Pan American Institute of Geography and History

GUIDE FOR THE POSITIONAL ACCURACY ASSESSMENT OF GEOSPATIAL DATA

2021

Pub. 563

Pan American Institute of Geography and History

GUIDE FOR THE POSITIONAL ACCURACY ASSESSMENT OF GEOSPATIAL DATA

2021

Pub. 563

Occasional Publications

This document is a result of the Project "Proposal for the adoption of methodologies and procedures for the evaluation of the quality of geographic information for the member states of the Pan American Institute of Geography and History", financed by the Pan American Technical Assistance Program (PAT) of the year 2018, belonging to the 2010-2020 Agenda of the PAIGH.

This Project was started under the direction of Mr. Edison Rojas and, later, Ms. Rosario Casanova, both from the Institute of Surveying of the National University of the Republic (Uruguay)

This Project was presented by the national section of Uruguay and had the participation of the following countries: Argentina, Brazil, Chile, Ecuador, Mexico and Spain.

Edited by: Pan American Institute of Geography and History - General Secretariat

Authors: Francisco Javier Ariza-López, José Luis García-Balboa, José Rodríguez-Avi, Joselyn Robledo Ceballos

Editor and coordinator: Francisco Javier Ariza López

To refer to this document, please use the following:

Ariza-López F.J., García-Balboa, J.L., Rodríguez-Avi, J., Robledo J., (2021). *Guide for the Positional Accuracy Assessment of Geospatial Data.* Pan American Institute of Geography and History, Occasional Publication # 563

To send suggestions, comments and report errors write to:fjariza@ujaen.es

Pan American Institute of Geography and History
General Secretary
P.O.B. 18879
11870 México, D.F.
Telephone: (52-55)5277-5791 / 5277-5888 / 5515-1910 Fax: (52-55)5271-6172
e-mail: info@ipgh.org / publicaciones@ipgh.org

Contents

INTRODUCTION

Positional accuracy has always been considered a defining and essential element of the quality of any cartographic product (Ariza-López 2002) as it affects factors such as geometry, topology, and thematic quality, and is directly related to the interoperability of spatial data.

Considering the widespread use of geospatial information and the interoperability requirements of different geomatics applications and spatial data infrastructures (SDIs), it is crucial to ensure information quality, as this is the only means of guaranteeing reliable solutions when making decisions.

The quality of spatial data products has to be defined within the appropriate specification framework (e.g., ISO 19131; ISO 2007). The production processes should be properly designed and managed so that the intended quality is assured during the production process, and thus attained in the product. Quality assessments are carried out to confirm such quality, which should be based on standardized and well-defined methods. The results of these assessments should be used by producers to understand, improve, and manage their production processes. These results should also be published as metadata for users to be aware of the quality of the products they choose to use. This represents a very broad general quality framework, and this guide only addresses one of the indicated stages, namely quality assessment, specifically positional accuracy assessment.

In line with these broad objectives, it was considered relevant to provide the geomatics sector with a guide for positional accuracy assessment[1] as a critical component of official types of spatial data[2]. This is the defining component of spatial data, and, although there are many publications, it is also true that gaps remain, especially concerning application. Moreover, from a pan-American regional perspective, there are notable differences between countries, and so this guide provides an instrument for achieving a greater degree of standardization[3] in data-assessment processes.

Positional accuracy assessment methods are standardized processes to either estimate or control quality. This estimation consists of determining a reliable value of the property of interest in the data product, while quality control involves deciding whether or not the property of interest in that data product reaches a certain quality level.

Positional accuracy assessment methods (PAAMs) have also evolved over time, from the simple National Map Accuracy Standard (NMAS; USBB 1947) to the more recent, complex, and sophisticated American Society for Photogrammetry and Remote Sensing (ASPRS) Positional Accuracy Standards for Digital Geospatial Data (ASPRS 2015). Furthermore, new acquisition technologies, such as Global Navigation Satellite Systems (GNSS), enable the collection of absolute coordinates in the field with high accuracy, which increases confidence in these assessments.

---

[1] This activity was developed within the project 'Diagnosis of the current situation of methodologies and procedures used for quality assessment of Geographic Information' (*Diagnóstico de la situación actual sobre las metodologías y procedimientos empleados para la evaluación de la calidad de la Información Geográfica*), funded under the 2018 call 'Technical Assistance Projects' (*Proyectos de Asistencia Técnica* - PAT). See article 'Proposal of a guide for the positional accuracy assessment of spatial data', published in the journal Cartográfica No. 99 (2019).

[2] The terms spatial data, geospatial data, geographic data, and geographic information are considered synonymous in this document.

[3] The term 'standardization' is used in this document to refer to the activity of normalization within an organization, as opposed to the term 'normalization', which is considered more appropriate for the activities of a national standardization body or agency.

The sources of uncertainty arising from the different stages of the production process are multiple, varied, and dependent on the process itself (e.g., image capture, geodetic process, orientation process, restitution, acquisition and standardization of databases, etc.). Considering this complex scenario, it is impossible to guarantee a 'perfect' product, which is, in our case, the absence of positional errors. The key aspect is to control the levels of uncertainty by the end of the production process, and comply with international standards and their parameters and tolerances for each scale (resolution) of interest.

OBJECTIVE

This report aims to: a) provide a theoretical-practical guide that supports the development of absolute positional accuracy assessments[4] in a correct and reliable manner, focusing on improving the metaquality of results and processes. This requires that both theoretical and practical aspects are considered together, as a correct assessment cannot be carried out without both components; b) develop this guide within the framework established by the ISO 19100 standards of ISO Technical Committee 211, clarifying, as far as possible, the aspects that typically give rise to user uncertainty; and c) propose a template report that summarizes the characteristics and results of positional accuracy assessments. The report is intended to be useful for both producers and users. As far as is reasonable and as a first approximation, effectiveness and efficiency criteria are considered throughout the entire guide so that the processes proposed for quality assessment and description are both optimal and cost-effective.

SCOPE

This guide focuses primarily on processes for office-based application but also provides guidelines for reference data collection processes in the field. In addition, this guide specifically includes considerations based on four existing standards (NMAS, EMAS, NSSDA, UNE 148002), which often provide different and complementary perspectives.

CONFORMITY

As it is only intended to provide general guidance, conformity requirements are intentionally excluded. Nevertheless, any positional accuracy assessment processes based in whole or in part on this guide (or on other considerations), and implemented by an organization should provide a mechanism for verifying conformity[5].

NORMATIVE REFERENCES

The standards listed below have provisions applicable to this guide. All standards are subject to revision, so the dates refer to the documents in force when this guide was published.

---

[4] However, all information in this guide is also valid for the case of relative positional accuracy.

[5] The Spanish UNE 148002 (UNE 2016) standard presents an example in this regard.

- FGDC (1998). FGDC-STD-007: Geospatial Positioning Accuracy Standards, Part 3. National Standard for Spatial Data Accuracy, Federal Geographic Data Committee, Reston, USA.
- ISO (1985). ISO 2859-2:1985 Sampling procedures for inspection by attributes – Part 2: Sampling plans indexed by limiting quality (LQ) for isolated lot inspection.
- ISO (1999). ISO 2859-1:1999 Sampling procedures for inspection by attributes – Part 1: Sampling schemes indexed by acceptance quality limit (NCA) for lot-by-lot inspection.
- ISO 19115-1:2014 Geographic information – Metadata – Part 1: Fundamentals.
- ISO 19131:2007 Geographic information – Data product specifications.
- ISO 19157:2013 Geographic information – Data quality.
- ISO 3534-1:2006 Statistics – Vocabulary and symbols – Part 1: General statistical terms and terms used in probability.
- ISO 3534-2:2006 Statistics – Vocabulary and symbols – Part 2: Applied statistics.
- ISO 5725-1:1994 Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions.
- JCGM 200:2012 International vocabulary of metrology – Basic and general concepts and associated terms (VIM). 3rd edition.

TERMS AND DEFINITIONS

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

- Quality: Degree to which a set of inherent characteristics fulfils requirements [ISO 19101-1:2014]

  Note 1: The term “quality” can be used with adjectives such as poor, good, or excellent.

