of the data quality measure applied to the data	M
2	Alias [a]	Another recognized name, an abbreviation or a short name for the same data quality measure	O
3	Data quality element	Name of the data quality element for which quality is reported	M
4	Data quality subelement	Name of the data quality subelement for which quality is reported	M
5	Data quality basic measure	Name of the data quality basic measure from which the data quality measure is derived	C/if derived from basic measure
6	Definition	Definition of the fundamental concept for the data quality measure	M
7	Description	Description of the data quality measure, including all formulae and/or illustrations needed to establish the result of applying the measure	C/if the definition is not sufficient for the understanding of the data quality measure concept
8	Parameter [a]	Auxiliary variable used by the data quality measure, including its name, definition and optionally its description	C/if required
9	Data quality value type [a]	Value type for reporting a data quality result	M
10	Data quality value structure	Structure for reporting a complex data quality result	O
11	Source reference [a]	Reference to the source of an item that has been adopted from an external source	C/if an external source exists
12	Example [a]	Illustration of the use of a data quality measure	O
13	Identifier	Integer number, uniquely identifying a data quality measure	C/if data quality measures are administered in a register

[a] Multiple entries are allowed. When values for the optional or conditional elements are not present, this should be indicated by assigning the character "—" to the appropriate component.

B.2 Mapping of the components to ISO 19115 and ISO 19135

Table B.2 — Mapping of the components to ISO 19115 and ISO 19135

Line	Component	ISO 19115 element name	ISO 19135 element name
1	Name	nameOfMeasure	name
2	Alias	–	alternativeExpressions
3	Data quality element	DQ_Element	–
4	Data quality subelement	lines 108-127 [B.2.4.3 Data quality element information]	–
5	Data quality basic measure	–	–
6	Definition	–	definition
7	Description	measureDescription	description
8	Parameter	–	–
9	Data quality value type	–	–
10	Data quality value structure	–	–
11	Source reference	–	source
12	Example	–	–
13	Identifier	measureIdentification	itemIdentifier

B.3 UML-diagram for data quality measure

Figure B.1 defines the components for data quality measures and Figure B.2 defines the relationship of data quality measures and registered items from ISO 19135. Both figures are in UML notation.

The UML models describe the content model, if a register for data quality measures is implemented.

The class RE_RegisteredItem is defined in ISO 19135.

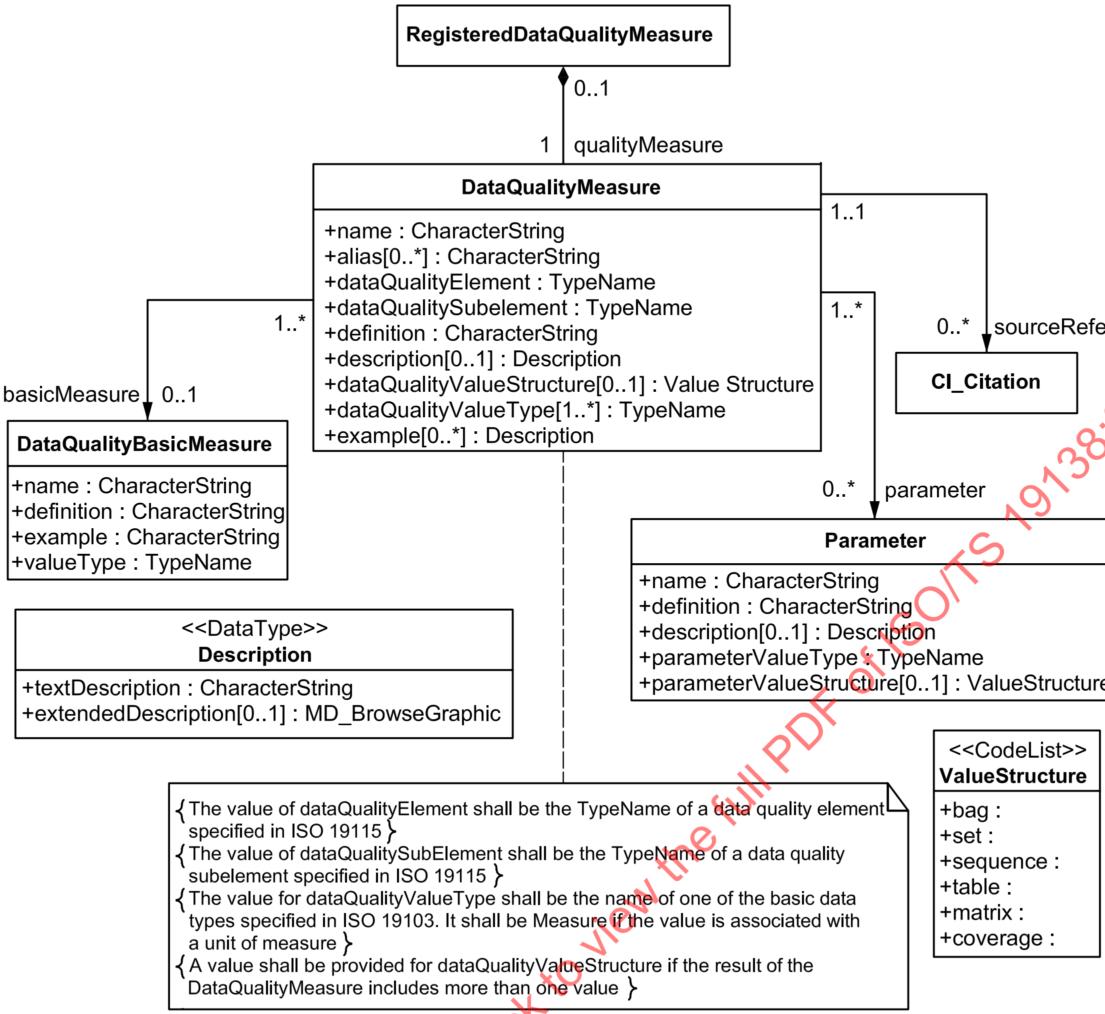


Figure B.1 — Data quality measure

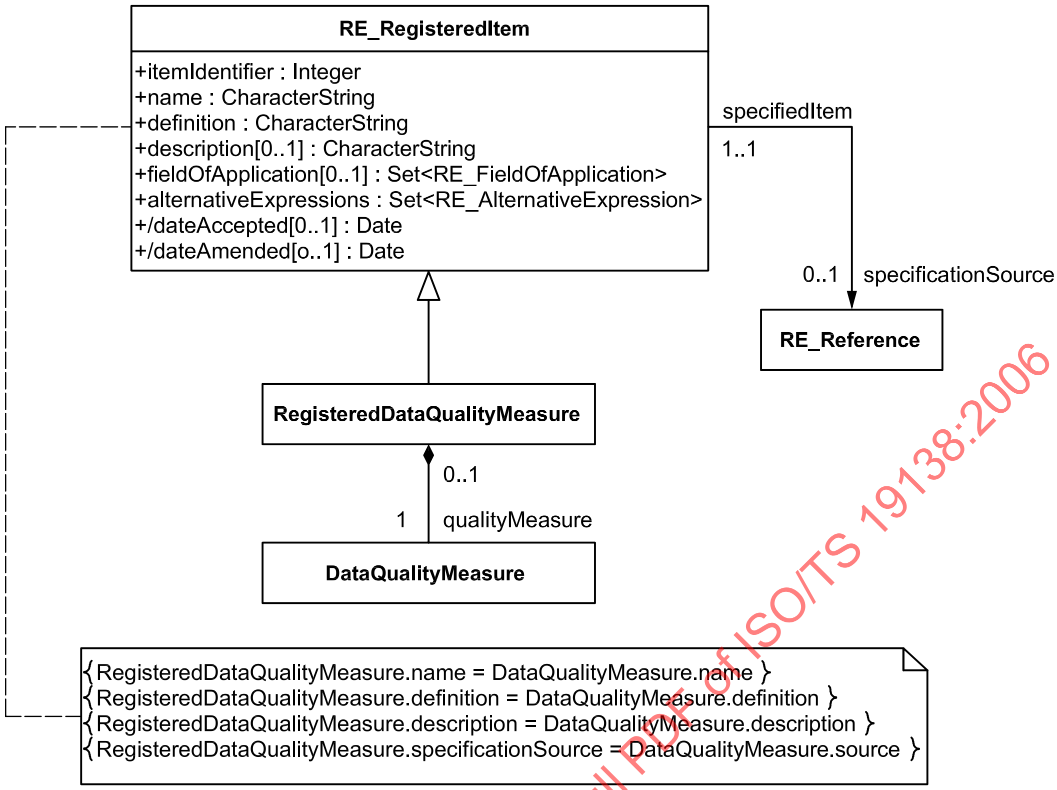


Figure B.2 — Relationship between registered item of ISO 19135 and data quality measure

