INTERNATIONAL STANDARD

ISO 19113

First edition
2002-12-01

Geographic information — Quality principles

Information géographique — Principes qualité

Reference number
ISO 19113:2002(E)

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

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 ... 2

5 Principles for describing the quality of geographic data ... 4
5.1 Components of data quality description ... 4
5.2 Data quality elements and data quality subelements ... 5
5.3 Data quality overview elements ... 7

6 Identifying the quality of geographic information ... 8
6.1 Identifying quantitative quality information ... 8
6.2 Identifying non-quantitative quality information ... 10

7 Reporting quality information ... 10
7.1 Reporting quantitative quality information ... 10
7.2 Reporting non-quantitative quality information ... 10

Annex A (normative) **Abstract test suite** ... 11

Annex B (informative) **Data quality concepts and their use** ... 14

Annex C (informative) **Data quality elements, data quality subelements and data quality overview elements** ... 19

Foreword

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

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

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

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

ISO 19113 was prepared by Technical Committee ISO/TC 211, *Geographic information/Geomatics*.

Introduction

Geographic datasets are increasingly being shared, interchanged and used for purposes other than their producers' intended ones. Information about the quality of available geographic datasets is vital to the process of selecting a dataset in that the value of data is directly related to its quality. Data users confront situations requiring different levels of data quality. Extremely accurate data is required by some data users for certain needs and less accurate data are sufficient for other needs. Information about the quality of geographic data is becoming a decisive factor for its utilization as technological advances allow the collection and use of geographic datasets whose quality can exceed that which is needed and requested by data users.

The purpose of describing the quality of geographic data is to facilitate the selection of the geographic dataset best suited to application needs or requirements. Complete descriptions of the quality of a dataset will encourage the sharing, interchange and use of appropriate geographic datasets. A geographic dataset can be viewed as a commodity or product. Information on the quality of geographic data allows a data producer or vendor to validate how well a dataset meets the criteria set forth in its product specification and assists a data user in determining a product's ability to satisfy the requirements for their particular application.

The objective of this International Standard is to provide principles for describing the quality for geographic data and concepts for handling quality information for geographic data.

Geographic information — Quality principles

1 Scope

This International Standard establishes the principles for describing the quality of geographic data and specifies components for reporting quality information. It also provides an approach to organizing information about data quality.

This International Standard is applicable to data producers providing quality information to describe and assess how well a dataset meets its mapping of the universe of discourse as specified in the product specification, formal or implied, and to data users attempting to determine whether or not specific geographic data is of sufficient quality for their particular application. This International Standard should be considered by organizations involved in data acquisition and purchase, in such a way that it makes it possible to fulfil the intentions of the product specification. It can additionally be used for defining application schemas and describing quality requirements.

As well as being applicable to digital geographic data, the principles of this International Standard can be extended to identify, collect and report the quality information for a geographic dataset, its principles can be extended and used to identify, collect and report quality information for a dataset series or smaller groupings of data that are a subset of a dataset.

Although this International Standard is applicable to digital geographic data, its principles can be extended to many other forms of geographic data such as maps, charts and textual documents.

This International Standard does not attempt to define a minimum acceptable level of quality for geographic data.

2 Conformance

Any product claiming conformance with this International Standard shall pass all the requirements described in the abstract test suite presented in 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 19108:2002, *Geographic information — Temporal schema*

ISO 19109:— [1)], *Geographic information — Rules for application schema*

ISO 19114:— [1)], *Geographic information — Quality evaluation procedures*

ISO 19115:— [1)], *Geographic information — Metadata*

1) To be published.

4 Terms and definitions

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

4.1
accuracy
closeness of agreement between a test result and the accepted reference value [ISO 3534-1]

NOTE A test result can be observations or measurements.

4.2
conformance
fulfilment of specified requirements [ISO 19105]

4.3
conformance quality level
threshold value or set of threshold values for data quality results used to determine how well a dataset meets the criteria set forth in its product specification or user requirements [ISO 19114]

4.4
data quality date
date or range of dates on which a data quality measure is applied

4.5
data quality element
quantitative component documenting the quality of a dataset [ISO 19101]

NOTE The applicability of a data quality element to a dataset depends on both the dataset's content and its product specification, the result being that all data quality elements may not be applicable to all datasets.

4.6
data quality evaluation procedure
operation(s) used in applying and reporting quality evaluation methods and their results

4.7
data quality measure
evaluation of a data quality subelement

EXAMPLE The percentage of the values of an attribute that are correct.

4.8
data quality overview element
non-quantitative component documenting the quality of a dataset [ISO 19101]

NOTE Information about the purpose, usage and lineage of a dataset is non-quantitative quality information.

4.9
data quality result
value or set of values resulting from applying a data quality measure or the outcome of evaluating the obtained value or set of values against a specified conformance quality level

EXAMPLE A data quality result of "90" with a data quality value type of "percentage" reported for the data quality element and its data quality subelement "completeness, commission" is an example of a value resulting from applying a data quality measure to the data specified by a data quality scope. A data quality result of "true" with a data quality value type of "boolean variable" is an example of comparing the value (90) against a specified acceptable conformance quality level (85) and reporting an evaluation of a kind, pass or fail.

4.10
data quality scope
extent or characteristic(s) of the data for which quality information is reported

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.

4.11
data quality subelement
component of a data quality element describing a certain aspect of that data quality element

4.12
data quality value type
value type for reporting a data quality result

EXAMPLE "boolean variable", "percentage", "ratio"

NOTE A data quality value type is always provided for a data quality result.

4.13
data quality value unit
value unit for reporting a data quality result

EXAMPLE "metre"

NOTE A data quality value unit is provided only when applicable for a data quality result.

4.14
dataset
identifiable collection of data [ISO 19115]

NOTE A dataset may be a smaller grouping of data which, though limited by some constraint such as spatial extent or feature type, is located physically within a larger dataset. Theoretically, a dataset may be as small as a single feature or feature attribute contained within a larger dataset.

4.15
dataset series
collection of datasets sharing the same product specification [ISO 19115]

4.16
feature
abstraction of real world phenomena [ISO 19101]

NOTE A feature may occur as a type or an instance. Feature type or feature instance should be used when only one is meant.

4.17
feature attribute
characteristic of a feature [ISO 19101]

NOTE A feature attribute has a name, a data type and a value domain associated with it. A feature attribute for a feature instance also has an attribute value taken from the value domain.

4.18
feature operation
operation that every instance of a feature type may perform [ISO 19110]

EXAMPLE 1 An operation upon the feature type "dam" is to raise the dam. The result of this operation is to raise the level of water in a reservoir.

EXAMPLE 2 An operation by the feature type "dam" might be to block vessels from navigating along a watercourse.

NOTE Feature operations provide a basis for feature type definitions.

4.19
metadata
data about data [ISO 19115]

4.20
product specification
description of the universe of discourse and a specification for mapping the universe of discourse to a dataset

4.21
quality
totality of characteristics of a product that bear on its ability to satisfy stated and implied needs [ISO 19101]

4.22
universe of discourse
view of the real or hypothetical world that includes everything of interest [ISO 19101]

5 Principles for describing the quality of geographic data

5.1 Components of data quality description

This International Standard can be used when

— identifying and reporting quality information;

— evaluating the quality of a dataset;

— developing product specifications and user requirements;

— specifying application schemas.

ISO 19114 and ISO 19115 describe schemas for reporting quality information.

ISO 19114 provides the framework for evaluating the quality of a dataset.

ISO 19109 describes the development of application schemas.

A quality description can be applied to a dataset series, a dataset or a smaller grouping of data located physically within the dataset sharing common characteristics so that its quality can be evaluated.

The quality of a dataset shall be described using two components:

— data quality elements;

— data quality overview elements.

Data quality elements, together with data quality subelements and the descriptors of a data quality subelement, describe how well a dataset meets the criteria set forth in its product specification and provide quantitative quality information.