  Note 2: “Inherent”, as opposed to “assigned”, means existing in something, especially as a permanent characteristic. [SOURCE: ISO 9000:2005, 3.1.1]

- Dataset: Identifiable collection of data [ISO 19101-1:2014].
- Quality control: Part of quality management focused on fulfilling quality requirements [ISO 9000:2015].
- ‘Data quality element’: Quantitative component documenting the quality of a dataset [19113:2002].
- En 19157:2013 no aparece ‘data quality element’ en el listado de definiciones, pero en el apartado 7.4.1. lo cita como “component describing a certain aspect of the quality of geographic data”. Propongo poner esta por ser más reciente e indicar entre corchete [ISO 19157:2014, 7.4.1.]
- Measurement error (error of measurement, error): Measured quantity value minus a reference quantity value [JCGM 200:2012].

  NOTE 1: The concept of ‘measurement error’ can be used both:

  a) when there is a single reference quantity value to refer to, which occurs if a calibration is made by means of a measurement standard with a measured quantity value having a negligible measurement uncertainty or if a conventional quantity value is given, in which case the measurement error is known, and

  b) if a measurand is supposed to be represented by a unique true quantity value or a set of true quantity values of negligible range, in which case the measurement error is not known.

  NOTE 2: Measurement error should not be confused with production error or mistake.

- Random measurement error (random error of measurement, random error): Component of measurement error that in replicate measurements varies in an unpredictably manner [JCGM 200:2012].

  NOTE 1 A reference quantity value for a random measurement error is the average that would ensue from an infinite number of replicate measurements of the same measurand.

  NOTE 2 Random measurement errors of a set of replicate measurements form a distribution that can be summarized by its expectation, which is generally assumed to be zero, and its variance.

  NOTE 3 Random measurement error equals measurement error minus systematic measurement error.

- Systematic measurement error (systematic error of measurement, systematic error): Component of measurement error that in replicate measurements remains constant or varies in a predictable manner [JCGM 200:2012].

  NOTE 1 A reference quantity value for a systematic measurement error is a true quantity value, or a measured quantity value of a measurement standard of negligible measurement uncertainty, or a conventional quantity value.

  NOTE 2 Systematic measurement error, and its causes, can be known or unknown. A correction can be applied to compensate for a known systematic measurement error.

  NOTE 3 Systematic measurement error equals measurement error minus random measurement error.

  - Measurement accuracy (accuracy of measurement, accuracy): Closeness of agreement between a measured quantity value and a true quantity value of a measurand [JCGM 200:2012].

  NOTE 1 The concept 'measurement accuracy' is not a quantity and is not given a numerical quantity value. A measurement is said to be more accurate when it offers a smaller measurement error.

  NOTE 2 The term "measurement accuracy" should not be used for measurement trueness and the term "measurement precision" should not be used for 'measurement accuracy', which, however, is related to both these concepts.

  NOTE 3 'Measurement accuracy' is sometimes understood as closeness of agreement between measured quantity values that are being attributed to the measurand.

- Positional accuracy: Accuracy of the position of features within a spatial reference system [ISO19157:2014, 7.4.4].
- Measurement uncertainty (uncertainty of measurement, uncertainty): Non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used [JCGM 200:2012].

  NOTE 1 Measurement uncertainty includes components arising from systematic effects, such as components associated with corrections and the assigned quantity values of measurement standards, as well as the definitional uncertainty. Sometimes estimated systematic effects are not corrected for but, instead, associated measurement uncertainty components are incorporated.

  NOTE 2 The parameter may be, for example, a standard deviation called standard measurement uncertainty (or a specified multiple of it), or the half-width of an interval, having a stated coverage probability.

  NOTE 3 Measurement uncertainty comprises, in general, many components. Some of these may be evaluated by Type A evaluation of measurement uncertainty from the statistical distribution of the quantity values from series of measurements and can be characterized by standard deviations. The other components, which may be evaluated by Type B evaluation of measurement uncertainty, can also be characterized by standard deviations, evaluated from probability density functions based on experience or other information.

NOTE 4 In general, for a given set of information, it is understood that the measurement uncertainty is associated with a stated quantity value attributed to the measurand. A modification of this value results in a modification of the associated uncertainty.

- Standard measurement uncertainty (standard uncertainty of measurement, standard uncertainty): Measurement uncertainty expressed as a standard deviation [JCGM 200:2012].
- Measurement bias (bias): Estimate of a systematic measurement error [JCGM 200:2012].
- Measurement trueness (trueness of measurement, trueness): Closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value [JCGM 200:2012].

NOTE 1 Measurement trueness is not a quantity and thus cannot be expressed numerically, but measures for closeness of agreement are given in ISO 5725.

NOTE 2 Measurement trueness is inversely related to systematic measurement error, but is not related to random measurement error.

NOTE 3 "Measurement accuracy" should not be used for 'measurement trueness'.

- Measurement precision (precision): Closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions [JCGM 200:2012].

NOTE 1 Measurement precision is usually expressed numerically by measures of imprecision, such as standard deviation, variance, or coefficient of variation under the specified conditions of measurement.

NOTE 2 The 'specified conditions' can be, for example, repeatability conditions of measurement, intermediate precision conditions of measurement, or reproducibility conditions of measurement (see ISO 5725-1:1994).

NOTE 3 Measurement precision is used to define measurement repeatability, intermediate measurement precision, and measurement reproducibility.

NOTE 4 Sometimes "measurement precision" is erroneously used to mean measurement accuracy.

- Standalone quality report: free text document providing fully detailed information about data quality, evaluations, results and measures used [ISO 19157:2013].
- Metaquality: Information describing the quality of data quality [ISO 19157:2013].
- Direct evaluation method: method of evaluating the quality of a dataset based on inspection of the items within the dataset [ISO 19157:2013].
- Sample: Subset of a population made up of one or more sampling units [ISO 3534-2:2006].
- Sampling unit: One of the individual parts into which a population is divided [ISO 3534-2:2006].
- Data product: dataset or dataset that conforms to a data product specification [ISO 19131:2007].
- Universe of discourse: View of the real or hypothetical world that includes everything of interest [ISO 19101-1:2014].
- Homoscedasticity: An equal behavior of the variances that characterize each of the components of the position.
- Outlier: Member of a small subset of observations that appears to be inconsistent with the remainder of a given sample [ISO 16269-4:2010].

NOTE 1 The classification of an observation or a subset of observations as outlier(s) is relative to the chosen model for the population from which the data set originates. This or these observations are not to be considered as genuine members of the main population.

NOTE 2 An outlier may originate from a different underlying population, or be the result of incorrect recording or gross measurement error.

NOTE 3 The subset may contain one or more observations.

- Easily identifiable point: A point easy to identify at a glance.
- Well-defined point: A point for which the position is unambiguous (e.g., there is no problem of dilution of geometric precision).
- Standard: A type, model, norm, rule or reference (RAE, 2021). A technical solution that has a dominant position in a sector of activity.
- Estimation: Procedure that obtains a statistical representation of a population from a random sample drawn from this population [ISO 3534-1:2006].

NOTE 1 In particular, the procedure involved in progressing from an estimator to a specific estimate constitutes estimation.

NOTE 2 Estimation is understood in a rather broad context to include point estimation, interval estimation or estimation of properties of populations.

NOTE 3 Frequently, a statistical representation refers to the estimation of a parameter or parameters or a function of parameters from an assumed model. More generally, the representation of the population could be less specific, such as statistics related to impacts from natural disasters (casualties, injuries, property losses and agricultural losses — all of which an emergency manager might wish to estimate).

NOTE 4 Consideration of descriptive statistics could suggest that an assumed model provides an inadequate representation of the data, such as indicated by a measure of the goodness of fit of the model to the data. In such cases, other models could be considered and the estimation process continued.

- Statistical process control (SPC): Activities focused on the use of statistical techniques to reduce variation, increase knowledge about the process and steer the process in the desired way [ISO 3534-2:2006].

NOTE 1 SPC operates most efficiently by controlling variation of a process characteristic or an in-process product characteristic that is correlated with a final product characteristic and/or by increasing the robustness of the process against this variation. A supplier's final product characteristic can be a process characteristic to the next downstream supplier's process.

NOTE 2 Although SPC originally was concerned primarily with manufactured goods, it is also equally applicable to processes producing services or transactions, for example, those involving data, software, communications and movement of material.

NOTE 3 SPC involves both process control and process improvement.

ABBREVIATIONS

The following abbreviated terms are used in this document:

- ASCE: American Society of Civil Engineers.
- ASPRS: American Society for Photogrammetry and Remote Sensing.
- DQU: Data quality unit.
- ADS: Assessed Data Set
- EMAS: Engineering map accuracy standard.
- FGDC: Federal Geographic Data Committee.
- GNSS: Global Navigation Satellite System.
- PAIGH: Pan-American Institute of Geography and History.
- ISO: International Organization for Standardization.