Annex C
(normative)

Data quality basic measures

C.1 Purpose of data quality basic measures

The concept of data quality basic measures is introduced in this Technical Specification to avoid the repetitive definition of the same concept. There are data quality measures that have certain communalities. For example, the counting-related data quality measures are dealing with the concept of counting errors. The number of errors may be used to construct different kind of data quality measures. The concept of constructing these data quality measures is defined for the generic data quality basic measures and shall be used for the creation of data quality measures that share these communalities.

Counting- and uncertainty-related data quality measures can be identified. Therefore two principle categories of data quality basic measures are listed in this annex. The counting-related data quality basic measures are based on the concept of counting errors or correct items. The uncertainty-related data quality basic measures are based on the concept of modelling the uncertainty of measurements with statistical methods. The measured quantity can be embedded in different dimensions. Depending on the dimension of the measured quantity, different types of data quality basic measures shall be used to construct data quality measures.

Annex D uses the data quality basic measures of Annex C where appropriate. When appropriate, the construction of new data quality measures shall be derived from one of the following data quality basic measures.

C.2 Counting-related data quality basic measures

The data quality basic measures based on different methods of counting errors or counting the number of correct values are listed in Table C.1.

Table C.1 — Data quality basic measures for counting-related data quality measures

Data quality basic measure name	Data quality basic measure definition	Example	Data quality value type
Error indicator	Indicator that an item is in error	False	Boolean (if the value is true the item is not correct)
Correctness indicator	Indicator that an item is not in error	True	Boolean (if the value is true the item is correct)
Error count	Total number of items that are subject to an error of a specified type	11	Integer
Correct items count	Total number of items that are free of errors of a specified type	571	Integer
Error rate	Number of the erroneous items with respect to the total number of items that should have been present	0,0189 1,89% 11,582	Error rate can either be presented as real, percentage or as ratio
Correct items rate	Number of the correct items with respect to the total number of items that should have been present	0,9811 98,11% 571:582	Correct items rate can either be presented as real, percentage or as ratio

NOTE 1 Number of items is defined using number of items in the universe of discourse for the dataset specified by data quality scope.

EXAMPLE Use number of items found in the real world or reference dataset.

NOTE 2 A list of data quality value types is provided in Table 2 (see 7.2.9).

C.3 Uncertainty-related data quality basic measures

C.3.1 General

Numerical values that are obtained by some kind of measuring procedure can only be observed to a certain accuracy. By treating the measured quantity (measurand) as random variable, this uncertainty can be quantified. The different ways of describing uncertainty with statistical methods are used for the definition of uncertainty-related data quality basic measures.

The statistical methods used for the definition of uncertainty-related data quality measures are based on certain assumptions:

- uncertainties are homogeneous for all observed values;
- the observed values are not correlated;
- the observed values have normal distribution.

C.3.2 One-dimensional random variable, Z

For a continuous measurand (i.e. the value domain of the measured quantities is the real numbers), it is impossible to give the probability of a single value to be the true value. But it is possible to give the probability for the true value to be within a certain interval. This interval is called the confidence interval. It is given by the probability P of the true value being between the lower and the upper limit. This probability P is also called the significance level.

$$P(\text{lower limit} \leqslant \text{true value} \leqslant \text{upper limit}) = \text{P}$$

If the standard deviation σ is known, the limits are given by the quantiles u of the normal (Gaussian) distribution $P(z_t - u \cdot \sigma \leqslant \text{true value} \leqslant z_t + u \cdot \sigma) = \text{P}$.

Table C.2 — Relation between the quantiles of the normal distribution and the significance level

Probability P	Quantile	Data quality basic measure	Name	Data quality value type
P = 68,3 %	$u_{68,3\%} = 1$	$u_{68,3\%} \cdot \sigma_Z$	LE68.3	measure
P = 50 %	$u_{50\%} = 0{,}6745$	$u_{50\%} \cdot \sigma_Z$	LE50	measure
P = 90 %	$u_{90\%} = 1{,}645$	$u_{90\%} \cdot \sigma_Z$	LE90	measure
P = 95 %	$u_{95\%} = 1{,}960$	$u_{95\%} \cdot \sigma_Z$	LE95	measure
P = 99 %	$u_{99\%} = 2{,}576$	$u_{99\%} \cdot \sigma_Z$	LE99	measure
P = 99,8 %	$u_{99,8\%} = 3$	$u_{99,8\%} \cdot \sigma_Z$	LE99.8	measure

If the standard deviation σ is unknown, but the one-dimensional random variable Z is measured redundantly by N independent observations, it is possible to estimate the standard deviation from the observations.

z_{mi} represents the i^{th} measurement for the value. If the true value z_t for Z is known, the standard deviation can be estimated by

$$s_Z = \sqrt{\frac{1}{r}\sum_{i=1}^{N}(z_{mi} - z_t)^2}$$

with redundancy r being the number of observations $r = N$. If the true value is unknown, it may be estimated as the arithmetic mean of the observations $z_t = \sum_{i=1}^{N} z_{mi}$.

The standard deviation may then be estimated using the same formula, with $r = N - 1$.

If the standard deviation is estimated by redundant measurements, the confidence interval can be derived from the Student's t-distribution with parameter r:

$$P\left(-t \cdot s_z \leqslant Z - z_t \leqslant t \cdot s_z\right) = \mathrm{P} \text{ with } (Z - z_t)/s_z \sim t(r)$$

Table C.3 — Relation between the quantiles of the Student's t-distribution and the significance level for different redundancies r

Probability P	Quantile for r = 10	Quantile for r = 5	Quantile for r = 4	Quantile for r = 3	Quantile for r = 2	Quantile for r = 1
P = 50 %	t = 1,221	t = 1,301	t = 1,344	t = 1,423	t = 1,604	t = 2,414
P = 68,3 %	t = 1,524	t = 1,657	t = 1,731	t = 1,868	t = 2,203	t = 3,933
P = 90 %	t = 2,228	t = 2,571	t = 2,776	t = 3,182	t = 4,303	t = 12,706
P = 95 %	t = 2,634	t = 3,163	t = 3,495	t = 4,177	t = 6,205	t = 25,452
P = 99 %	t = 3,581	t = 4,773	t = 5,598	t = 7,453	t = 14,089	t = 127,321
P = 99,8 %	t = 4,587	t = 6,869	t = 8,610	t = 12,924	t = 31,599	t = 636,619

Table C.4 — Data quality basic measures for different probabilities P of a one-dimensional quantity, where the standard deviation is estimated from redundant measurements

Probability P	Data quality basic measure	Name	Data quality value type
P = 50,0 %	$t_{50\%}(r) \cdot s_Z$	LE50(r)	measure
P = 68,3 %	$t_{68,3\%}(r) \cdot s_Z$	LE68.3(r)	measure
P = 90,0 %	$t_{90\%}(r) \cdot s_Z$	LE90(r)	measure
P = 95,0 %	$t_{95\%}(r) \cdot s_Z$	LE95(r)	measure
P = 99,0 %	$t_{99\%}(r) \cdot s_Z$	LE99(r)	measure
P = 99,8 %	$t_{99,8\%}(r) \cdot s_Z$	LE99.8(r)	measure
NOTE The values of t for a number of redundancies r can be obtained from Table C.3.			