Data quality overview elements provide general, non-quantitative information.

NOTE Data quality overview elements are critical for assessing the quality of a dataset for a particular application that differs from the intended application.

This International Standard recognizes that quantitative and non-quantitative quality information may have associated quality.

The quality about quality information may include a measure of the confidence or the reliability of the quality information. This type of information is recorded in ISO 19114's quality evaluation report.

Figure 1 provides an overview of data quality information.

Annex B provides a discussion of data quality concepts used to establish the components for describing the quality of geographic data.

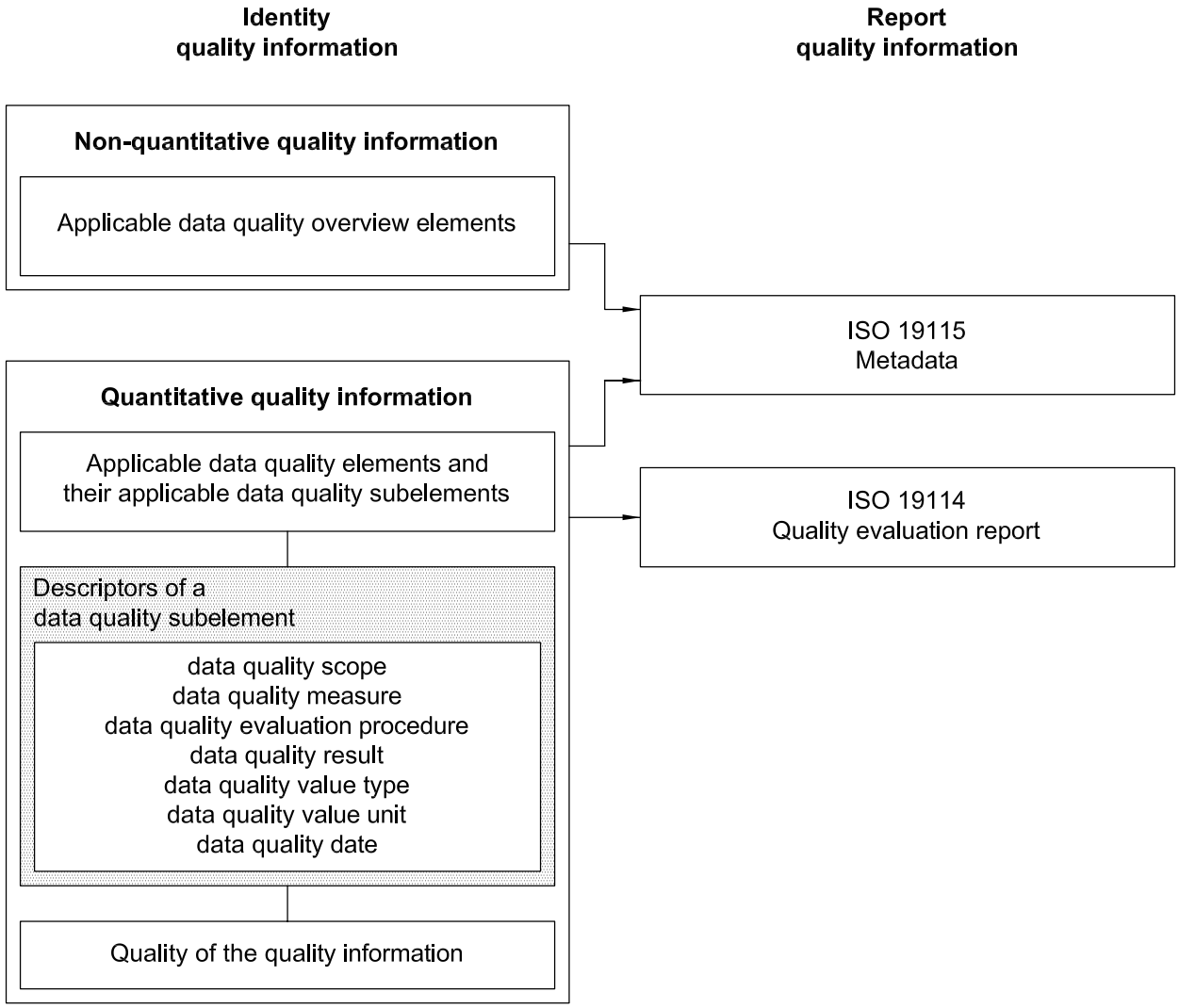

Figure 1 — An overview of data quality information

5.2 Data quality elements and data quality subelements

5.2.1 Data quality elements

The following data quality elements, where applicable, shall be used to describe how well a dataset meets the criteria set forth in its product specification:

— completeness: presence and absence of features, their attributes and relationships;

— logical consistency: degree of adherence to logical rules of data structure, attribution and relationships (data structure can be conceptual, logical or physical);

— positional accuracy: accuracy of the position of features;

— temporal accuracy: accuracy of the temporal attributes and temporal relationships of features;

— thematic accuracy: accuracy of quantitative attributes and the correctness of non-quantitative attributes and of the classifications of features and their relationships.

Additional data quality elements may be created to describe a component of the quantitative quality of a dataset not addressed in this International Standard.

5.2.2 Data quality subelements

For the data quality elements identified in 5.2.1, the following data quality subelements where applicable shall be used to describe aspects of the quantitative quality of a dataset:

— 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 attributes,

  — quantitative attribute accuracy: accuracy of quantitative attributes.

Additional data quality subelements may be created for any of the data quality elements.

5.2.3 Descriptors of a data quality subelement

Quality information shall be recorded for each applicable data quality subelement. The mechanism for completely recording information for a data quality subelement shall be the use of the seven descriptors of a data quality subelement:

— data quality scope;

— data quality measure;

— data quality evaluation procedure;

— data quality result;

— data quality value type;

— data quality value unit;

— data quality date.

NOTE The descriptors of a data quality subelement are defined in Clause 4.

5.3 Data quality overview elements

The following data quality overview elements where applicable shall be used to describe the non-quantitative quality of a dataset:

— purpose;

— usage;

— lineage.

Purpose shall describe the rationale for creating a dataset and contain information about its intended use.

NOTE A dataset's intended use is not necessarily the same as its actual use. Actual use is described using the data quality overview element usage.

Usage shall describe the application(s) for which a dataset has been used. Usage describes uses of the dataset by the data producer or by other, distinct, data users.

Lineage shall describe the history of a dataset and, in as much as is known, recount the life cycle of a dataset from collection and acquisition through compilation and derivation to its current form.

Lineage may contain two unique components:

— source information shall describe the parentage of a dataset;

— process step or history information shall describe a record of events or transformations in the life of a dataset, including the process used to maintain the dataset whether continuous or periodic, and the lead time.

Additional data quality overview element(s) shall describe an area of non-quantitative quality of a dataset not addressed in this International Standard.

6 Identifying the quality of geographic information

6.1 Identifying quantitative quality information

6.1.1 General

Clause 6.1 describe the general process for identifying quantitative quality information. Some of the subclauses may not be relevant in all cases.

6.1.2 Identifying applicable data quality elements

All data quality elements applicable to a dataset shall be identified. Some data quality elements may not be applicable for a particular type of dataset.

NOTE 1 Applicability of a data quality element should be determined by reference to a dataset's product specification.

EXAMPLE A dataset whose spatial references are postal references only will not have a data quality element of positional accuracy.

NOTE 2 Annex C contains examples of identifying applicable data quality elements.

6.1.3 Creating additional data quality elements

New data quality element(s) may be named and defined if the data quality elements listed in this International Standard do not sufficiently address a component of quality. The name and definition of an additional data quality element shall be included as a part of a dataset's quality information.

6.1.4 Identifying applicable data quality subelements

All applicable data quality subelements for each applicable data quality element shall be identified (at least one data quality subelement shall be identified as applicable for each applicable data quality element). Some of an applicable data quality element's data quality subelements may not be applicable to a particular type of dataset.