- MSE: Mean squared error
- NMAS: National Map Accuracy Standard.
- NSSDA: National Standard for Spatial Data Accuracy.
- PAAM: Positional accuracy assessment method.
- RDS: Reference data set.
- SDI: Spatial data infrastructure.
- SDS: Spatial data set.
- TAP: Technical Assistance Program of the IPGH.

SPATIAL DATA QUALITY BASED ON ISO/TC 211 STANDARDS

Under the normative framework established by the 19100 standards of the ISO Technical Committee 211, any quality assessment must adequately reconcile aspects related to product specifications, the formalization of quality aspects, their assessment, and the reporting of results. This is intended to ensure that producers and users work under the same criteria and concepts with the aim of achieving interoperability and consensus when generating data, assessing its quality, understanding the assessment process, and, finally, visualizing and interpreting the results presented as metadata. Therefore, the international standards directly related to quality and its assessment are ISO 19157 (ISO 2013), ISO 19131 (ISO 2007), and ISO 19115-1 (ISO 2014). These standards are outlined below.

ISO 19157[6]. This standard establishes a framework to: i) identify the relevant aspects of spatial data quality through quality elements (i.e., which aspects to assess?); ii) link quality elements to the data sets of interest (scope) through the data quality unit (DQU) (i.e., which geographic objects to assess?); iii) develop well-described assessment methods (i.e., how to assess?); iv) use well-determined measures (i.e., how to express quality?); v) adequately express the result (i.e., what is the outcome of the quality assessment?); and vi) appropriately report the results (i.e., how to report?). Figure 1 depicts the relationships between all these aspects.

---

[6] ISO 19157 is currently in the process of being revised.

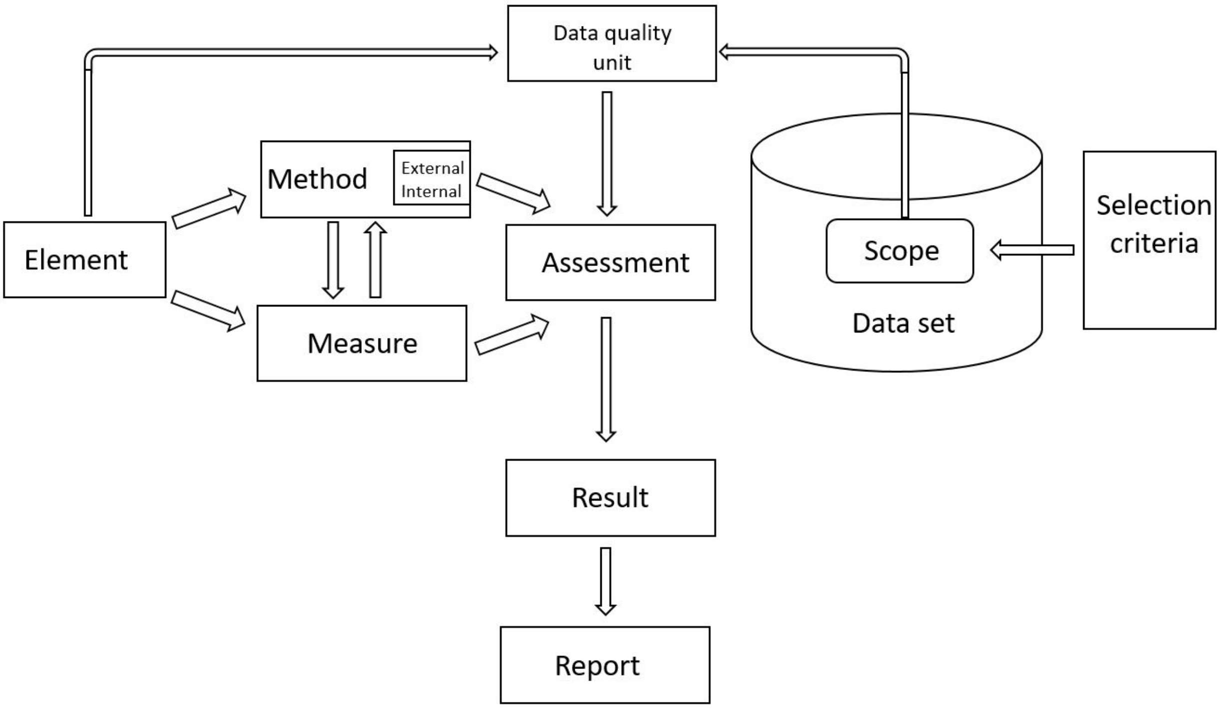


Figure 1. Relationships and intervening elements in a data quality assessment (Source: Ariza-López, 2017)

The main objectives of this standard are to:

- Provide principles for describing the quality of geographic data and concepts for managing quality information of a data set.
- Provide guidance for quantitative data quality assessment procedures.
- Provide a mechanism for reporting on data quality (metadata).
- Provide a set of data quality measures for use in data quality assessment and reporting.

ISO 19131. This standard establishes guidelines for the development of data product specifications. A data product is the abstraction of all data sets generated as a result of a well-established production process. The specifications define the theoretical, ideal, or intended data set, that is, the nominal domain, which, if it is a series, does not vary with each version of the product or with each unit. These specifications inform both the producer and user of the intended qualities of the data, such as the reference system, the resolution/scale of the data, the data model, the acquisition processes, and the quality aspects of interest and their levels (e.g., maximum percentage of omissions). The contents of the specifications should be logically adjusted to the technical capabilities of the processes (the voice of the process) and the use requirements (the voice of the user). Therefore, the application of the standard to a specific data product implies establishing explicit levels of conformity for each quality element under consideration. As such, the specifications in ISO 19131 regarding the quality aspects of data products are based on the ISO 19157 standard, and, therefore, the specifications of any product conforming to ISO 19131 must adequately cover the concepts of ISO 19157, and, in addition, establish the quality requirements, clearly indicating all the related aspects (Figure 1). The assessment of the quality of a product is determined by its specifications (i.e., established quality levels, scope, quality measures, and assessment methods to be applied).

ISO 19115-1. This standard provides a metadata model for spatial data that includes information on use, purpose, and lineage, and is relevant to understand the quality of an SDS. Metadata conforming to ISO 19115-1 will add zero, one, or more DQUs (ISO 19157) to report the quality of an SDS. In addition, these metadata can be complemented with an standalone quality report proposed in the international standard ISO 19157.

Finally, the technical specification ISO/TS 19158 (ISO 2012) should be mentioned. This provides a quality assurance framework for the production of spatial data based on the principles of quality management systems (e.g., ISO 9000 (ISO 2015b)) and the aforementioned ISO 19157.

THE ISO 19157 STANDARD FOR POSITIONAL ACCURACY ASSESSMENT

The International Standard ISO 19157 (ISO 2013) focuses on what are termed quantitative elements of spatial data quality, that is, those that can be expressed in a numerical (quantitative) quantity. The quality elements are components that describe certain aspects of spatial data quality and have been organized into categories, such as completeness, logical consistency, and temporal quality. These are described below with an indication of the data quality elements they include.

Usability[7]. This category is based on the combination of the quality elements listed below to better express the user requirements.

Completeness. This refers to being complete (completeness), and considers the following quality elements: omission (missing elements) and commission (excess elements).

Logical consistency. This refers to compliance with explicit rules of the logical models with which a spatial data set works, and considers the following quality elements: conceptual consistency, domain consistency, format consistency, and topological consistency.

Temporal quality. This refers to the temporal aspects of spatial data, and considers the following quality elements: time measurement accuracy, temporal consistency, and temporal validity.

Thematic accuracy[8]. This refers to the thematic aspects of spatial data, and considers the following quality elements: classification correctness, non-quantitative (qualitative) attribute correctness, and quantitative attribute accuracy.

Positional accuracy. This refers to the accuracy of the position of spatial data in a reference system, and considers the following quality elements:

- Absolute or external positional accuracy: proximity of the collected coordinate values to the true or accepted values.
- Relative or internal positional accuracy: proximity of the relative positions of the geographic objects in a data set to their respective true or accepted relative positions.
- Positional accuracy of grid data: proximity of the position values of the data in a regular grid structure to the true or accepted values.

In addition to the categories and quality elements already presented, ISO 19157 provides other relevant contributions to standardize definitions and quality assessment processes, which is of great practical importance. Specifically, ISO 19157 proposes the use of standard quality measures and a general process for assessing spatial data quality, which are outlined below.