The data quality basic measures for the uncertainty of one-dimensional quantities are given in Tables C.2 and C.4. They both aim to measure the uncertainty by giving the upper and lower limit of a confidence interval. The difference is in how the standard deviation is obtained. If it is known *a priori*, then Table C.2 is relevant. If the standard deviation is estimated from redundant measurements, then Table C.4 in conjunction with Table C.3 is relevant.

C.3.3 Two-dimensional random variable X and Y

The case of the one-dimensional random variable Z can be expanded to two dimensions where the measurand is always observed by two values. The measurand is given by the tuple X, Y. This has the same assumptions as in the case of the one-dimensional random variable.

The observations are x_{mi} and y_{mi}. The equivalence of the confidence interval in one dimension is the confidence area, which is usually described as a circle around the best estimation for the true value. The probability for the true value to lie in this area is calculated by area integration over the two-dimensional density function of the normal distribution. A circular area is characterized by its radius. This radius, R, is used as measure for the accuracy of two-dimensional random variables:

$$P(R,\sigma_X,\sigma_Y)=\frac{1}{2\pi\sigma_X\sigma_Y}\iint\limits_{(x-x_t)^2+(y-y_t)^2=R^2} \mathrm{e}^{-\frac{1}{2}\left(\frac{(x-x_t)^2}{\sigma_X^2}+\frac{(y-y_t)^2}{\sigma_Y^2}\right)}\mathrm{d}x\mathrm{d}y$$

For some particular probabilities, the radius can be calculated depending on the standard deviations σ_x and σ_y.

Table C.5 — Relationship between the probability P and the corresponding radius of the circular area

Probability P	Data quality basic measure	Name	Data quality value type
P = 39,4 %	$\frac{1}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE39.4	measure
P = 50 %	$\frac{1{,}1774}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE50	measure
P = 90 %	$\frac{2{,}146}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE90	measure
P = 95 %	$\frac{2{,}4477}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE95	measure
P = 99,8 %	$\frac{3{,}5}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE99.8	measure

C.3.4 Three-dimensional random variable X, Y, Z

The case of the one-dimensional random variable Z can be expanded to three dimensions where the measurand is always observed by three values. The measurand is given by the tuple X, Y, Z. They underlay the same assumptions as in the case of the one-dimensional random variable.

The observations are x_{mi}, y_{mi} and z_{mi}. The equivalence of the confidence interval in one dimension is the confidence volume, which is usually described as a sphere around the best estimation for the true value. The probability for the true value to lie in this volume is calculated by volume integration over the three-dimensional density function of the normal distribution. A spherical volume is characterized by its radius. This radius is used as measure for the accuracy of three-dimensional random variables.

Table C.6 — Relationship between the probability P and the corresponding radius of the spherical volume

Probability P	Data quality basic measure	Name	Data quality value type
P = 50 %	$0{,}51 \cdot (\sigma_x + \sigma_y + \sigma_z)$	spherical error probable (SEP)	measure
P = 61 %	$\sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$	mean radial spherical error (MRSE)	measure
P = 90 %	$0{,}833 \cdot (\sigma_x + \sigma_y + \sigma_z)$	90 % spherical accuracy standard	measure
P = 99 %	$1{,}122 \cdot (\sigma_x + \sigma_y + \sigma_z)$	99 % spherical accuracy standard	measure

Annex D
(normative)

List of data quality measures

D.1 Completeness

D.1.1 Overview

This annex defines data quality measures. In order to achieve well defined and comparable quality information, it is strongly recommended to carry out the evaluation and reporting of data quality using these data quality measures. Due to the nature of quality and geospatial information, this list cannot be complete. Therefore, there may be cases where the user of this Technical Specification has to come up with user-defined data quality measures. In cases where user-defined data quality measures are related to error counts or to uncertainty, they shall be defined using the data quality basic measures as provided in Annex C. In any case, a data quality measure shall be defined using the structure as given in Annex B.

D.1.2 Commission

The data quality measures for the data quality subelement commission are provided in Tables D.1 to D.4.

Table D.1 — Excess item

Line	Component	Description
1	Name	excess item
2	Alias	–
3	Data quality element	completeness
4	Data quality subelement	commission
5	Data quality basic measure	error indicator
6	Definition	indication that an item is incorrectly present in the data
7	Description	–
8	Parameter	–
9	Data quality value type	Boolean (true indicates that the item is in excess)
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	1

Table D.2 — Number of excess items

Line	Component	Description
1	Name	number of excess items
2	Alias	–
3	Data quality element	completeness
4	Data quality subelement	commission
5	Data quality basic measure	error count
6	Definition	number of items within the dataset that should not have been in the dataset
7	Description	–
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	2

Table D.3 — Rate of excess items

Line	Component	Description
1	Name	Rate of excess items
2	Alias	–
3	Data quality element	completeness
4	Data quality subelement	commission
5	Data quality basic measure	error rate
6	Definition	number of excess items in the dataset in relation to the number of items that should have been present
7	Description	–
8	Parameter	–
9	Data quality value type	real, percentage, ratio
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	3

Table D.4 — Number of duplicate feature instances

Line	Component	Description
1	Name	number of duplicate feature instances
2	Alias	–
3	Data quality element	completeness
4	Data quality subelement	commission
5	Data quality basic measure	error count
6	Definition	total number of exact duplications of feature instances within the data
7	Description	count of all items in the data that are incorrectly extracted with duplicate geometries
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	Features with identical attribution and identical coordinates: ▲ two (or more) points collected on top of each other; — two (or more) curves collected on top of each other; □ two (or more) surfaces collected on top of each other.
13	Identifier	4

D.1.3 Omission

The data quality measures for the data quality subelement omission are provided in Tables D.5 to D.7.

Table D.5 — Missing item

Line	Component	Description
1	Name	missing item
2	Alias	–
3	Data quality element	completeness
4	Data quality subelement	omission
5	Data quality basic measure	error indicator
6	Definition	indicator that shows that a specific item is missing in the data
7	Description	–
8	Parameter	–
9	Data quality value type	Boolean (true indicates that an item is missing)
10	Data quality value structure	–
11	Source reference	–
12	Example	A product specification requires all towers higher than 300 m to be captured. The data quality measure “missing item” allows a data quality evaluator or a data user to report that a specific item, in this case a feature of type “tower” (name depends on the application schema), is missing. Data quality scope: all towers with height > 300 m Example result of a completeness evaluation of a particular data set: missing item = true for • tower.name = “Eiffel Tower, Paris, France” • tower.name = “Beijing Tower, Beijing, China”
13	Identifier	5

Table D.6 — Number of missing items

Line	Component	Description
1	Name	number of missing items
2	Alias	–
3	Data quality element	completeness
4	Data quality subelement	omission
5	Data quality basic measure	error count
6	Definition	count of all items that should have been in the dataset and are missing
7	Description	–
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	6

Table D.7 — Rate of missing items

Line	Component	Description
1	Name	rate of missing items
2	Alias	–
3	Data quality element	completeness
4	Data quality subelement	omission
5	Data quality basic measure	error rate
6	Definition	number of missing items in the dataset in relation to the number of items that should have been present
7	Description	–
8	Parameter	–
9	Data quality value type	real, percentage, ratio
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	7

D.2 Logical consistency

D.2.1 Conceptual consistency

The data quality measures for the data quality subelement conceptual consistency are provided in Tables D.8 to D.13.