NOTE 1 Applicability of a data quality subelement should be determined by reference to a dataset's product specification.

NOTE 2 Annex C contains examples of identifying applicable data quality subelements.

6.1.5 Creating additional data quality subelements

New data quality subelement(s) may be named and defined if the data quality subelements listed in this International Standard do not sufficiently address an aspect of quality. The name and definition of an additional data quality subelement shall be included as a part of a dataset's quality information.

6.1.6 Using the descriptors of a data quality subelement

6.1.6.1 Data quality scope

At least one data quality scope shall be identified for each applicable data quality subelement. A data quality scope may be a dataset series to which a dataset belongs, the dataset or a smaller grouping of data located physically within the dataset sharing common characteristics. If a data quality scope cannot be identified, the data quality scope shall be the dataset.

NOTE Data quality scope(s) should be determined by reference to a dataset's product specification and the non-quantitative quality information provided for data quality overview elements.

Quality can vary within a dataset. Multiple data quality scopes may be identified for each applicable data quality subelement to more completely describe quantitative quality information. A data quality scope shall be adequately described. The following can be used to describe a data quality scope:

— the level (a dataset series to which a dataset belongs, the dataset or a smaller grouping of data located physically within the dataset sharing common characteristics);

— the types of items (lists of feature types, feature attributes and feature relationships) or specific items (lists of feature instances, attribute values and instances of feature relationships);

— the geographic extent;

— the temporal extent (the time frame of reference and accuracy of the time frame).

6.1.6.2 Data quality measure

One data quality measure shall be provided for each data quality scope. A data quality measure shall briefly describe and name, where a name exists, the type of test being applied to the data specified by a data quality scope and shall include bounding or limiting parameters.

NOTE 1 Examples of bounding or limiting parameters are confidence intervals and error rates.

This International Standard recognizes that the quality of a dataset is measured using a variety of tests. A single data quality measure might be insufficient for fully evaluating the quality of the data specified by a data quality scope and providing a measure of quality for all possible utilizations of a dataset. A combination of data quality measures can give useful information. Multiple data quality measures may be provided for the data specified by a data quality scope.

NOTE 2 ISO 19114 includes examples of names and descriptions of types of data quality measures.

6.1.6.3 Data quality evaluation procedure

One data quality evaluation procedure shall be provided for each data quality measure. A data quality evaluation procedure shall describe, or reference documentation describing, the methodology used to apply a data quality measure to the data specified by a data quality scope and shall include the reporting of the methodology.

NOTE 1 Examples of documentation are published articles or accepted industry standards.

NOTE 2 ISO 19114 includes a data quality evaluation procedure framework applicable to datasets and further clarifies the type of information to be reported in a data quality evaluation procedure.

6.1.6.4 Data quality result

One data quality result shall be provided for each data quality measure. The data quality result shall be either

— the value or set of values obtained from applying a data quality measure to the data specified by a data quality scope, or

— the outcome of evaluating the value or set of values obtained from applying a data quality measure to the data specified by a data quality scope against a specified acceptable conformance quality level. This type of data quality result is referred to in this International Standard as pass-fail.

Both types of data quality results identified in this International Standard may be provided.

NOTE ISO 19114 addresses the determination of conformance quality levels.

6.1.6.5 Data quality value type

One data quality value type shall be provided for each data quality result.

NOTE The data quality value type for pass-fail is “boolean variable”.

6.1.6.6 Data quality value unit

One data quality value unit, if applicable, shall be provided for each data quality result.

6.1.6.7 Data quality date

One data quality date shall be provided for each data quality measure in conformance with the requirements of ISO 19108’s temporal schema.

6.2 Identifying non-quantitative quality information

6.2.1 Identifying applicable data quality overview elements

Purpose of a dataset shall always be applicable.

All usage of a dataset that the producer is aware of shall be applicable.

Lineage of a dataset shall always be applicable. In extreme cases, information about lineage may not be known. Either lineage or an explanation of the lack of lineage information shall be reported.

Lineage for smaller groupings of data within a dataset specified by a data quality scope can be collected for and differ from the rest of the dataset’s lineage. Differing lineage may be provided for smaller groupings of data within a dataset specified by a data quality scope as a part of a dataset’s non-quantitative quality information for more complete non-quantitative quality information.

6.2.2 Creating additional data quality overview elements

New data quality overview element(s) may be named and defined if the data quality overview elements identified in this International Standard do not address an area of general non-quantitative quality. The name and definition of an additional data quality overview element shall be included as a part of its quality information.

7 Reporting quality information

7.1 Reporting quantitative quality information

Quantitative quality information shall be reported as metadata in conformance with the requirements of ISO 19115.

Quantitative quality information shall additionally be reported using a quality evaluation report in conformance with the requirements of ISO 19114.

7.2 Reporting non-quantitative quality information

Non-quantitative quality information shall be reported as metadata in conformance with the requirements of ISO 19115.

NOTE Non-quantitative quality information is not reported in ISO 19114’s quality evaluation report.

Annex A
(normative)

Abstract test suite

A.1 Abstract test suite

A.1.1 General

All of the test cases in this annex are of the Test Type: Basic.

A.1.2 Test case identifier: Component test

a) Test Purpose: to determine conformance by ensuring the components of quality are used in the quality description.

b) Test Method: examine the quality description and verify data quality elements (together with data quality subelements and the descriptors of a data quality subelement) have been used to provide quantitative quality information.

   Examine the quality description and verify data quality overview elements have been used to provide non-quantitative quality information.

c) Reference: ISO 19113:2002, 5.1.

A.1.3 Test case identifier: Validity test

a) Test Purpose: to determine conformance by ensuring the validity of the quality description.

b) Test Method: examine the quality description and verify its data quality elements and data quality subelements are listed in this International Standard or are additional and describe a component or aspect of quantitative quality that is not specifically identified in this International Standard.

   Examine the quality description and verify the descriptors of a data quality subelement identified in this International Standard have been used to describe quantitative quality.

   Examine the quality description and verify its data quality overview elements are listed in this International Standard or are additional and describe an area of non-quantitative quality that is not specifically identified in this International Standard.

c) Reference: ISO 19113:2002, 5.2 and 5.3.

A.1.4 Test case identifier: Quantitative quality applicability test

a) Test Purpose: to determine conformance by ensuring the applicability of the quantitative quality description.

b) Test Method: identify the product specification statements relevant to quantitative quality and use them to identify the applicable data quality elements and their applicable data quality subelements. Compare the applicable data quality subelements with the data quality subelements used in the quality description to ensure all data quality subelements applicable to the dataset have been identified and used in the quality description.

NOTE Conformance is valid if nonapplicable data quality subelements are additionally used to describe quantitative quality. However, the non-applicable data quality subelements cannot be subjected to further conformance testing.

c) Reference: ISO 19113:2002, 6.1.

A.1.5 Test case identifier: Non-quantitative quality applicability test

a) Test Purpose: to determine conformance by ensuring the applicability of the non-quantitative quality description.

b) Test Method: verify the applicable data quality overview elements are used to describe non-quantitative quality.

c) Reference: ISO 19113:2002, 6.2.

A.1.6 Test case identifier: Exclusiveness test

a) Test Purpose: to determine conformance by ensuring additional items in the quality description are exclusive and that sufficient information about an additional item is provided.

b) Test Method: examine all additional data quality elements and ensure each addresses a component of quantitative quality that is not specifically listed and described in this International Standard.

Examine all additional data quality subelements and ensure each addresses an aspect of quantitative quality that is not specifically listed and described in this International Standard.

Examine all additional data quality overview elements and ensure each addresses an area of non-quantitative quality that is not specifically listed and described in this International Standard.

Ascertain the name and a description of the additional item are a part of the quality description.

c) Reference: ISO 19113:2002, 6.1.3, 6.1.5 and 6.2.2.