---

[7] The September 2020 draft of ISO 19157-1 proposes that this quality element is no longer considered.

[8] The September 2020 draft of ISO 19157-1 proposes this element to be called 'thematic quality'.

Quality measures. ISO 19157 defines a set of basic measures to account for the presence of errors or to quantify their magnitude, depending on the type of error. From these basic measures, a broad set of measures is defined that can be applied to specific quality elements. This set of standard measures linked to the quality elements are presented in Annex D of the standard, and are freely available for use. Table 1 presents the set of measures defined for positional accuracy, indicating the identifier (ID), the name, the quality element to which they apply, and the basic measure on which each of them is based.

Table 1. Quality measures proposed by ISO 19157 for positional accuracy (Source: ISO 19157)

ID	Name	Element	Basic measure
28	Mean value of positional uncertainties (1D, 2D, and 3D)	Absolute or external	Not applicable
128	Bias of positions (1D, 2D, and 3D)	Absolute or external	Not applicable
29	Mean value of positional uncertainties excluding outliers (2D)	Absolute or external	Not applicable
30	Number of positional uncertainties above a given threshold	Absolute or external	Error count
31	Rate of positional uncertainties above a given threshold	Absolute or external	Not applicable
32	Covariance matrix	Absolute or external	Not applicable
33	Linear error probable	Absolute or external	LEP50.0 or LE50.0(r)
34	Standard linear error	Absolute or external	LE68.3 or LE68.3(r)
35	Linear map accuracy at 90% significance level	Absolute or external	LE90 or LE90(r)
36	Linear map accuracy at 95% significance level	Absolute or external	LE95 or LE95(r)
37	Linear map accuracy at 99% significance level	Absolute or external	LE99 or LE99(r)
38	Near certainty linear error	Absolute or external	LE99.8 or LE99.8(r)
39	Root mean square error	Absolute or external	Not applicable
40	Absolute linear error at 90% significance level of biased vertical data (alternative 1)	Absolute or external	Not applicable
41	Absolute linear error at 90% significance level of biased vertical data (alternative 2)	Absolute or external	Not applicable
42	Circular standard deviation	Absolute or external	CE39.4
43	Circular error probable	Absolute or external	CE50
44	Circular error at 90% significance level	Absolute or external	CE90
45	Circular error at 95% significance level	Absolute or external	CE95
46	Circular near certainty error	Absolute or external	CE99.8
47	Root mean square error of planimetry	Absolute or external	Not applicable
48	Absolute circular error at 90% significance level of biased data	Absolute or external	Not applicable
49	Absolute circular error at 90% significance level of biased data	Absolute or external	Not applicable
50	Uncertainty ellipse	Absolute or external	Not applicable
51	Confidence ellipse	Absolute or external	Not applicable
52	Relative vertical error	Relative or internal	Not applicable
53	Relative horizontal error	Relative or internal	Not applicable

The specifications of any data product should use measures defined in ISO 19157. The purpose of using standard measures is to provide transparency in quality specification and assessment processes, and provide greater confidence to users and producers as well as greater comparability of results and interoperability of data and results. For this, as already indicated, the measures established in ISO 19157 can be used but it is also possible to develop other measures following the guidelines provided in the standard for the standardized descriptions. In any case, such measures must always be identified and included in a catalog of measures that must be publicly available and organized according to the guidelines of ISO 19135-1 (ISO 2015a).

General spatial data quality assessment process. If a data product is well specified in terms of its quality aspects, the product specifications will provide all the key elements, such as the quality elements, measures, and methods for its quality assessment. Any quality assessment involves a general logical process, which is included in the product specifications or may be established subsequently for a new assessment requirement. Figure 2 presents the flow chart included in ISO 19157 as a general quality assessment process. If the data to be assessed are heterogeneous, with different qualities in different parts, such assessments will be carried out on each of those parts. Table 2 specifies the steps of this process in greater detail. When reflecting on Table 2, it can be concluded that the most difficult step is likely the 3rd step, in which the quality assessment method must be specified. For example, if one intends to establish the method for assessing absolute positional accuracy, one should consider that there are several positional accuracy standards (e.g., NMAS, EMAS, and NSSDA), each with different statistical characteristics. Moreover, it should be considered that the applied sampling method (e.g., number and distribution of control points) has a great influence on the result. It is also necessary to specify the fieldwork methodology for field data collection (e.g., topographic methods, GNSS methods, etc.). For an assessment to be well specified, all these aspects should be adequately documented and applied in the execution. Despite the apparent complexity of this, positional accuracy assessment is not the least favorable option as there are numerous standards in fairly widespread application and it is a component that works with well-known methods (e.g., topographic, GNSS, etc.). However, the same is not true for the assessment of other quality elements, such as completeness and thematic accuracy.

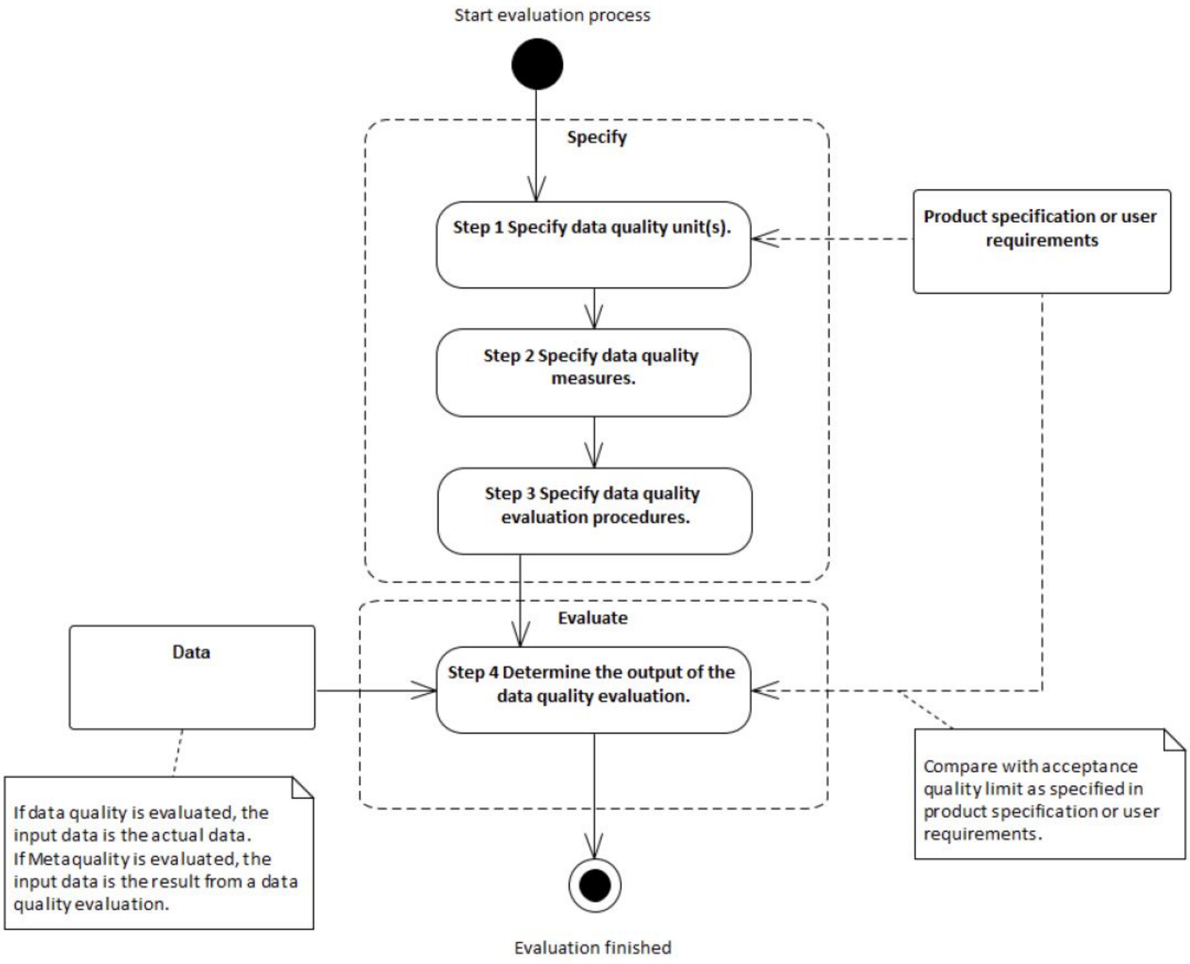


Figure 2. UML diagram of data quality evaluation (Source: ISO 19157)

Table 2. Steps of quality evaluation process according to ISO 19157 (Source: ISO 19157)

Step of the process	Action	Description
1	Specify the data quality unit(s) (DQU)	A DQU is composed of a specific scope and data quality element(s). All elements relevant to the data for which quality is to be described should be used.
2	Specify the data quality measures	If applicable, a measure should be specified for each data quality element. If no measure is identified, a descriptive result can be provided.
3	Specify the data quality assessment procedures	A data quality assessment procedure consists of the application of one or more assessment methods.
4	Determine the output of the data quality assessment	The output of the assessment process is a result.