Table D.8 — Conceptual schema noncompliance

Line	Component	Description
1	Name	conceptual schema noncompliance
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	conceptual consistency
5	Data quality basic measure	error indicator
6	Definition	indication that an item is not compliant to the rules of the relevant conceptual schema
7	Description	–
8	Parameter	–
9	Data quality value type	Boolean (true indicates that an item is not compliant with the rules of the conceptual schema)
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	8

Table D.9 — Conceptual schema compliance

Line	Component	Description
1	Name	conceptual schema compliance
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	conceptual consistency
5	Data quality basic measure	correctness indicator
6	Definition	indication that an item complies with the rules of the relevant conceptual schema
7	Description	–
8	Parameter	–
9	Data quality value type	Boolean (true indicates that an item is in compliance with the rules of the conceptual schema)
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	9

Table D.10 — Number of items not compliant with the rules of the conceptual schema

Line	Component	Description
1	Name	number of items not compliant with the rules of the conceptual schema
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	conceptual consistency
5	Data quality basic measure	error count
6	Definition	count of all items in the dataset that are not compliant with the rules of the conceptual schema
7	Description	If the conceptual schema explicitly or implicitly describes rules, these rules shall be followed. Violations against such rules can be, for example, invalid placement of features within a defined tolerance, duplication of features and invalid overlap of features.
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–

Table D.10 (*continued*)

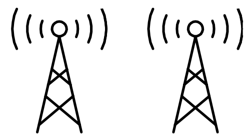
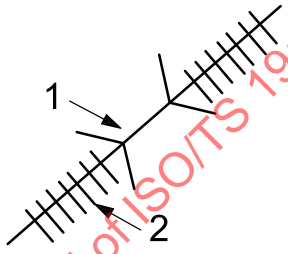
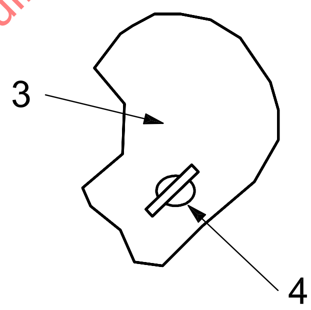
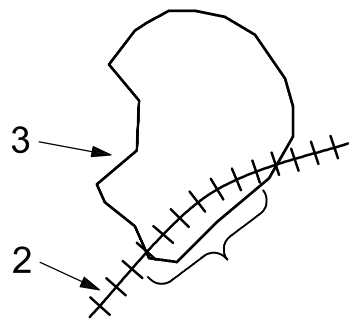
Line	Component	Description
12	Example	**Example 1**: Towers with identical attribution and within search tolerance (search tolerance = 10 m)  **Example 2**: Bridge has invalid Transportation. Use Category of Road  **Example 3**: Invalid placement of Airport inside a Lake  **Example 4**: Invalid overlap of area feature Lake within line feature Railroad  **Key** 1 Bridge 2 Railroad 3 Lake 4 Airport
13	Identifier	10

Table D.11 — Number of invalid overlaps of surfaces

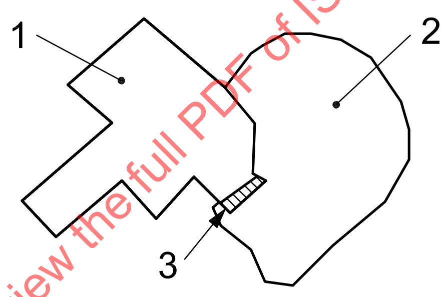
Line	Component	Description
1	Name	number of invalid overlaps of surfaces
2	Alias	overlapping surfaces
3	Data quality element	logical consistency
4	Data quality subelement	conceptual consistency
5	Data quality basic measure	error count
6	Definition	total number of erroneous overlaps within the data
7	Description	Which surfaces may overlap and which shall not is application dependent. Not all overlapping surfaces are necessarily erroneous. When reporting this data quality measure, the types of feature classes corresponding to the illegal overlapping surfaces shall be reported as well.
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	 **Key** 1 Surface 1 2 Surface 2 3 Overlapping Area
13	Identifier	11

Table D.12 — Noncompliance rate with respect to the rules of the conceptual schema

Line	Component	Description
1	Name	noncompliance rate with respect to the rules of the conceptual schema
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	conceptual consistency
5	Data quality basic measure	error rate
6	Definition	number of items in the dataset that are not compliant with the rules of the conceptual schema in relation to the total number of these items supposed to be in the dataset
7	Description	–
8	Parameter	–
9	Data quality value type	real, percentage, ratio
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	12

Table D.13 — Compliance rate with the rules of the conceptual schema

Line	Component	Description
1	Name	compliance rate with the rules of the conceptual schema
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	conceptual consistency
5	Data quality basic measure	correct items rate
6	Definition	number of items in the dataset in compliance with the rules of the conceptual schema in relation to the total number of items
7	Description	–
8	Parameter	–
9	Data quality value type	real, percentage, ratio
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	13

D.2.2 Domain consistency

The data quality measures for the data quality subelement domain consistency are provided in Tables D.14 to D.18.

Table D.14 — Value domain nonconformance

Line	Component	Description
1	Name	value domain nonconformance
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	domain consistency
5	Data quality basic measure	error indicator
6	Definition	indication of if an item is not in conformance with its value domain
7	Description	–
8	Parameter	–
9	Data quality value type	Boolean (true indicates that an item is not in conformance with its value domain)
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	14

Table D.15 — Value domain conformance

Line	Component	Description
1	Name	value domain conformance
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	domain consistency
5	Data quality basic measure	correctness indicator
6	Definition	indication of if an item is conforming to its value domain
7	Description	–
8	Parameter	–
9	Data quality value type	Boolean (true indicates that an item is conforming to its value domain)
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	15

Table D.16 — Number of items not in conformance with their value domain

Line	Component	Description
1	Name	number of items not in conformance with their value domain
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	domain consistency
5	Data quality basic measure	error count
6	Definition	count of all items in the dataset that are not in conformance with their value domain
7	Description	–
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	16

Table D.17 — Value domain conformance rate

Line	Component	Description
1	Name	value domain conformance rate
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	domain consistency
5	Data quality basic measure	correct items rate
6	Definition	number of items in the dataset that are in conformance with their value domain in relation to the total number of items in the dataset
7	Description	–
8	Parameter	–
9	Data quality value type	real, percentage, ratio
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	17

Table D.18 — Value domain nonconformance rate

Line	Component	Description
1	Name	value domain nonconformance rate
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	domain consistency
5	Data quality basic measure	error ratio
6	Definition	number of items in the dataset that are not in conformance with their value domain in relation to the total number of items
7	Description	–
8	Parameter	–
9	Data quality value type	real, percentage, ratio
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	18

D.2.3 Format consistency

The data quality measures for the data quality subelement format consistency are provided in Tables D.19 and D.20.

Table D.19 — Physical structure conflicts

Line	Component	Description
1	Name	physical structure conflicts
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	format consistency
5	Data quality basic measure	error count
6	Definition	count of all items in the dataset that are stored in conflict with the physical structure of the dataset
7	Description	–
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	19

Table D.20 — Physical structure conflict rate

Line	Component	Description
1	Name	physical structure conflict rate
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	format consistency
5	Data quality basic measure	error rate
6	Definition	number of items in the dataset that are stored in conflict with the physical structure of the dataset divided by the total number of items
7	Description	–
8	Parameter	–
9	Data quality value type	real, percentage, ratio
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	20

D.2.4 Topological consistency

The data quality measures in Tables D.21 to D.27 are designed to test the topological consistency of geometric representations of features. They will not serve as measures of the consistency of explicit descriptions of topology using the topological objects specified in ISO 19107.