A.1.7 Test case identifier: Correct use of the descriptors of a data quality subelement

a) Test Purpose: to determine conformance by verifying that the descriptors of a data quality subelement have been correctly used in the quality description.

b) Test Method: compare this International Standard and the quality information supplied for each applicable data quality subelement (including additional data quality subelements) to determine the occurrence rules for using descriptors of a data quality subelement have been followed.

c) Reference: ISO 19113:2002, 6.1.6.

A.1.8 Test case identifier: Reporting quality information as metadata

a) Test Purpose: to determine conformance by verifying the quality description is reported as metadata.

b) Test Method: verify that quantitative quality information has been reported as metadata in conformance with ISO 19115.

Verify that non-quantitative quality information has been reported as metadata in conformance with ISO 19115.

c) Reference: ISO 19113:2002, Clause 7.

A.1.9 Test case identifier: Reporting quantitative quality information using a quality evaluation report

a) Test Purpose: to determine conformance by verifying the quantitative quality of the quality description is reported as a quality evaluation report.

b) Test Method: verify that quantitative quality information is reported in a quality evaluation report in conformance with the requirements of ISO 19114.

c) Reference: ISO 19113:2002, 7.1.

Annex B
(informative)

Data quality concepts and their use

B.1 Background

A dataset is defined as an identifiable collection of data. Those data represent entities of the real-world which are characterized by having spatial, thematic and temporal aspects. The process of abstracting from the real-world to the universe of discourse involves modelling the potentially infinite characteristics of real-world entities into an ideal form defined by a position, a theme and a time for the reason of making intelligible and representable these entities. The universe of discourse is described by a product specification, against which the content of [parts of] a dataset is tested for its quality.

B.2 Purpose of data quality concepts

Since a dataset is not generally produced for a specific application but rather for a set of supposed applications, the quality of the dataset can only be assessed by knowing the data quality elements and the data quality overview elements. The data quality elements evaluate the difference between the dataset produced and the universe of discourse (that is the perfect dataset that corresponds to the product specification). The data quality overview elements provide general, non-quantitative information. The purpose gives information on the reasons for creating the dataset and on the intended use of the dataset. The usage provides information on the kind of application for which the dataset has been used. Lineage describes the history of the dataset.

Data quality concepts provide an important framework for data producers and data users. A data producer is given the means for specifying how well the mapping used to create a dataset reflects its universe of discourse. Data producers can validate how well a dataset meets the criteria set forth in its product specification. Data users are given the means for assessing a dataset derived from a universe of discourse identified as being coincident with requirements of a data user's application. Data users can assess quality to ascertain if a dataset can satisfy the requirements of an application (see Figure B.1).

B.3 The structure of datasets and components for quality description

A dataset may belong to a dataset series. The quality of all member datasets belonging to a dataset series is often the same. Data quality concepts recognize dataset series and allow for substituting and reporting the quality of a dataset series for a dataset.

A dataset can be viewed as containing a large but finite number of smaller groupings of data. Smaller groupings of data which share a commonality such as belonging to the same feature type, feature attribute or feature relationship or sharing a collection criteria or geographic extent can be expected to have similar quality. A smaller grouping of data can be as small as a feature instance, attribute value or occurrence of a feature relationship and, theoretically, data quality concepts allow each feature instance, attribute value and occurrence of a feature relationship of a dataset to have its own quality. The quality of smaller groupings of data cannot be assumed to be the same as the quality of the rest of the dataset to which they belong. Data quality concepts allow for reporting the quality of a dataset and additionally the differing quality of smaller groupings of data by identifying these groupings as the data specified by data quality scopes. The quality information reported for multiple data quality scopes provide a more complete picture of quality.

NOTE For a data producer, a product specification describes the universe of discourse and contains the parameters for constructing a dataset. For a data user, user requirements describe a universe of discourse, which may or may not match the dataset's universe of discourse. The true quality of a dataset is how well it represents a universe of discourse.

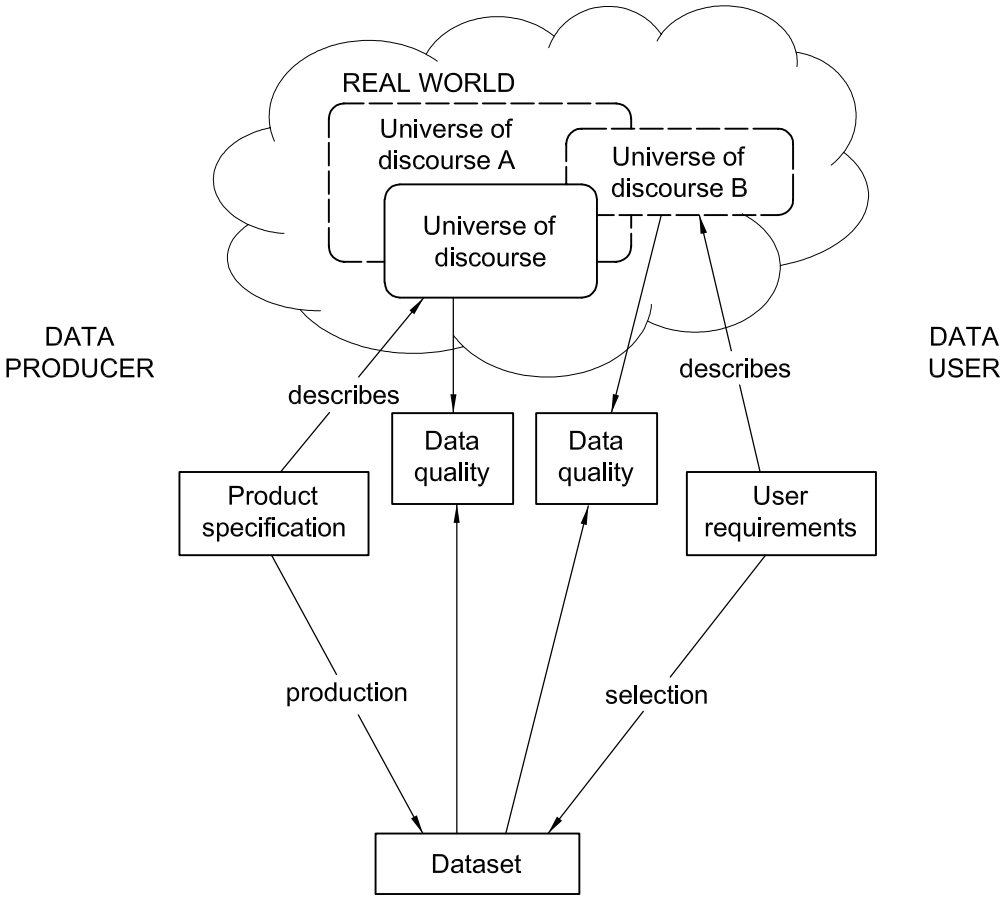

Figure B.1 — The framework provided by data quality concepts

To describe the quality of a dataset, two unique components of data quality information are recognized: quantitative quality components and non-quantitative quality components. Data quality elements are quantitative components of quality information; data quality overview elements are non-quantitative components of quality information.

Data quality elements allow for the measurement of how well a dataset meets the criteria set forth in its product specification. Data quality elements have distinct aspects known as data quality subelements. Data quality subelements can be measured or tested in various ways. Data quality concepts recognize that not all data quality elements nor all data quality subelements and their subsequent means of measurement and testing are applicable to a particular type of dataset. Additionally, some data quality subelements are applicable to and measured or tested for a dataset while others are applicable to and measured or tested for smaller groupings of data in a dataset specified by a data quality scope.

This International Standard identifies data quality elements primarily as a means of detecting and reporting separate categories of quality information. However, this International Standard additionally recognizes that frequently data quality subelements are interrelated. For example, a coordinate error may generate at least two kinds of errors, a positional error and a topological error. The meaning of the data quality subelements in terms of the product and manner in which the data quality subelements are handled are the purview of the quality evaluator.

Whereas data quality elements allow for the measurement of how well a dataset meets the criteria set forth in its product specification, data quality overview elements allow for additionally evaluating a dataset for a particular application by providing purpose, usage and lineage information.