A key aspect when applying spatial data quality standards is to be clear about the difference between measurement and method (assessment procedure). Both elements (measure and method) interact but are different; the same measure can be used by several methods, and the same general method could be customized using different measures if these require modifying the method. Figure 3 presents a diagram with the key defining aspects. As already mentioned, ISO 19157 provides a set of standard measures and also a process for creating new measures; however, methods are rather neglected as no standardized options are provided nor a basis for standardizing. Indeed, given that it is advocated that organizations have a catalog of standard quality measures, a catalog of standard quality assessment methods should also be advocated. To assist with this, this guide should help better define assessment methods.

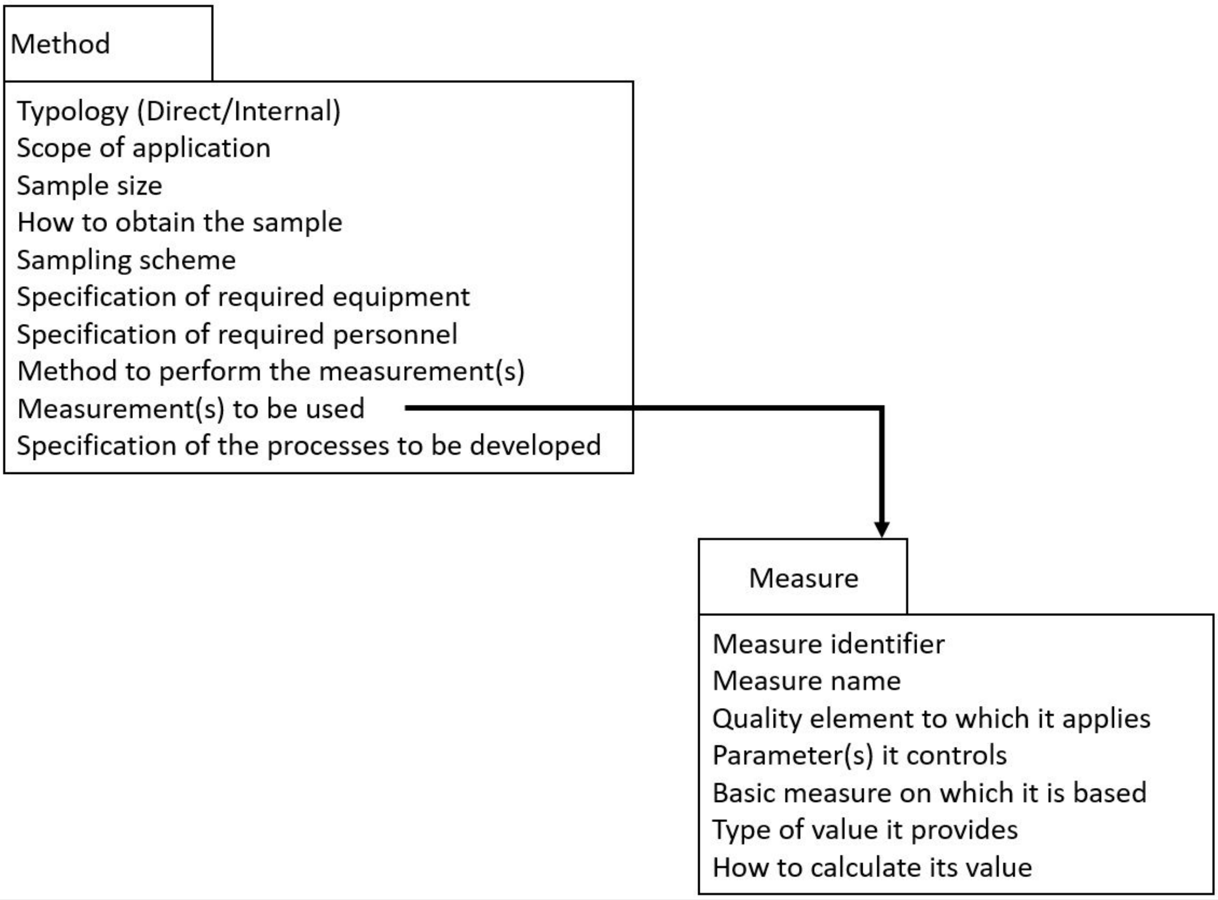


Figure 3. Relationship between methods and measures (Source: Ariza-López, 2017)

POSITIONAL ACCURACY AND ERROR

As previously indicated, positional accuracy refers to the accuracy of the position of spatial data in a reference system, that is, the proximity of the reported coordinates of some geographical objects(s) in the data product (e.g., the corner of a house in a data set) to the true or accepted values (e.g., the coordinates of the corner of that house based on GNSS data acquisition). This 'proximity between values' has traditionally been termed error, although it would be more appropriate to call it discrepancy.

It should be noted that the term 'error' is not entirely appropriate but is used as common practice in cartography and geosciences, with the justification that for a value V_p corresponding to an estimated quantitative characteristic, if its associated error (E_p) is known, V_p can be corrected (by adding or subtracting E_p). Theoretically, an error-free value of V could, therefore, be obtained. However, this is not typically the case. Therefore, the term 'discrepancy' is more appropriate than 'error', which itself should be more appropriately termed 'uncertainty' when used in a statistical estimation framework.

Positional error is usually considered to refer to a specific point with some coordinates and can have independent values for each of the X, Y, and Z components ($\{e_x\}$, $\{e_y\}$, and $\{e_z\}$) (1D case). It can also be measured and expressed planimetrically or radially with two ($\{e_x$ and $e_y\}$) (2D case) or, very rarely, three ($\{e_x$, e_y, and $e_z\}$) (3D case) combined components. For a quantitative characteristic of interest (i.e., the position), the error is the difference between an estimated value (observed) and present in the data set to be assessed (assessed data set, ADS)[9] and the true or reference value (the reference data set, RDS[10]). Equation 1 presents the mathematical definition[11] of the positional error in each of the components X, Y, and Z.

$$e_X = X_{DSA} - X_{RDS} \qquad e_Y = Y_{DSA} - Y_{RDS} \qquad e_Z = Z_{DSA} - Z_{RDS} \qquad \text{Equation 1}$$

where:

e	error (discrepancy)
X	coordinate X
Y	coordinate Y
Z	coordinate Z

The typology of objects that are commonly used in positional accuracy assessments to obtain these error or discrepancy values are well-defined points[12], that is, those which geometric definition facilitates the unequivocal identification of the point (e.g., there is no dilution of geometric precision when two lines cross). They are also isolated points (e.g., corners or intersections, generally of human-made elements) that are easily identifiable in the context of their neighborhood, both in the ADS and RDS. In addition, both ADS and RDS coordinates must be expressed in the same coordinate reference system. Finally, to ensure that possible biases are detected and the accuracy of the RDS coordinates does not significantly affect the intended assessment, the RDS source must be independent and, at the same time, of higher accuracy[13].

The errors involved are classified into the following categories:

- **Gross errors.** This term is common in Geomatics but it is in fact erroneous. Measurement error should not be confused with production or human error (JCGM 2012). These are mistakes, and must be eliminated prior to any statistical analysis. In many cases these types of mistakes derive from data transcriptions (e.g., when copying numerical figures, 75 can be miss-transcribed as 57). All work methods must be designed to eliminate such mistakes.
- **Systematic errors.** These are errors that, either in a constant manner (e.g., in time or space) or following a specific function, affect the measured values. A simple example is that of an incorrectly sized tape measure with centimeter divisions. Therefore, assuming that it is not affected by any external parameter (e.g., temperature or humidity), all measurements are affected by a systematic error as a consequence of applying an erroneous measurement standard. Such systematic errors can be of a constant type (as in the tape measure example) and affect the entire set of measurements by the same value, or of a functional type, whereby the generated errors are constant but only affect a subset of the measurements. The latter category includes, for example, terrain effects and local distortions. Systematic errors can be detected statistically and can be corrected by appropriate methods.
- **Random errors**. These are errors that occur randomly merely by producing data and performing measurements. They typically follow a statistical model, but the model does not

---

[9] Hereinafter, the 'assessed data set' (DSA) or 'product to be assessed' is used equivalently to refer to the data set for which a positional accuracy parameter is to be estimated or controlled.