Table D.21 — Number of faulty point-curve connections

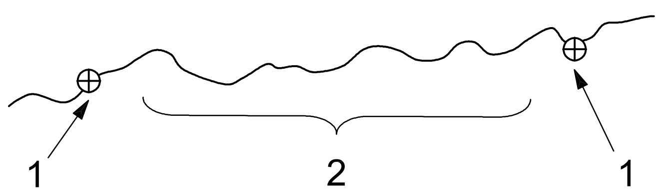
Line	Component	Description
1	Name	number of faulty point-curve connections
2	Alias	extraneous nodes
3	Data quality element	logical consistency
4	Data quality subelement	topological consistency
5	Data quality basic measure	error count
6	Definition	number of faulty point-curve connections in the dataset
7	Description	A point-curve connection exists where different curves touch. These curves have an intrinsic topological relationship that shall reflect the true constellation. If the point-curve connection contradicts the universe of discourse, the point-curve connection is faulty with respect to this data quality measure. The data quality measure counts the number of errors of this kind.
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	**Example 1**: Two-point curve connections exist where only one should be present  **Key** 1 Junction of two roads should be at a "+" intersection **Example 2**: System automatically places point-curve based on vertices limitation built into software code where no spatial justification for point-curve exists  **Key** 1 Link-node 2 500 vertices limit
13	Identifier	21

Table D.22 — Rate of faulty point-curve connections

Line	Component	Description
1	Name	rate of faulty point-curve connections
2	Alias	–
3	Data quality element	logical consistency
4	Data quality subelement	topological consistency
5	Data quality basic measure	error rate
6	Definition	number of faulty link node connections in relation to the number of supposed link node connections
7	Description	A point-curve connection exists where different curves touch. These curves have an intrinsic topological relationship that shall reflect the true constellation. If the point-curve connection contradicts the universe of discourse, the point-curve connection is faulty with respect to this data quality measure. This data quality measure gives the erroneous point-curve connections in relation to the total number of point-curve connections.
8	Parameter	–
9	Data quality value type	real, percentage, ratio
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	22

Table D.23 — Number of missing connections due to undershoots

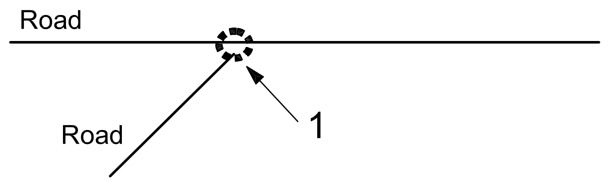
Line	Component	Description
1	Name	number of missing connections due to undershoots
2	Alias	undershoots
3	Data quality element	logical consistency
4	Data quality subelement	topological consistency
5	Data quality basic measure	error count
6	Definition	count of items in the dataset, within the parameter tolerance, that are mismatched due to undershoots
7	Description	–
8	Parameter	search distance from the end of a dangling line
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	 **Key** 1 Search tolerance = 3 m
13	Identifier	23

Table D.24 — Number of missing connections due to overshoots

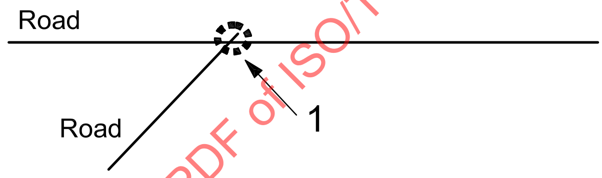
Line	Component	Description
1	Name	number of missing connections due to overshoots
2	Alias	overshoots
3	Data quality element	logical consistency
4	Data quality subelement	topological consistency
5	Data quality basic measure	error count
6	Definition	count of items in the dataset, within the parameter tolerance, that are mismatched due to overshoots
7	Description	–
8	Parameter	search tolerance of minimum allowable length in the dataset
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	 **Key** 1 Search tolerance = 3 m
13	Identifier	24

Table D.25 — Number of invalid slivers

Line	Component	Description
1	Name	number of invalid slivers
2	Alias	slivers
3	Data quality element	logical consistency
4	Data quality subelement	topological consistency
5	Data quality basic measure	error count
6	Definition	count of all items in the dataset that are invalid sliver surfaces
7	Description	A sliver is an unintended area that occurs when adjacent surfaces are not digitized properly. The borders of the adjacent surfaces may unintentionally gap or overlap by small amounts to cause a topological error.
8	Parameter	This data quality measure has 2 parameters: • maximum sliver area size • thickness quotient The thickness quotient shall be a real number between 0 and 1. This quotient is determined by the following formula: T is the thickness quotient $T = 4\pi\,[\text{area}]/[\text{perimeter}]^2$ $T = 1$ value corresponds to a circle that has the largest area/perimeter^2 value. $T = 0$ value corresponds to a line that has the smallest area/perimeter^2 value. The thickness quotient is independent of the size of the surface, and the closer the value is to 0, the thinner the selected sliver surfaces shall be. The maximum area determines the upper limit of a sliver. This is to prevent surfaces with sinuous perimeters and large areas from being mistaken as slivers.
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	Environmental Systems Research Institute, Inc. (ESRI) GIS Data ReViewer 4.2 User Guide

Table D.25 (*continued*)

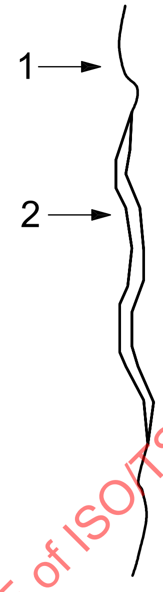
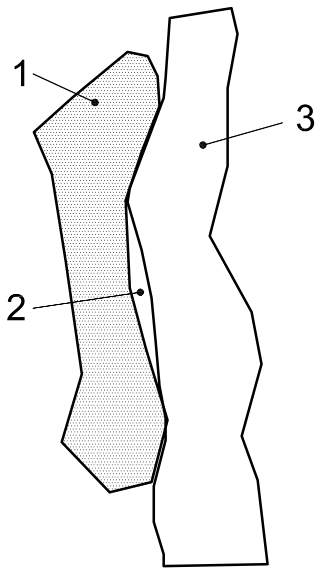
Line	Component	Description
12	Example	 **Key** 1 Single line drain 2 Double line drain a) Maximum area parameter prevents correct double line drain portrayal from being flagged as an error  **Key** 1 Sand 2 Sliver 3 Double line drain b) Sliver is less than the maximum parameter and is flagged for evaluation of possible error
13	Identifier	25

Table D.26 — Number of invalid self-intersect errors

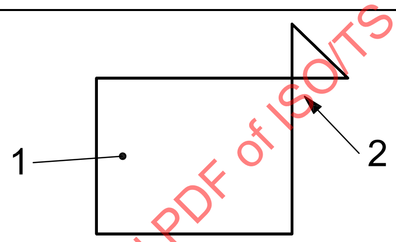
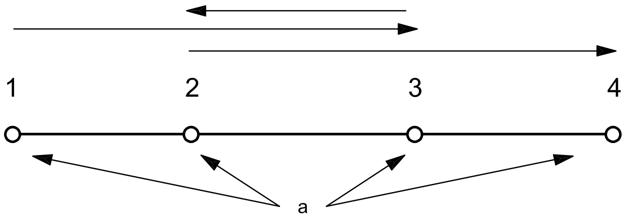
Line	Component	Description
1	Name	number of invalid self-intersect errors
2	Alias	loops
3	Data quality element	logical consistency
4	Data quality subelement	topological consistency
5	Data quality basic measure	error count
6	Definition	count of all items in the data that illegally intersect with themselves
7	Description	–
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	 **Key** 1 Building 1 2 Illegal intersection (loop)
13	Identifier	26

Table D.27 — Number of invalid self-overlap errors

Line	Component	Description
1	Name	number of invalid self-overlap errors
2	Alias	kickbacks
3	Data quality element	logical consistency
4	Data quality subelement	topological consistency
5	Data quality basic measure	error count
6	Definition	count of all items in the data that illegally self overlap
7	Description	–
8	Parameter	–
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	 [a] Vertices.
13	Identifier	27

D.3 Positional accuracy

D.3.1 Absolute or external accuracy

D.3.1.1 General measures for positional uncertainties

The data quality measures for positional uncertainty in general of the data quality subelement absolute or external accuracy are provided in Tables D.28 to D.32.