B.4 Reporting quality information

B.4.1 When to report quality information

Datasets are continually being created, updated and merged with the result that the quality or a component of the quality of a dataset may change. The quality information of a dataset can be affected by three conditions:

a) when any quantity of data is deleted from, modified or added to a dataset;

b) when a dataset's product specification is modified;

c) when the real world has changed.

The first condition, a modification to a dataset, may occur quite frequently. Many datasets are not static. There is an increase in the interchange of information, the use of datasets for multiple purposes and an accompanying update and refinement of datasets to meet multiple purposes. If the reported quality of a dataset is likely to change with modifications to the dataset, the quality of a dataset should be reassessed and updated as required when changes occur.

Complete knowledge of all applicable data quality elements and all data quality overview elements with the exception of the data quality overview element usage should be available when a dataset is created. Only the data producer's usage (assuming the data producer actually uses the dataset) of a dataset can initially be reported. There is a reliance on data users to report uses of a dataset that differ from its intended purpose so that continual updates to this particular data quality overview element can be made to reflect occurring, unforeseen uses.

The second condition, a modification to a dataset's product specification, is most likely to occur before initial dataset construction and prior to the release of quality information. It is conceivable, however, that as a dataset is used its product specification is updated so that future modifications to the dataset will better meet the actual need. As the product specification changes, the quality of the current dataset also changes. The quality information for a dataset should always reflect the current dataset given its current product specification.

The third condition, a change of the real-world, occurs continuously. Change may be caused by natural phenomena such as movements in the earth's crust or erosion, but it is most often a result of human activity. Changes are often very rapid and dramatic. For this reason, the date of data collection is important when judging the quality of a dataset. In some cases, when known, even the rate of change is of interest.

The amount of, and storage requirements for, quality information can exceed that for the dataset. It is important to present quality information in a succinct, easily understood and easily retrievable format.

Quality information for a dataset series, a dataset or a grouping of data larger than a feature instance, attribute value or occurrence of a feature relationship specified by a data quality scope is generally contained in a metadata file or metadata repository.

B.4.2 Reporting quality information as metadata

B.4.2.1 Reporting quantitative quality information as metadata

Quantitative quality information may be recorded for multiple data quality scopes for a dataset. The data specified by a data quality scope may include a dataset series to which a dataset belongs, the dataset itself and smaller groupings of data physically located within the dataset.

Data quality concepts allow for the substituting of quantitative quality information for a dataset series to which a dataset belongs, as the quality of all members of the dataset series may be equal and be best measured at the dataset series level. The quality information may be stored as metadata with the dataset series, in which case the metadata of the dataset must provide a pointer to it; or the quality information may be repeated in, and be a part of, the dataset's metadata. If the quality information for a dataset is known to be unique and

differs from the quality for the rest of the dataset series, the dataset's unique and differing quality information should be provided for the dataset and a substitution is not recommended.

Quantitative quality information may be collected for and differ between a dataset and the many smaller groupings of data specified by a data quality scope. The amount of quantitative quality information being recorded is partially dependent on the number of identified data quality scopes. Quantitative quality information is typically recorded for the data specified by a data quality scope only when differing from quality information reported at a "higher" level. It is typical to begin at the uppermost levels of a dataset and work down through a dataset when reporting quality information. This is illustrated in Table B.1.

Table B.1 — Reporting hierarchical quantitative quality information

Data specified by a data quality scope	Positional accuracy/absolute or external accuracy data quality result	Reported positional accuracy/absolute or external accuracy data quality result
The dataset	1,35	1,35
Roads only	1,10	1,10
Streams only	1,35	Not reported
Railroads only	1,20	1,20
Pipelines only	1,80	1,80

NOTE The data quality value type of all of the data quality results in the above example is "distance". The data quality value unit for all of the data quality results in the above example is "metres".

In order to minimize the effort of creating, storing and interpreting metadata for the example, it is suggested that the quality information be reported as metadata only for the dataset and the data specified by the data quality scopes consisting of "roads", "railroads", and "pipelines". Quality information for the data specified by the data quality scope consisting of "streams" would be omitted.

ISO 19115 does not explicitly provide for the recording of quantitative quality information as metadata for feature instances, attribute values or single occurrences of feature relationships (this may be circumvented by identifying the data specified by a data quality scope as consisting of a feature instance, attribute value or single occurrence of a feature relationship). Quantitative quality information for single occurrences of items, when differing from their parent types, may be implemented by carrying the quality information as an attribute of the occurrence within a dataset. The hierarchical principle for reporting may also be applied between types and occurrences. It is suggested quantitative quality information for a feature instance be reported only when differing from that of its parent feature type, quality information for an attribute value be reported only when differing from that of its parent feature attribute and quality information for a single occurrence of a feature relationship be reported only when differing from that of its parent feature relationship. As the manner in which the attribution within a dataset occurs is dataset dependent, guidelines for the actual attribution of quality information on feature instances, attribute values and single occurrences of feature relationships is not provided.

Figure B.2 illustrates the hierarchical levels of a dataset for which quantitative quality information is reported as metadata and the suggested reporting method (either as part of a metadata file or repository or as attribution within a dataset).

ISO 19114 addresses the problem of an overabundance of quantitative quality information and the need to compress this information as well as the problem of requiring more detailed quantitative quality information through the use of a quality evaluation report.

B.4.2.2 Reporting non-quantitative quality information as metadata

Quality information for the data quality overview elements purpose and usage are considered essential information pertaining only to the dataset and are recorded only for the dataset. If the dataset belongs to a

dataset series and shares similar purpose and usage information, this International Standard allows the substitution of the dataset series' purpose and usage. In the case of substitution, the dataset's metadata may point to the purpose and usage in the dataset series' metadata rather than repeat the information.

Quality information for the data quality overview element lineage is considered essential information to the dataset. Additionally, lineage information may be collected for and differ between a dataset and the many smaller groupings of data specified by a data quality scope. It is suggested lineage for a smaller grouping of data specified by a data quality scope be reported only if it differs from the lineage being reported for the dataset. (Again, this International Standard allows the reporting of lineage information for a dataset series to which a dataset belongs rather than the actual lineage of the dataset. In the case of substitution, the dataset's metadata may point to the lineage in the dataset series' metadata rather than actually repeat the information.)