[10] Hereinafter, 'reference data set' (RDS) or 'reference' will be used equivalently.

[11] In principle, it makes no difference whether the product coordinates are subtracted from the reference coordinates or *vice versa*. This does not invalidate the results of the analysis. It is only necessary to know what is subtracted from what to be clear about the real direction of the quantity (missing or excess).

[12] However, linear elements can also be used for positional accuracy controls (see Ariza-López 2013).

[13] The PAAMs generally indicate that accuracy should be at least three-times higher. A higher accuracy implies closer proximity between measured values and true values, that is, smaller discrepancies and thus lower uncertainty.

have to be unique or common—it can be a mixture of models. The base model is generally the normal model, that is, a Gaussian distribution. Random errors must be within an acceptable variability range for the use of the product. All production processes generate such errors. Therefore, production processes (methods, technologies, and organization schemes) must be suitably designed so that these errors do not prove inconvenient. The extent of random errors can be statistically assessed. A well-designed production process is one that is capable, that is, can manage production in such a way as to minimize the production of these random errors.

In addition to these error types, it is important to recognize and properly define the concepts of 'outliers' and 'accuracy' to have good understanding of the situations that can occur when dealing with positional accuracy and developing quality assessment analyses and reports.

Outlier. Statistical outliers should never be confused with gross errors. An outlier is an extremely high or low value with respect to the data set (or population) to which it belongs, indicating that it has a low probability of occurrence (atypicality). As an example, in the case of the height of individuals, an outlier is a value that could correspond to a basketball player with a height of 2.32 m; whilst this is a person with full rights and obligations, such a value within a small sample collected to estimate the height of the population can distort the overall analysis. Therefore, although outliers are values that truly belong to the population and should, therefore, be included in statistical reports, they are generally not included in the calculations of parameters such as the mean and standard deviation where the sampled population displays a normal distribution[14]. It should be emphasized that outliers are not gross errors, and the latter should be eliminated before statistical analysis. Outliers should be analyzed by a particular analysis process[15] and appropriately reflected in result reporting.

Accuracy. Accuracy is considered at the data set level. ISO 5725-1 (ISO 1994) uses two terms, 'trueness' and 'precision', to describe the accuracy[16] of a measurement method. 'Trueness' refers to how close the arithmetic mean of a large number of test results is to the true value or the accepted reference value. 'Precision' refers to the proximity between different results and has thus no relation to the true value or accepted reference value.

$$Accuracy = Trueness + Precision \qquad \text{Equation 2}$$

Trueness. According to the definition of accuracy, trueness is equivalent to a mean bias value. For example, individual data can have positional discrepancies (errors), whereas data sets can have bias. Bias, which is typically indicated in terms of a direction (i.e., left/right, up/down, etc.) is linked to the presence of systematic errors that may or may not be assigned a cause. If a cause can be assigned, it can be eliminated.

Precision. Precision is considered at the data set level. For a quantitative characteristic of interest, precision is the degree of mutual proximity in a set of repeated measurements that are considered independent. This proximity is generally measured by variability or dispersion in the measurements, and is quantified by the standard deviation ($\pm\sigma$), which establishes an interval of

---

[14] This situation could be adequately managed with so-called robust statistics, but their application is not common for PAAMs. The median and mode are parameters that are less affected by the presence of outliers. The median absolute deviation (MAD) is a substitute for the variance in this circumstance.

[15] Outliers are of great interest for the improvement of processes. If their presence is high, it may indicate the mixture of several juxtaposed processes in the process under analysis. On the other hand, the presence of extremely high/low values allows the detection of processes that perform better/worse than the expected process, thus leading to improvement.

[16] The International Standard ISO 5725-1 indicates that the term 'accuracy' has been used in the past only to describe the component now called 'trueness', but many specialists consider that it should refer to the total displacement of a result with respect to its reference value, due to both random and systematic effects.

possible values [-σ, +σ]. Precision is linked to the variability of the processes involved; all processes have some variability, which indicates that this can be reduced by appropriate decisions (e.g., changing processes) but never be completely eliminated. The precision of specific data cannot be improved; if it is not sufficient, it should be created again using methods that reduce this issue. When characterizing an acquisition device (e.g., a digitizing tablet, GNSS equipment, etc.), it is common for the manufacturer to provide a value for precision and not for accuracy. This implies that it is assumed that there are no biases in the process from which the results are obtained (e.g., there is an adequate calibration process that eliminates the bias).

Figure 4 shows examples of the general cases that can occur with regard to the relationship between trueness (presence or absence of bias) and precision (more precise or less precise data).

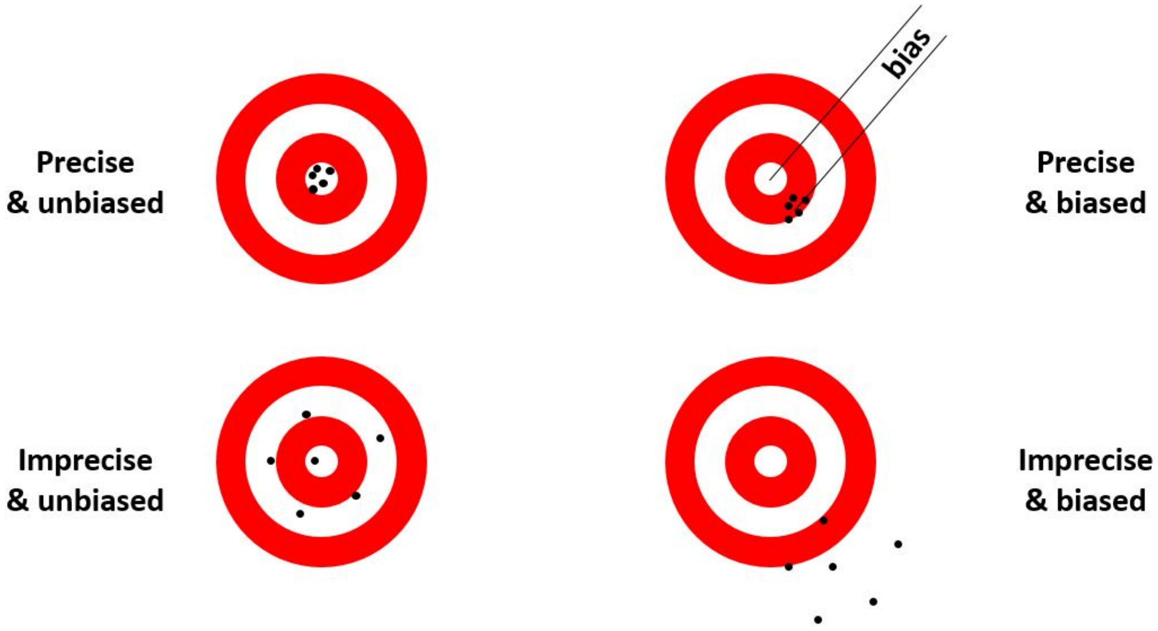


Figure 4. Relationship between trueness and accuracy

POSITIONAL ACCURACY STANDARDS

Given the importance of positional accuracy in cartographic products, its assessment has been of interest since the development of science-based cartography. Therefore, several methods have been proposed and applied for this purpose throughout the 20th century (e.g., see Maling 1989; Ariza-López and Atkinson Gordo 2008a). These methods are referred to here as 'standards' as many of them are so named (e.g., the National Map Accuracy **Standard**, Engineering Map Accuracy **Standard**, and National **Standard** for Spatial Data Accuracy). These are standards and not norms, since almost none of them were developed by a standardization organization, and derive from professional associations or administrative bodies in the cartographic sector. The exception is UNE 148002, which bases its section on statistics on two other international standards of ISO.