Table D.28 — Mean value of positional uncertainties

Line	Component	Description
1	Name	mean value of positional uncertainties (1D, 2D and 3D)
2	Alias	–
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	mean value of the positional uncertainties for a set of positions where the positional uncertainties are defined as the distance between a measured position and what is considered as the corresponding true position
7	Description	For a number of points (N), the measured positions are given as x_{mi}, y_{mi} and z_{mi} coordinates depending on the dimension in which the position of the point is measured. A corresponding set of coordinates, x_{ti}, y_{ti} and z_{ti}, are considered to represent the true positions. The errors are calculated as 1D: $e_i = \lvert x_{mi} - x_{ti} \rvert$ 2D: $e_i = \sqrt{(x_{mi} - x_{ti})^2 + (y_{mi} - y_{ti})^2}$ 3D: $e_i = \sqrt{(x_{mi} - x_{ti})^2 + (y_{mi} - y_{ti})^2 + (z_{mi} - z_{ti})^2}$ The mean positional uncertainties of the horizontal absolute or external positions are then calculated as $\bar{e} = \frac{1}{N}\sum_{i=1}^{N} e_i$ A criterion for the establishing of correspondence should also be stated (e.g. allowing for correspondence to the closest position, correspondence on vertices or along lines). The criterion/criteria for finding the corresponding points shall be reported with the data quality evaluation result. This data quality measure is different from the standard deviation.
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	28

Table D.29 — Mean value of positional uncertainties excluding outliers

Line	Component	Description
1	Name	mean value of positional uncertainties excluding outliers (2D)
2	Alias	–
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	for a set of points where the distance does not exceed a defined threshold, the arithmetical average of distances between their measured positions and what is considered as the corresponding true positions
7	Description	For a number of points (N), the measured positions are given as x_{mi}, y_{mi} and z_{mi} coordinates depending on the dimension in which the position of the point is measured. A corresponding set of coordinates, x_{ti}, y_{ti} and z_{ti}, are considered to represent the true positions. All positional uncertainties above a defined threshold e_{max} are then removed from the set. The positional uncertainties are calculated as $e'_i = \begin{cases} e_i, \text{ if } e_i \leqslant e_{\text{max}} \\ 0, \text{ if } e_i > e_{\text{max}} \end{cases}$ The calculation of e_i is given by the data quality measure "mean value of positional uncertainties" in one, two and three dimensions. For the remaining number of errors (N_{R}), the mean of the horizontal absolute positions is calculated as $\bar{e}_{\text{excluding outliers}} = \frac{1}{N_{\text{R}}} \sum_{i=1}^{N} e'_i$ A criterion for the establishing of correspondence should also be stated (e.g. allowing for correspondence to the closest position, correspondence on vertices or along lines). The criteria for finding the corresponding points shall be reported with the data quality evaluation result.
8	Parameter	e_{max} is the threshold for accepted positional uncertainties
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	29

Table D.30 — Number of positional uncertainties above a given threshold

Line	Component	Description
1	Name	number of positional uncertainties above a given threshold
2	Alias	–
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	error count
6	Definition	number of positional uncertainties above a given threshold for a set of positions The errors are defined as the distance between a measured position and what is considered as the corresponding true position.
7	Description	For a number of points (N), the measured positions are given as x_{mi}, y_{mi} and z_{mi} coordinates depending on the dimension in which the position of the point is measured. A corresponding set of coordinates, x_{ti}, y_{ti} and z_{ti}, are considered to represent the true positions. The calculation of e_i is given by the data quality measure “mean value of positional uncertainties” in one, two and three dimensions. All positional uncertainties above a defined threshold e_{max} $(e_i > e_{\mathrm{max}})$ are then counted as error. A criterion for the establishing of correspondence should also be stated (e.g. allowing for correspondence to the closest position, correspondence on vertices or along lines). The criterion/criteria for finding the corresponding points shall be reported with the data quality evaluation result.
8	Parameter	e_{max} is the threshold for accepted positional uncertainties
9	Data quality value type	integer
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	30

Table D.31 — Rate of positional errors above a given threshold

Line	Component	Description
1	Name	rate of positional uncertainties above a given threshold
2	Alias	–
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	number of positional uncertainties above a given threshold for a set of positions in relation to the total number of measured positions The errors are defined as the distance between a measured position and what is considered as the corresponding true position
7	Description	For a number of points (N), the measured positions are given as x_{mi}, y_{mi} and z_{mi} coordinates depending on the dimension in which the position of the point is measured. A corresponding set of coordinates, x_{ti}, y_{ti} and z_{ti}, are considered to represent the true positions. The calculation of e_i is given by the data quality measure "mean value of positional uncertainties" in one, two and three dimensions. All positional uncertainties above a defined threshold e_{max} ($e_i > e_{\text{max}}$) are then counted as error. The number of errors is set in relation to the total number of measured points. A criterion for the establishing of correspondence should also be stated (e.g. allowing for correspondence to the closest position, correspondence on vertices or along lines). The criterion/criteria for finding the corresponding points shall be reported with the data quality evaluation result.
8	Parameter	e_{max} is the threshold above which the positional uncertainties are counted
9	Data quality value type	real, percentage, ratio
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	31

Table D.32 — Covariance matrix

Line	Component	Description
1	Name	covariance matrix
2	Alias	variance-covariance matrix
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	symmetrical square matrix with variances of point coordinates on the main diagonal and covariances between these coordinates as off-diagonal elements
7	Description	The covariance matrix generalizes the concept of variance from one to n dimensions, i.e. from scalar-valued random variables to vector-valued random variables (tuples of scalar random variables). (1) 1D coordinates (e.g. height data) Vector-valued random variable: $x = \begin{bmatrix} x_1 \\ \vdots \\ x_{1n} \end{bmatrix}$ Its covariance matrix: $\Sigma_{xx} = \begin{bmatrix} \sigma_{x1}^2 & \cdots & \sigma_{x1xn} \\ \vdots & \ddots & \vdots \\ \sigma_{xnx1} & \cdots & \sigma_{xn}^2 \end{bmatrix}$, with $\sigma_{x1xn} = \sigma_{xnx1}$ σ_{x1}^2 denotes the variance of the element x_1, its square root gives the standard deviation of this element $\sigma_{x1} = \sqrt{\sigma_{x1}^2}$. The correlation between 2 elements can be calculated by $\rho_{xixj} = \frac{\sigma_{xixj}}{\sigma_{xi}\sigma_{xj}}$. If the coordinates are uncorrelated, the off-diagonal elements are of value 0. (2) 2D coordinates Vector-valued random variable: $x = \begin{bmatrix} x_1 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}$ Its covariance matrix: $\Sigma_{xx} = \begin{bmatrix} \sigma_{x1}^2 & \sigma_{x1y1} & \cdots & \sigma_{x1yn} \\ \sigma_{y1x1} & \sigma_{y1}^2 & \cdots & \sigma_{y1yn} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{ynx1} & \sigma_{yny1} & \cdots & \sigma_{yn}^2 \end{bmatrix}$,

Table D.32 (*continued*)

Line	Component	Description
		(3) 3D coordinates Vector-valued random variable: $x = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ \vdots \\ y_n \\ z_n \end{bmatrix}$ Its covariance matrix: $\Sigma_{xx} = \begin{bmatrix} \sigma_{x1}^2 & \sigma_{x1y1} & \sigma_{x1z1} & \cdots & \sigma_{x1yn} & \sigma_{x1zn} \\ \sigma_{x1y1} & \sigma_{y1}^2 & \sigma_{y1z1} & \cdots & \sigma_{y1yn} & \sigma_{y1zn} \\ \sigma_{x1z1} & \sigma_{y1z1} & \sigma_{z1}^2 & \cdots & \sigma_{z1yn} & \sigma_{z1zn} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{x1yn} & \sigma_{y1yn} & \sigma_{z1yn} & \cdots & \sigma_{yn}^2 & \sigma_{ynzn} \\ \sigma_{x1zn} & \sigma_{y1zn} & \sigma_{z1zn} & \cdots & \sigma_{ynzn} & \sigma_{zn}^2 \end{bmatrix},$ (4) arbitrary observables Vector-valued random variable: $x = \begin{bmatrix} a \\ b \\ \vdots \\ z \end{bmatrix}$ Its covariance matrix: $\Sigma_{xx} = \begin{bmatrix} \sigma_a^2 & \sigma_{ba} & \cdots & \sigma_{za} \\ \sigma_{ab} = \sigma_{ba} & \sigma_b^2 & \cdots & \sigma_{zb} \\ & \vdots & \ddots & \vdots \\ \sigma_{az} = \sigma_{za} & \sigma_{bz} = \sigma_{zb} & \cdots & \sigma_z^2 \end{bmatrix}$
8	Parameter	–
9	Data quality value type	measures
10	Data quality value structure	matrix
11	Source reference	–
12	Example	–
13	Identifier	32

D.3.1.2 Vertical positional uncertainties

Height measurements are position observations in one dimension. The measurand height may therefore be treated as a one-dimensional random variable. The data quality measures for positional uncertainties are therefore based on the data quality basic measure "one-dimensional random variable".