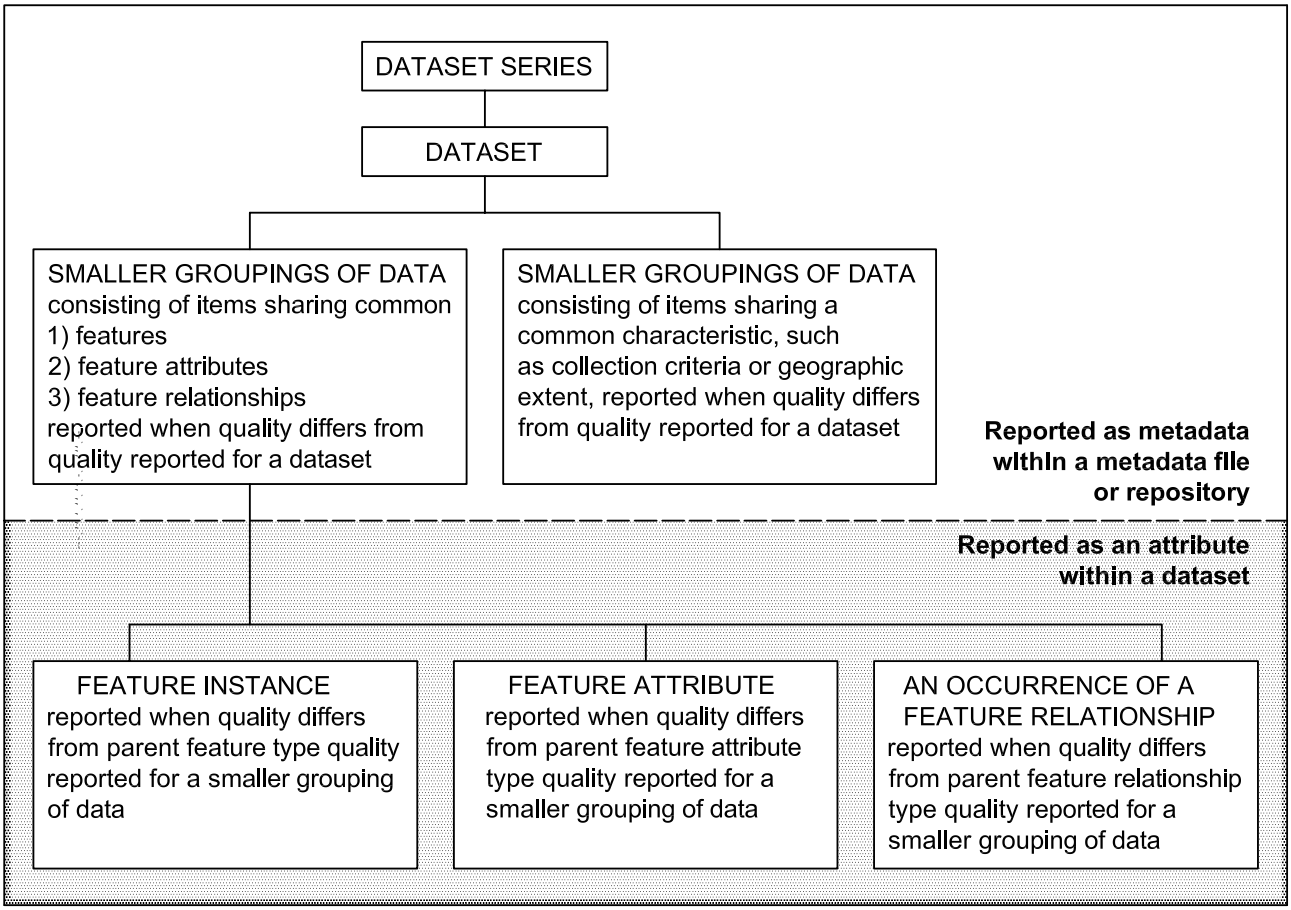

Figure B.2 — A suggested method of reporting quantitative quality information as metadata

Annex C
(informative)

Data quality elements, data quality subelements and data quality overview elements

C.1 Example 1 – Digital Chart of the World (DCW)

C.1.1 Overview

Example 1 illustrates a data producer's assessment of relevant quantitative quality information by identifying applicable data quality elements and data quality subelements using the product specification. Once a data quality subelement is determined to be applicable, the product specification is additionally used to identify appropriate data quality scope(s).

Example 1 also includes the data producer's assessment of non-quantitative quality information by compiling information for applicable data quality overview elements.

Example 1 does not include the actual reporting of the relevant quality information as metadata or using the quality evaluation report.

C.1.2 Background information

Dataset: Digital Chart of the World (DCW)

Product specification: Military Specification MIL-D-89009, 13 April 1992

Product description: (extracted and condensed from the product specification)

The DCW is a general-purpose global digital database designed to support Geographic Information System (GIS) applications. The DCW database contains five libraries represented on four CD-ROMs. The BROWSE library contains world-wide data at approximately 1:31 000 000 scale, supporting overview displays at a global scale. The four "detailed data" libraries, one for each CD-ROM, contain data for (1) North America, (2) Europe and Northern Asia, (3) South America, Africa, and Antarctica, and (4) Southern Asia and Australia at 1:1 000 000 scale.

The DCW utilizes the Vector Product Format (VPF) georelational data model to support a vector-based, thematically layered database. The BROWSE library contains eight thematic layers; each of the four detailed data libraries contains seventeen thematic layers. These thematic layers are: (1) Aeronautical, (2) Cultural Landmarks, (3) Data Quality, (4) Drainage, (5) Drainage-Supplemental, (6) Hypsography, (7) Hypsography-Supplemental, (8) Land Cover, (9) Ocean Features, (10) Physiography, (11) Political/Oceans, (12) Populated Places, (13) Railroads, (14) Roads, (15) Transportation Structure, (16) Utilities and (17) Vegetation.

Attributes and attribute value code combinations define each feature in the DCW.

C.1.3 Assessment of relevant quantitative quality information

NOTE The data producer did not find all data quality elements and data quality subelements to be applicable as not all were referenced in the product specification. Only applicable data quality elements and applicable data quality subelements are listed (see Table C.1).

Table C.1 — Quantitative quality information assessment, Example 1

Relevant paragraphs from the product specification documenting applicability		Applicable data quality element/data quality subelement	Identified data quality scope(s)
Number	**Text**		
4.1.2.e	Review a random sample of tiles in various coverages to ensure cartographic and attribute data completeness *(for all 17 data quality scopes).*	completeness commission	17 data quality scopes, each consisting of one thematic layer
4.1.2.g	Review the final vendor data format prior to Vector Product Format (VPF) conversion and record a final frequency count of all features for each thematic layer.		
3.1.2	Ensure the digital marginalia [information that originally appeared in notes, tables, and graphs on the borders of the Operational Navigation Chart (ONC) map sheets (the source material)] is included in the area attribute values of the data quality coverage and in the data quality tables.	completeness commission	the dataset
3.3	The unit of measure for the DCW database shall be the English measurement system.	logical consistency domain consistency	the dataset
4.1.2.g	A review of the final vendor data format prior to Vector Product Format (VPF) conversion. All attribute names and attribute definitions are verified.		
3.2.1	The horizontal datum shall be the current World Geodetic System. *(identified data quality scope = the dataset)*	logical consistency format consistency	1. the dataset 2. features encoded as points and polygons 3. all text strings
3.9	Where polygons on Operational Navigational Charts (source material) were below the minimum capturable size of less than or equal to 3,05 mm (0,12 in) in circumference, and consist of a single edge, they are represented as a point feature. *(identified data quality scope = encoded features)*		
4.1.2.d	A quality assurance check shall be used as a means for identifying and correcting problems with text placement. For example, a check shall be made to verify that text strings do not overlap. *(identified data quality scope = text strings)*		

Table C.1 *(continued)*

Relevant paragraphs from the product specification documenting applicability		Applicable data quality element/data quality subelement	Identified data quality scope(s)
Number	**Text**		
4.1.2.d	A quality assurance review of all thematic data plotted together to verify proper integration or feature positioning between coverages. *(identified data quality scope = the dataset)*	logical consistency topological consistency	1. the dataset 2. all connectors
4.1.2.g	A check shall be conducted to verify that the correct topology is present. Note – Refer to Military Standard 600006 for definitions of topology "levels" and testing requirements for evaluation procedures. *(identified data quality scope = the dataset)*		
3.9.d	Connection rules apply for linear road and railroad features in the DCW database. Where roads and railroads are broken by text on an Operation Navigational Chart (ONC), gaps wider than 2,54 mm (0,1 inch) in the network are coded as special connectors with TYPE attribute code of 8, and STATUS attribute code of 4, 6, or 9. Smaller gaps along a road or railroad are coded with the TYPE attribute code of 1 or 2, and STATUS attribute code 1, 2, 3, or 5. Note – Connectivity rules do not apply in situations where gaps are clearly due to natural obstructions. *(identified data quality scope = connectors)*		
3.1.2	The absolute vertical accuracy of the DCW is the same as for the original Operation Navigational Chart (ONC) and Jet Navigation Chart (JNC) lithographs at 90 % linear error, mean sea level. *(identified data quality scope = the dataset)*	position accuracy absolute or external accuracy (vertical)	1. the dataset 2. all contours 3. all spot elevations
3.1.2.a	The accuracy of the contours collected from the Operation Navigational Chart (ONC) source is ½ the contour interval of the original source, or plus or minus 150 m (plus or minus 500 feet). Note – Not applicable to feature symbols that are displaced as identified in Military Standard 6000003 (reference displacement rules). *(identified data quality scope = contours)*		
3.1.2.b	The accuracy of spot elevations collected from the Operation Navigational Chart (ONC) source is plus or minus 30 m (plus or minus 100 feet). Note – Not applicable to feature symbols that are displaced as identified in Military Standard 6000003 (reference displacement rules). *(identified data quality scope = spot elevations)*		