As it is impossible to assess the positional component of 100% of the objects of a data product, a sample is also worked on and, therefore, the statistical method applied should be robust and reliable in all matters related to sampling to ensure the representativeness and validity of the results. This is related to the possibility of applying assessment methods focused either on quality estimation or quality control:

- **Estimation methods.** These are intended to reliably establish the value of a quality parameter (e.g., mean bias, standard deviation, proportion, etc.) regarding the population of interest. These methods provide a value and its corresponding confidence interval as a result (e.g., a mean value and its deviation such as 5.27m ± 0.15 m). The sample size is linked to the population size, and a larger population size requires a larger sample size, although there is an asymptotic behavior in the latter with respect to population size.
- **Control methods.** These are intended to provide a statistical basis for making an acceptance/rejection decision as a consequence of compliance/non-compliance with a specification. For example, given the specification that no more than 5% of the elements present 2D positional errors greater than 1 m, a decision is made to accept/reject according to the evidence found in the sample, trying to minimize type I (producer's risk, that is, to reject a data set that meets the specifications) and type II (user's risk, that is, to accept a data set that does not meet the specifications) errors. In this case, hypothesis testing techniques are applied and, therefore, the sample size is not directly linked to the population size. A random sample size ensures type I errors are adequately controlled, although a specific sample size is required to ensure that type II errors are avoided.

Both the sample size and its distribution (spatial, thematic, and temporal) and randomness are key elements to ensure the representativeness and validity of the results. In addition, from a statistical point of view, there are two types of methods:

- **Parametric based.** These methods assume that the positional errors under analysis conform to a parametric statistical distribution function, which is known. It is common to assume the normal distribution, in which case the mean (location parameter) and the standard deviation (scale parameter) are applied. These are currently the most frequently used methods (e.g., NSSDA).
- **Non-parametric based.** These methods do not require the positional error to fit an underlying parametric statistical distribution function (i.e., a non-normal distribution). The error distribution is given by the observed data[17]. The non-parametric approach is appropriate when the data cannot be assumed to fit a known distribution. These methods are based on percentiles or proportions, and their application is common in the case of LiDAR altimetry data of vegetated land. The PAAM based on error counting from tolerances can also be applied in this case (e.g., NMAS).

Four standards for positional accuracy assessment are presented below, providing different perspectives on estimation/control and parametric/non-parametric-based methods. In each case, a brief introduction and summary table are provided with a focus on their respective origin, comparison method, positional components on which the calculation can be performed (e.g., X, Y, and Z), existence of a standard value, a brief description of the method, and the reference to the original source.

NATIONAL MAP ACCURACY STANDARDS (NMAS)

The NMAS standard (USBB 1947) has been used by US mapping agencies since 1947. This has resulted in the extension of its use and application by numerous institutions and official bodies in many other countries. The method proposed by the PAIGH (IPHG 1978) is to some extent based on the NMAS. The standard sets out a method of positional accuracy control that establishes an

---

[17] This is a current trend in the case of 'big data', and there is no need for a model with such an abundance of data; the data itself are the best model.

acceptance/rejection rule in a very simple manner. Therefore, its statistical basis is hypothesis testing. As a control method, excessively large sample sizes are not required. Ariza-López and Rodríguez-Avi (2014) indicate that this approach adopts a binomial-based model and is thus based on counting errors. Therefore, the NMAS does not require underlying normality. If a normal model is assumed, then tolerances can be established based on this distribution, which allows a comparison[18] of the results with other methods based on the normal distribution function. This standard is outdated, however, as it refers to tolerances defined on paper, that is, to the representation scale, but its conceptual basis can be applied to any tolerance value. Moreover, the NMAS is directly related to the modern lot control method proposed by the UNE 148002 (UNE 2016) Standard, which is also based on counting. As an extension, a method was recently proposed (Ariza-López and Rodriguez-Avi 2018) that allows the control of positional accuracy through two or more tolerances.

The NMAS method is very briefly explained in just one page, so it is not explicit in many aspects. The greatest advantage of the NMAS standard is its simplicity regarding calculations and ease of understanding, as the results are expressed as compliant/non-compliant, which can be easily interpreted by the user. The disadvantages of using this standard relate more to the producer, as it does not provide much insight on the statistical behavior of errors (e.g., bias and deviation for the normal case), and no information is obtained on how the effects of the production processes, making it difficult to improve them. Furthermore, the standard provides slightly permissive test, as the stated tolerances are wide. This may be explained by the age of the standard, as information acquisition methods were much less developed than today. Table 3 presents a description of this standard.

Table 3. NMAS standard

Comparison method	With sources having higher accuracy.
Positional component	Horizontal and vertical, separately.
Element type	Well-defined points.
Accuracy standard	The standard proposes accuracies in relation to the map scales by means of tolerances. Horizontal accuracy: • H_{Tol1} = 1/30 of an inch for maps at publication scales greater than E20k[19]. • H_{Tol2} = 1/50 of an inch for maps at publication scales of E20k or smaller. Vertical accuracy: • V_{Tol} = One-half of the interval between contour lines for all publication scales. When verifying elevations, the apparent vertical error can be decreased by assuming a horizontal offset within the acceptable horizontal error for a map of that scale. **Report**: The products complying with these requirements shall note this fact in their legends as follows: «*this map complies with NMAS*»
Description	Product (*p*) is compared to a reference with higher accuracy. The horizontal and vertical components can be assessed separately. The vertical assessment is subject to the horizontal assessment. A sample of *n* well-defined control points is used (without specifying the value of *n*). Horizontal accuracy: no more than 10% of the tested points shall have an error greater than Tol1 or Tol2 (depending on the scale). Vertical accuracy: no more than 10% of the tested points shall have an error greater than V_{Tol}.
Procedure	1. Select a sample. 2. Calculate the error of each point in each component: $e_{x_i} = x_{p_i} - x_i \quad e_{y_i} = y_{p_i} - y_i \quad e_{z_i} = z_{p_i} - z_i$ where:

---

[18] This is considered to analyze the relationship between this method and others, such as the NSSDA.

[19] The scale is denoted by the notation 'ExxxK', where E stands for scale, xxxx is the denominator of the scale, and k represents thousands. For example, a scale of 1/10,000 is denoted 'E10k'.

	x_i, y_i, z_i are the coordinates in the reference (RDS). $x_{p_i}, y_{p_i}, z_{p_i}$ are the coordinates in the product (ADS). 3. Calculate the horizontal component of the errors in x, y at each point: $e_{H_i} = \sqrt{e_{X_i}^2 + e_{Y_i}^2}$ 4. Establish which are the maximum tolerable errors: Horizontal: H_{Tol1} = 0.085 cm (1/30") in maps of scale greater than E20K or H_{Tol2} = 0.05 cm (1/50") in maps at scale smaller or equal to E20K. Vertical: Half of the equidistance (interval) between contour lines (V_{Tol}). 5. Count how many points have a horizontal error e_H greater than the tolerance that applies to the scale case. The control is surpassed in the horizontal component if the number of points having an error above the tolerance does not exceed 10% of the cases. 6. Count how many points have a vertical error e_z greater than the vertical tolerance. The control is surpassed in the vertical component if the number of points that have an error above the tolerance does not exceed 10% of the cases.
Source	USBB (1947). *United States National Map Accuracy Standards*. U.S. Bureau of the Budget. Washington, USA.

ENGINEERING MAP ACCURACY STANDARD (EMAS)

The EMAS standard (ASCE 1983) was developed by the American Society of Civil Engineers (ASCE) during the American Congress on Surveying and Mapping and emerged as a response to dissatisfaction with the existing assessment methods at the time. Its dissemination as the EMAS is largely restricted to the United States, but it proposes a logical process from a statistical perspective, enabling more widespread application (e.g., Seville, 1991, proposes something similar in Spain). The EMAS specifies the accuracy of large-scale topographic maps and also establishes a statistical procedure for positional accuracy control. It refers to tolerances defined on paper, that is, to the representation scale, but its conceptual basis can be applied to any tolerance value.

As the standard proposed a control method, excessively large sample sizes are not required. It assumes that positional errors are normally distributed (parametric model) and proposes a set of statistical hypothesis tests that must be overcome for the product to be accepted. Specifically, it establishes two statistical tests per component, one focused on the detection of biases (Student's t-test) and the other on the behavior of dispersion (Chi-square test). As several hypothesis tests should be addressed (one test on systematisms and another on the dispersion in each component), the method is rather restrictive, which can cause problems for both the producer and the user. In the former case, it is problematic that a large number of correct data sets can be rejected, and in the latter case, a high rejection rate can result from the administrative, temporal, and economic consequences of a deficient supply. As an example, in an exclusively planimetric control, under the assumption of independence of the X and Y components, and a significance level of $\alpha = 5\%$, given that four hypothesis tests are performed, only 81.5% of the cases will pass together (0.95 × 0.95 × 0.95 × 0.95 = 0.815). This situation has been analyzed by Ariza-López et al. (2008), who propose the solution of applying Bonferroni correction to limit type I error to the desired significance ($\alpha = 5\%$ overall). The standard provides adequate results for the producer as it informs, in detail, what happens in the case of each component, and, therefore, the producer can take improvement actions if necessary. It may not, however, be relevant for the end-user to have so much information. Table 4 presents a description of this standard.