The data quality measures for vertical positional uncertainty of the data quality subelement absolute or external accuracy are provided in Tables D.33 to D.41.

Table D.33 — Linear error probable

Line	Component	Description
1	Name	linear error probable
2	Alias	LEP
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	LE50 or LE50I, depending on the evaluation procedure
6	Definition	half length of the interval defined by an upper and a lower limit, in which the true value lies with probability 50 %
7	Description	See C.3.2
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	33

Table D.34 — Standard linear error

Line	Component	Description
1	Name	standard linear error
2	Alias	SD
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	LE68.3 or LE68.3(r), depending on the evaluation procedure
6	Definition	half length of the interval defined by an upper and a lower limit, in which the true value lies with probability 68,3 %
7	Description	See C.3.2
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	34

Table D.35 — Linear map accuracy at 90 % significance level

Line	Component	Description
1	Name	linear map accuracy at 90 % significance level
2	Alias	LMAS 90 %
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	LE90 or LE90I, depending on the evaluation procedure
6	Definition	half length of the interval defined by an upper and a lower limit, in which the true value lies with probability 90 %
7	Description	See C.3.2
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	35

Table D.36 — Linear map accuracy at 95 % significance level

Line	Component	Description
1	Name	linear map accuracy at 95 % significance level
2	Alias	LMAS 95 %
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	LE95 or LE95I, depending on the evaluation procedure
6	Definition	half length of the interval defined by an upper and a lower limit, in which the true value lies with probability 95 %
7	Description	See C.3.2
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	36

Table D.37 — Linear map accuracy at 99 % significance level

Line	Component	Description
1	Name	linear map accuracy at 99 % significance level
2	Alias	LMAS 99 %
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	LE99 or LE99(r), depending on the evaluation procedure
6	Definition	half length of the interval defined by an upper and a lower limit, in which the true value lies with probability 99 %
7	Description	See C.3.2
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	37

Table D.38 — Near certainty linear error

Line	Component	Description
1	Name	near certainty linear error
2	Alias	–
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	LE99.8 or LE99.8(r), depending on the evaluation procedure
6	Definition	half length of the interval defined by an upper and a lower limit, in which the true value lies with probability 99,8 %
7	Description	See C.3.2
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	38

Table D.39 — Root mean square error

Line	Component	Description
1	Name	root mean square error
2	Alias	RMSE
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	standard deviation, where the true value is not estimated from the observations but known *a priori*
7	Description	The true value of an observable Z is known as z_{t}. From this, the estimator $\sigma_z = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Z_{mi} - z_{\mathrm{t}})^2}$ yields to the linear root mean square error RMSE = σ_z.
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	39

Table D.40 — Absolute linear error at 90 % significance level of biased vertical data (Alternative 1)

Line	Component	Description
1	Name	absolute linear error at 90 % significance level of biased vertical data (Alternative 1)
2	Alternative name	LMAS
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	absolute vertical accuracy of the data's coordinates, expressed in terms of linear error at 90 % probability given that a bias is present
7	Description	A comparison of the data (source) and the control (reference) is calculated in the following manner: 1. Calculate the absolute error in the vertical dimension at each point: $\delta V_i = \text{source}V_i - \text{reference}V_i$ for $i = 1 \ldots N$ 2. Calculate absolute value of the bias: $\left\|\overline{\delta V}\right\| = \left\|\frac{1}{N}\sum_{i=1}^{N}\delta V_i\right\|$ 3. Calculate the linear standard deviation of measured differences between the tested product and the reference source: $\sigma_\text{M} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\Delta V_i^2}$ 4. Calculate the standard linear standard deviation of errors in the reference source: σ_R 5. Calculate the linear standard deviation of errors in the tested product: $\sigma = \sqrt{\sigma_\text{M}^2 + \sigma_\text{R}^2}$ 6. Calculate the ratio of the absolute value of the mean error to the standard deviation: $\text{ratio} = \frac{\left\|\overline{\delta V}\right\|}{\sigma_V}$ 7. If $\text{ratio} > 1{,}4$, then $\text{LMAS} = \sigma_V \cdot \left[1{,}282 + \text{ratio}\right]$ 8. If $\text{ratio} \leqslant 1{,}4$, then calculate k based on the ratio of the vertical bias to the standard deviation of the heights using $\text{LMAS} = \sigma_V \cdot \left[1{,}6435 + 0{,}92 \times \text{ratio}^2 - 0{,}28 \times \text{ratio}^3\right]$
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	NATO STANAG 2215 IGEO (Reference [3])
12	Example	–
13	Measure identifier	40

Table D.41 — Absolute linear error at 90 % significance level of biased vertical data

Line	Component	Description
1	Name	absolute linear error at 90 % significance level of biased vertical data
2	Alternative name	ALE
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	absolute vertical accuracy of the data's coordinates, expressed in terms of linear error at 90 % probability given that a bias is present
7	Description	A comparison of the data (source) and the control (reference) is calculated in the following manner: 1. Calculate the absolute error in the vertical dimension at each point: $\delta V_i = \text{source}V_i - \text{reference}V_i$ for $i = 1 \ldots N$ 2. Calculate the mean vertical error: $\overline{\lvert\delta V\rvert} = \left\lvert \frac{1}{N}\sum_{i=1}^{N}\delta V_i \right\rvert$ 3. Calculate the standard deviation of the vertical errors: $\sigma_V = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\Delta V_i^2}$ 4. Calculate the ratio of the absolute value of the mean error to the standard deviation: $\text{ratio} = \overline{\lvert\delta V\rvert} / \sigma_V$ 5. If $\text{ratio} > 1{,}4$, then $k = 1{,}2815$
		6. If $\text{ratio} \leqslant 1{,}4$, then calculate k based on the ratio of the vertical bias to the standard deviation of the heights using a cubic polynomial fit through the tabular values as defined in the *Handbook of Tables for Probability and Statistics* (Reference [4]). $k = 1{,}643\,5 - (0{,}999\,556 \times \text{ratio}) + (0{,}923\,237 \times \text{ratio}^2) - (0{,}282\,533 \times \text{ratio}^3)$ 7. Compute LE90 for the source: $\text{LE90}_{\text{source}} = \overline{\lvert\delta V\rvert} + (k \times \sigma_V)$ 8. Compute absolute LE90: $\text{LE90}_{\text{abs}} = \sqrt{\text{LE90}_{\text{reference}}^2 + \text{LE90}_{\text{source}}^2}$
8	Parameter	Sample size: minimum of 30 points is normally used but may not always be possible depending on identifiable control points. For feature level attribution sample 10 % of the feature population.
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	1. *Mapping, Charting and Geodesy Accuracy* (Reference [5]) 2. *Handbook of Tables for Probability and Statistics* (Reference [4])
12	Example	–
13	Measure identifier	41

D.3.1.3 Horizontal positional uncertainties

Horizontal point locations are defined by a 2D coordinates. The uncertainty of any point location can be described using the data quality basic measures for 2D random variables as described in C.3.3. The data quality measures for horizontal positional uncertainty of the data quality subelement absolute or external accuracy are provided in Tables D.42 to D.51.