Table C.1 (*continued*)

Relevant paragraphs from the product specification documenting applicability		Applicable data quality element/data quality subelement	Identified data quality scope(s)
Number	**Text**		
3.1.1.1	The absolute horizontal accuracy of the DCW for all features derived from Operation Navigational Charts (ONC) is 2 040 m (6 700 feet) rounded to the nearest 5 m at 90 % circular error, World Geodetic System (WGS84). The absolute horizontal accuracy of the DCW for all features derived from Jet Navigational Charts (JNC) is 4 270 m (14 006 feet) at 90 % circular error. *(applicable to all three identified data quality scopes. The Antarctic region is a unique data quality scope as it has a unique source. The rationale for identifying roads as a unique data quality scope is that as roads are usually some of the best-surveyed features, a higher accuracy is expected.)*	position accuracy absolute or external accuracy (horizontal)	1. all regions of the world excluding Antarctic region 2. the Antarctic region 3. roads
3.4.5	Attributes and attribute value code combinations define each DCW feature.	thematic accuracy classification correctness	the dataset
4.1.2.c	Attribute code frequencies are automatically tabulated to identify invalid codes and unusual behaviour in attribute tables. Plots of the data shall then be created and exhaustively reviewed for attribute code accuracy. In most cases, a second and third plot shall be created to verify that identified errors have been corrected.		

The data producer's summary of relevant quantitative quality information is presented in Table C.2.

Table C.2 — Summary of relevant quantitative quality information, Example 1

Data quality element	Data quality subelement	Relevant?	Number of identified data quality scopes
completeness	commission	yes	1
	omission	yes	1
logical consistency	conceptual consistency	no	—
	domain consistency	yes	1
	format consistency	yes	3
	topological consistency	yes	2
positional accuracy	absolute or external accuracy	yes	3 — vertical 3 — horizontal
	relative or internal accuracy	no	—
	gridded data position accuracy	no	—
temporal accuracy	accuracy of a time measurement	no	—
	temporal consistency	no	—
	temporal validity	no	—
thematic accuracy	classification correctness	yes	1
	non-quantitative attribute correctness	no	—
	quantitative attribute accuracy	no	—

Compilation of non-quantitative quality information

Purpose: "The DCW is a general-purpose global digital database designed to support Geographic Information System (GIS) applications." — MIL-D-89009, 3.4.1 product description.

Usage: Usage 1 **Developing databases:** "ESRI itself has used the DCW as a source for developing the ArcWorld database for use with its proprietary GIS software packages ARC/INFO and Arcview." — The Digital Chart of the World — A Review, this usage found on the Internet.

Usage 2 **Developing electronic map series:** "A series of electronic maps in WHEAT format was prepared from the Digital Chart of the World, a set of 1:1 000 000 electronic maps based on the Defense Mapping Agency's Operational Navigational Charts. These portions of the Digital Chart of the World database were imported for several purposes: to aid natural resource development in the developing world for basic needs development, to provide maps suitable for relief work in the Third World, and to provide example datasets for use with WHEAT. It is hoped that providing regional topographic maps in an easy-to-use format will facilitate groundwater exploration, agronomic planning, and the logistics of relief projects." — User's Manual for Digital Chart of the World 1 Quadrangles, Geohydrology Section, Kansas Geological Survey, this usage found on the Internet.

Usage 3 **Developing 3-D visualisations of a DEM:** "The image below was prepared in ARC/INFO. It was a fairly straightforward task. The steps involved were: 1. Prepare points with heights of the Holy Land in ARC/INFO from DCW dataset, 2. Create a TIN from the points of the Holy Land, 3. Generate contours from the TIN, 4. Use the contours to REGISTER the Landsat TM image of the Holy Land in ARC/INFO, 5. Use SURFACED FAULTS to do all the dirty work, 6. Use SURFACE OBSERVER RELATIVE to set up the azimuth and zenith angles of the observer (how high up you are, and from where you are looking e.g. from the south or the north), 7. Use SURFACE DRAPE to drape the TM image over the Holy Land TIN." — 3-D visualisation of the Holy Land, this usage found on the Internet.

Lineage: Source: The DCW database content is based primarily on the feature content of the 1:1 000 000-scale Defense Mapping Agency (now the National Imagery and Mapping Agency) Operation Navigational series (all regions excluding the Antarctic region). The Operation Navigational Charts used to create the product were produced by the Defense Mapping Agency between the years of 1974 and 1991.

Process Step: Stable-based positives were produced from the original reproduction negatives (up to 35 negatives per Operation Navigational Chart sheet) and either digitized through a scanning raster to vector conversion or hand digitized into vector form. The vector data were then tagged with attribute information using ARC/INFO software. Transformation to geographic coordinates was performed using the projection graticules for each sheet. Digital information was edge matched between sheets to create large regional datasets. The regional datasets were then subdivided into 5×5 tiles and converted from ARC/INFO into Vector Profile Format. The data was then pre-mastered for CD-ROM. Quality control was performed by a separate group for each step of this process. Processing was completed in January, 1991.

Source: The DCW database content for the Antarctic region only is based on the feature content of the 1:2 000 000-scale DMA Jet Navigational Chart series. The Jet Navigational Charts used to create the product were produced by the Defense Mapping Agency (now the National Imagery and Mapping Agency) between the years of 1974 and 1991.

Process Step: See Process Step for Operation Navigational series.

Source: The DCW airport information was taken from the Digital Aeronautical Flight Information File (DAFIF). The DAFIF was produced by the Defense Mapping Agency (now the National Imagery and Mapping Agency). The DAFIF consisted of airport records containing (1) name, (2) ICAO, (3) position, (4) elevation, and (5) type and was produced and published in 1991. The publisher was the Defense Mapping Agency (DMA, St. Louis, MO, United States). The DAFIF was released on magnetic tape.

Process Step: The DAFIF was transferred digitally directly into the VPF files by Environmental Science Research Institute staff. Processing was completed in January, 1991.

Source: Advance Very High Resolution Radimeter (AVHRR) data supplied by the USGS EROS Data Centre (EROS Data Centre, Sioux Falls, South Dakota, USA) was used to determine the six vegetation types covering the continental United States. The data was supplied in the form of remotely sensed imagery having a scale of 1:000 000 and datum of WGS 84.

Process Step: Daily AVHRR images were averaged for two-week time periods over the entire United States growing season. These averaged images, their rates of change, elevation information, and other data was used to produce a single land classification image of the continental United States.

Process Step: The EROS Data Centre image's raster files were converted to vector polygon, splined (stairstepping removed), thinned (all polygons under 2 km^2 deleted), and tied to existing DCW polygons (water bodies, built-up areas). The resulting file was tiled and converted to a VPF Vegetation layer. All production was accomplished by Environmental Science Research Institute staff. Processing was completed in January, 1991.

C.2 Example 2 — Digital Terrain Map (DTM)

C.2.1 Overview

Example 2 illustrates a data producer's assessment of relevant quantitative quality information by identifying applicable data quality elements and data quality subelements for a data quality scope equalling the dataset using the product specification.

Example 2 does not include the actual reporting of the relevant quality information as metadata or using the quality evaluation report.

C.2.2 Background information

Dataset: DTM of a hydrographic basin.

Product specification: Specifications on Geographic Information System (GIS) for use with hydrographic basin plans, National Water Institute, 1998.

Product description: The DTM should be build using the 1:25 000 topographic map produced by the National Mapping Agency and presented in a grid structure. Cell size should be 25 m and the origin of the grid should be placed at a multiple of 25 m. The DTM should allow for hydrological modelling operations.