Table D.42 — Circular standard deviation

Line	Component	Description
1	Name	circular standard deviation
2	Alias	circular standard error, Helmert's point error, CSE
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	CE39.4
6	Definition	radius describing a circle, in which the true point location lies with the probability of 39,4 %
7	Description	See C.3.3
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	42

Table D.43 — Circular error probable

Line	Component	Description
1	Name	circular error probable
2	Alias	CEP
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	CE50
6	Definition	radius describing a circle, in which the true point location lies with the probability of 50 %
7	Description	See C.3.3
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	43

Table D.44 — Circular map accuracy standard

Line	Component	Description
1	Name	circular map accuracy standard
2	Alias	CMAS
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	CE90
6	Definition	radius describing a circle, in which the true point location lies with the probability of 90 %
7	Description	See C.3.3
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	44

Table D.45 — Circular error at 95 % significance level

Line	Component	Description
1	Name	circular error at 95 % significance level
2	Alias	navigation accuracy
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	CE95
6	Definition	radius describing a circle, in which the true point location lies with the probability of 95 %
7	Description	See C.3.3
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	45

Table D.46 — Circular near certainty error

Line	Component	Description
1	Name	circular near certainty error
2	Alias	CNCE
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	CE99.8
6	Definition	radius describing a circle, in which the true point location lies with the probability of 99,8 %.
7	Description	See C.3.3
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	46

Table D.47 — Root mean square error of planimetry

Line	Component	Description
1	Name	root mean square error of planimetry
2	Alias	RMSEP
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	radius of a circle around the given point, in which the true value lies with probability P
7	Description	The true values of the observed coordinates X and Y are known as x_t and y_t From this the estimator $\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[(x_{mi}-x_t)^2+(y_{mi}-y_t)^2\right]}$ yields to the linear root mean square error of planimetry RMSEP = σ
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	47

Table D.48 — Absolute circular error at 90 % significance level of biased data (Alternative 2)

Line	Component	Description
1	Name	absolute circular error at 90 % significance level of biased data (Alternative 2)
2	Alternative name	absolute horizontal accuracy measure at the 90 % significance level of biased data | CMAS
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	absolute horizontal accuracy of the data's coordinates, expressed in terms of circular error at 90 % probability given that a bias is present
7	Description	A comparison of the data (source) and the control (reference) is calculated in the following manner: 1. Calculate the absolute error in the horizontal dimension at each point and each coordinate X_i and Y_i: $\delta X_i = (\text{source}X_i - \text{reference}X_i)$ and $\delta Yi = (\text{source}Y_i - \text{reference}Y_i)$ for i = 1…N 2. Calculate the mean horizontal error of each coordinate: $\overline{\delta X} = \frac{1}{N}\sum_{1}^{N}\delta Xi$ and $\overline{\delta Y} = \frac{1}{N}\sum_{1}^{N}\delta Yi$ 3. Calculate the circular standard deviation of measured differences between the tested product and the reference source: $\sigma_{\text{CM}} = \sqrt{\frac{1}{2(N-1)}\left(\sum_{i=1}^{N}\left(\delta Xi - \overline{\delta X}\right)^2 + \sum_{i=1}^{N}\left(\delta Xi - \overline{\delta X}\right)^2\right)}$ 4. Calculate the circular standard deviation of errors in the reference source: σ_{CR} 5. Calculate the circular standard deviation of errors in the tested product: $\sigma_{\text{C}} = \sqrt{\sigma_{\text{CM}}^2 + \sigma_{\text{CR}}^2}$ 6. Compute absolute circular error at 90 % confidence level of biased data (CMAS): $\text{CMAS} = \sigma_{\text{C}} \cdot \left[1{,}294\,3 + \sqrt{\left(\frac{\overline{\delta X}^2 + \overline{\delta Y}^2}{\sigma_{\text{C}}}\right)} + 0{,}725\,4\right]$
8	Parameter	–
9	Data quality value type	measure
10	Data quality value structure	–
11	Source reference	NATO STANAG 2215 IGEO (Reference [3])
12	Example	–
13	Measure identifier	48

Table D.49 — Absolute circular error at 90 % significance level of biased data

Line	Component	Description
1	Name	absolute circular error at 90 % significance level of biased data
2	Alternative name	ACE
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	absolute horizontal accuracy of the data's coordinates, expressed in terms of circular error at 90% probability given that a bias is present
7	Description	A comparison of the data (source) and the control (reference) is calculated in the following manner: 1. Calculate the absolute error in the horizontal dimension at each point: $\Delta H_i = \sqrt{(\text{source}X_i - \text{reference}X_i)^2 + (\text{source}Y_i - \text{reference}Y_i)^2}$ for $i = 1 \ldots N$ 2. Calculate the mean horizontal error: $\mu_\text{H} = \left(\sum \Delta H_i\right) / N$ 3. Calculate the standard deviation of the horizontal errors: $\sigma_\text{H} = \sqrt{\sum (\Delta H_i - \mu_\text{H})^2 / (N-1)}$ 4. Calculate the ratio of the absolute value of the mean error to the standard deviation: $\text{ratio} = \lvert\mu_H\rvert / \sigma_H$ 5. If $\text{ratio} > 1{,}4$, then $k = 1{,}2815$ 6. If $\text{ratio} \leqslant 1{,}4$, then calculate k, the ratio of the mean to the standard deviation, using a cubic polynomial fit through the tabular values as defined in the CRC Handbook of Tables for Probability and Statistics $k = 1{,}643\,5 - (0{,}999\,556 \times \text{ratio}) + (0{,}923\,237 \times \text{ratio}^2) - (0{,}282\,533 \times \text{ratio}^3)$ 7. Compute CE90 for the source: $\text{CE90}_\text{source} = \lvert\mu_\text{H}\rvert + (k \times \sigma_\text{H})$ 8. Compute absolute CE90: $\text{CE90}_\text{abs} = \sqrt{\text{CE90}_\text{reference}{}^2 + \text{CE90}_\text{source}{}^2}$
8	Parameter	Sample size: minimum of 30 points is normally used but may not always be possible depending on identifiable control points. For feature level attribution sample 10 % of the feature population.
9	Data quality value type	Measure
10	Data quality value structure	–
11	Source reference	1. *Mapping, Charting and Geodesy Accuracy* (Reference [5]) 2. *Handbook of Tables for Probability and Statistics* (Reference [4])
12	Example	–
13	Measure identifier	49

Table D.50 — Uncertainty ellipse

Line	Component	Description
1	Name	uncertainty ellipse
2	Alias	standard point error ellipse
3	Data quality element	positional accuracy
4	Data quality subelement	absolute or external accuracy
5	Data quality basic measure	not applicable
6	Definition	2D ellipse with the two main axes indicating the direction and magnitude of the highest and the lowest uncertainty of a 2D point
7	Description	From a given covariance matrix (data quality measure Table D.32) of 2D point coordinates, the elements describing the uncertainty ellipse can be determined by its eigenvalues. For a single point k, the covariance matrix is given by $\Sigma^{k}_{xx} = \begin{bmatrix} \sigma^2_{xk} & \sigma_{xkyk} \\ \sigma_{ykxk} & \sigma^2_{yk} \end{bmatrix}$, with $\sigma_{xkyk} = \sigma_{ykxk}$ The direction α (bearing) of the major semi-axis of the uncertainty ellipse can be computed by $\varphi = \frac{1}{2}\arctan\frac{2\sigma_{xkyk}}{\sigma^2_{xk} - \sigma^2_{yk}}$ and $a = \sqrt{\frac{1}{2}\left(\sigma^2_{xk} + \sigma^2_{yk} + \sqrt{\left(\sigma^2_{xk} - \sigma^2_{yk}\right)^2 + 4\sigma^2_{xkyk}}\right)}$ $b = \sqrt{\frac{1}{2}\left(\sigma^2_{xk} + \sigma^2_{yk} - \sqrt{\left(\sigma^2_{xk} - \sigma^2_{yk}\right)^2 + 4\sigma^2_{xkyk}}\right)}$
8	Parameter	–
9	Data quality value type	measures
10	Data quality value structure	list (a, b, φ)
11	Source reference	–
12	Example	–
13	Identifier	50