The DTM can be split into different files using hydrographic basin limits and the country border supplied by the National Water Institute.

C.2.3 Assessment of relevant quantitative quality information

NOTE The data producer identified the only relevant data quality scope to be the dataset. The data producer did not find all data quality elements and data quality subelements to be applicable (the product specification did not directly reference any of the data quality elements and their data quality subelements). Only applicable data quality elements and applicable data quality subelements are listed (see Table C.3).

Table C.3 — Quantitative quality information assessment, Example 2

Relevant paragraphs from the product specification documenting applicability	Applicable data quality element/ data quality subelement
The DTM can be split into different files using the hydrographic basin limits and the country boundary supplied by the National Water Institute.	completeness commission
The DTM should cover completely its corresponding hydrographic basin area.	completeness omission
The Product Specification does not include a reference to this requirement, however, the data producer has indicated the need to ensure there are no heights above 2 000 m in the dataset.	logical consistency domain consistency
Cell size should be of 25 m and the origin of the grid should be placed at a multiple of 25 m. The DTM should allow for hydrological modelling operations.	logical consistency format consistency
The DTM should support hydrological modelling.	logical consistency topological consistency
The Product Specification does not contain a direct reference to positional accuracy. The data producer assumes, however, as data from a 1:25 000 topographic map was used that elevation errors should be less than 4 m.	positional accuracy absolute or external accuracy

The data producer's summary of relevant quantitative quality information for data quality scope equalling the dataset is presented in Table C.4.

Table C.4 — Summary of relevant quantitative quality information, Example 2

Data quality element	Data quality subelement	Relevant?
completeness	commission	yes
	omission	yes
logical consistency	conceptual consistency	no
	domain consistency	yes
	format consistency	yes
	topological consistency	yes
positional accuracy	absolute or external accuracy	yes
	relative or internal accuracy	no
	gridded data position accuracy	no
temporal accuracy	accuracy of a time measurement	no
	temporal consistency	no
	temporal validity	no
thematic accuracy	classification correctness	no
	non-quantitative attribute correctness	no
	quantitative attribute accuracy	no

C.3 Example 3 – A Land Use Dataset

C.3.1 Overview

Example 3 illustrates a data producer's assessment of relevant quantitative quality information by identifying applicable data quality elements and data quality subelements for a data quality scope equalling the dataset using the product specification.

Example 3 does not include the actual reporting of the relevant quality information as metadata or using the quality evaluation report.

Example 3 includes a compilation of relevant non-quantitative quality information.

C.3.2 Background Information

Dataset: A Land Use Dataset comprised of a number of subsets (one per county) prepared to support generic Geographic Information System (GIS) projects.

Product specification: A requirement for the product to reproduce land use maps, originally built in analogue format, and approved according to legal procedures.

Product description: The dataset contains land use subsets resulting from digitizing paper source cartography. Statistical elements related to image registration are reported.

Each subset matches a county having a boundary officially established and supplied to the data producer.

Objects in each subset are polygons classified according to a common key of land use classes.

C.3.3 Assessment of relevant quantitative quality information

NOTE The data producer identified the only relevant data quality scope to be the dataset. The data producer did not find all data quality elements and data quality subelements to be applicable (the Product Specification did not directly reference any of the data quality elements and their data quality subelements). Only applicable data quality elements and applicable data quality subelements are listed (see Table C.5).

Table C.5 — Quantitative quality information assessment, Example 3

Relevant paragraphs from the product specification documenting applicability	Applicable data quality element/ data quality subelement
Number of missing polygons; area associated with missing polygons; identification of incomplete coverage of the area defined by the county limits due to misuse of the given county boundaries.	completeness commission
The identification of the roads should be made according to the names listed in a file supplied	completeness omission
Objects in the subsets are polygons classified according to a common key of land use classes.	logical consistency domain consistency
Less than 10 percent of the difference vectors should be over 1 mm the scale of the source.	positional accuracy absolute or external accuracy

The data producer's summary of relevant quantitative quality information for data quality scope equalling the dataset is presented in Table C.6.

Table C.6 — Summary of relevant quantitative quality information, Example 3

Data quality element	Data quality subelement	Relevant?
completeness	commission	yes
	omission	yes
logical consistency	conceptual consistency	no
	domain consistency	yes
	format consistency	no
	topological consistency	no
positional accuracy	absolute or external accuracy	yes
	relative or internal accuracy	no
	gridded data position accuracy	no
temporal accuracy	accuracy of a time measurement	no
	temporal consistency	no
	temporal validity	no
thematic accuracy	classification correctness	no
	non-quantitative attribute correctness	no
	quantitative attribute accuracy	no

Compilation of non-quantitative quality information

Purpose: The Land Use Dataset is intended to support land management and planning activities, being as similar as possible as the original legally approved analogue format.

Usage: The original paper sources have been used for land management in each county.

Lineage: Source: Analogue land use maps, drawn over 1:25 000 and 1:10 000 topographic maps.

Process Step: Original paper maps were scanned. Images were registered using at least nine points of known coordinates.

C.4 Example 4 — A 3 Dimensional (3D) Road Network Database

C.4.1 Overview

Example 4 illustrates a data producer's assessment of relevant quantitative quality information by identifying applicable data quality elements and data quality subelements for a data quality scope equalling the dataset using the product specification.

Example 4 does not include the actual reporting of the relevant quality information as metadata or using the quality evaluation report.

C.4.2 Background information

Dataset: A vector 3D Road Network Database. The dataset is intended to serve as source data to build new databases such as (1) a 2 dimensional road network database with attributes about paving condition (referenced by distance measured in 3D from the origin of each segment) and (2) a 2 dimensional road network database with references to speeds and (3) road profiles.

Product specification: A requirement for a product consisting of a set of 3D vector polylines describing the network of national roads.

Product description: Each 3D polyline corresponds to an arc of the road network. Each arc is stored in an individual layer with the name corresponding to its identification. The identification of the roads should be made according to the names listed in a file supplied to the data producer.

Positional errors, both planimetric and altimetric, should not be greater than 2 m. 3D polylines should connect to one another in crossings and be broken only in crossings or when there is a change of identification.

C.4.3 Assessment of relevant quantitative quality information

NOTE The data producer identified the only relevant data quality scope to be the dataset. The data producer did not find all data quality elements and data quality subelements to be applicable (the Product Specification did not directly reference any of the data quality elements and their data quality subelements). Only applicable data quality elements and applicable data quality subelements are listed (see Table C.7).

Table C.7 — Quantitative quality information assessment, Example 4

Relevant paragraphs from the product specification documenting applicability	Applicable data quality element/ data quality subelement
No specific statement; refer to completeness, omission statement.	completeness commission
The dataset should contain all roads classified as National roads.	completeness omission
The identification of the roads should be made according to the names listed in a file supplied to the data producer.	logical consistency domain consistency
3D polylines should connect each other in crossings and be broken only in crossings or when a change of identification occurs.	logical consistency topological consistency
3D polylines should connect each other in crossings and be broken only in crossings or when a change of identification occurs.	positional accuracy absolute or external accuracy

The data producer's summary of relevant quantitative quality information for data quality scope equalling the dataset is presented in Table C.8.

Table C.8 — Summary of relevant quantitative quality information, Example 4

Data quality element	Data quality subelement	Relevant?
completeness	commission	yes
	omission	yes
logical consistency	conceptual consistency	no
	domain consistency	yes
	format consistency	no
	topological consistency	yes
positional accuracy	absolute or external accuracy	yes
	relative or internal accuracy	no
	gridded data position accuracy	no
temporal accuracy	accuracy of a time measurement	no
	temporal consistency	no
	temporal validity	no
thematic accuracy	classification correctness	no
	non-quantitative attribute correctness	no
	quantitative attribute accuracy	no

ICS 35.240.70

Price based on 29 pages