Table 4. EMAS standard

Comparison method	With sources having higher accuracy.
Positional component	Each component (X, Y, and Z) separately.
Element type	Well-defined points. Well-distributed: separation between points in the range [1/12, 1/4] of the diagonal dimension of the map coverage. At least 15% of the points in each quadrant.

Accuracy standard	The standard proposes accuracies in relation to map scales (see Table 1 of ASCE 1983). **Report**: The products complying with these accuracy requirements shall note this fact in their legends as follows: *« This map complies with the EMAS for a scale____ with error limits not exceeding____ meters.* *Type of error* *X* *Y* *Z* σ_0 __ __ __ $\lvert \delta_0 \rvert$ __ __ __»
Description	The product (*p*) is compared with a reference with higher accuracy. The X, Y, and Z components can be assessed separately. A sample of at least 20 well-defined and well-distributed control points is used. A mean absolute error ($\lvert \delta_0 \rvert$) limit in a Student's t-test is used for bias. A standard deviation (σ_0) limit is taken as the null hypothesis of the Chi-square test for dispersion.
Procedure	1. Select a sample of n points, where $n \geq 20$. 2. Calculate the error for each point in each component: $e_{x_i} = x_{p_i} - x_i \qquad e_{y_i} = y_{p_i} - y_i \qquad e_{z_i} = z_{p_i} - z_i$ where: x_i, y_i, z_i are the coordinates in the reference (RDS). $x_{p_i}, y_{p_i}, z_{p_i}$ are the coordinates in the product (ADS). 3. Calculate the mean error of each component: $\bar{e}_x = \frac{1}{n}\sum_{i=1}^{n} e_{x_i}; \quad \bar{e}_y = \frac{1}{n}\sum_{i=1}^{n} e_{y_i}; \quad \bar{e}_z = \frac{1}{n}\sum_{i=1}^{n} e_{z_i}$ 4. Calculate the sampling standard deviation in each component: $S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(e_{x_i} - \bar{e}_x\right)^2}; \quad S_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(e_{y_i} - \bar{e}_y\right)^2}; \quad S_z = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(e_{z_i} - \bar{e}_z\right)^2}$ 5. Perform, for each component, the standard compliance test to determine if the mean error is acceptable (which implies absence of bias). For this, a test is performed on the mean, under the assumption of unknown population variance, establishing the following hypotheses: $H_0: \mu = 0 \qquad H_1: \mu \neq 0$ The map will pass the test with a significance level α if the following is met: $\lvert t_x \rvert \leq t_{n-1,\frac{\alpha}{2}}; \quad \lvert t_x \rvert \leq t_{n-1,\frac{\alpha}{2}}; \quad \lvert t_z \rvert \leq t_{n-1,\frac{\alpha}{2}}$ where: $t_{n-1,\alpha/2}$ Student's t-distribution value, with $n-1$ degrees of freedom. t_x, t_y, t_z Result of calculating the following statistics: $t_x = \frac{\bar{e}_x\sqrt{n}}{S_x}; \quad t_y = \frac{\bar{e}_y\sqrt{n}}{S_y}; \quad t_z = \frac{\bar{e}_z\sqrt{n}}{S_z}$ 6. Perform, for each component, the standard compliance test to determine if the sample standard deviation is within acceptable limits. For this purpose, a test is performed on the variance, establishing the following hypotheses in relation to a maximum variance value $\sigma_{0x}^2, \sigma_{0y}^2$ and σ_{0z}^2 pre-established and specified on each component: $H_0: \sigma^2 \leq \sigma_0^2; \quad H_1: \sigma^2 > \sigma_0^2$ The product will pass the control with a significance level α if the following is met: $\chi_x^2 \leq \chi_{n-1,\alpha}; \quad \chi_y^2 \leq \chi_{n-1,\alpha}; \quad \chi_z^2 \leq \chi_{n-1,\alpha}$ where: $\chi_{n-1,\alpha}$ Theoretical value of the Chi-square distribution, with *n-1* degrees of freedom. $\chi_x^2; \chi_y^2; \chi_z^2$ Result of calculating the following statistics: $\chi_x^2 = \frac{S_x^2(n-1)}{\sigma_{0x}^2}; \quad \chi_y^2 = \frac{S_y^2(n-1)}{\sigma_{0y}^2}; \quad \chi_z^2 = \frac{S_z^2(n-1)}{\sigma_{0z}^2}$
Source	ASCE (1983). *Map Uses, scales and accuracies for engineering and associated purposes.* American Society of Civil Engineers, Committee on Cartographic Surveying, Surveying and Mapping Division, New York, USA.

National Standard for Spatial Data Accuracy (NSSDA)

The NSSDA standard (FGDC 1998) was proposed by the Federal Geographic Data Committee to replace the NMAS and ASPRS standards (ASPRS 1990) due to the rise of digital data in the 1990s, becoming mandatory for all US federal bodies involved in cartographic production tasks. In addition, this method is the basis of the new ASPRS Accuracy Standards for Digital Geospatial Data (ASPRS 2015) in all matters related to the determination of positional accuracy when positional errors follow a normal distribution model. Consequently, this standard is widely applied across the world.

The NSSDA is based on the assumption of normality of the error data (parametric basis) but does not provide guidelines for the verification of this assumption. The NSSDA is not a positional accuracy control method as it does not establish acceptance or rejection; the result is a value and, therefore, is an estimation method. The NSSDA performs a unique estimation as it does not link sample size to population size[20]. This method provides a positional accuracy index in real units in the field, but does not indicate if a data set is accepted or rejected, as is the case with the previously outlined methods. Instead, the user must interpret whether the value is satisfactory or not. On the other hand, the estimated value of positional accuracy is not complemented by the uncertainty of the estimate, and so has limited statistical value. In this regard, according to Ariza-López and Atkinson (2008b), the NSSDA underestimates the real value and the uncertainty in its estimates is highly influenced by sample size. For example, in the case of 20 points, the NSSDA has a variability of ± 11%, and more than 100 control points are required to reduce the variability to ± 5%. Table 5 presents a description of this standard.

Table 5. NSSDA standard

Comparison method	With sources having higher accuracy.
Positional component	Horizontal and vertical, separately.
Element type	Well-defined points.
Accuracy standard	No tolerances, limiting values or similar restrictions. A result is provided, and the user has to consider whether or not it is fit for purpose. **Report**: The estimated accuracy at a confidence level of 95% is reported for both the horizontal and vertical cases. The product should include the legend: *« Tested __ meters horizontal accuracy at 95% confidence level* *Tested __ meters vertical accuracy at 95% confidence level %»*
Description	The product (*p*) is compared to a reference with higher accuracy. The horizontal and vertical components are assessed separately. A sample of at least 20 well-defined and well-distributed control points is used. The NSSDA uses the MSE to estimate positional accuracy. The accuracy is reported in ground distances at a confidence level of 95%. The systematic errors are assumed to have been eliminated as best as possible. The errors are assumed to be normally distributed and independent in each error component. Some level of heteroscedasticity is tolerated.
Procedure	1. Select a sample of n points, where $n \geq 20$. 2. Calculate the error for each point in each component: $e_{x_i} = x_{p_i} - x_i \qquad e_{y_i} = y_{p_i} - y_i \qquad e_{z_i} = z_{p_i} - z_i$ where: x_i, y_i, z_i are the coordinates in the reference (RDS). $x_{p_i}, y_{p_i}, z_{p_i}$ are the coordinates in the product (ADS). 3. Calculate the mean squared error of each component: $MSE_x = \sqrt{\frac{\sum e_{x_i}^2}{n}} \;;\; MSE_y = \sqrt{\frac{\sum e_{y_i}^2}{n}} \;;\; MSE_z = \sqrt{\frac{\sum e_{z_i}^2}{n}}$ 4. Obtain the horizontal NSSDA$_H$ value: if $MSE_x = MSE_y$ | $NSSDA_H = \frac{2.4477}{\sqrt{2}} MSE_r = 2.4477\, MSE_x$

---

[20] In a novel approach to positional accuracy assessment methods, the Positional Accuracy Standards for Digital Geospatial Data (ASPRS,2015) links the number of control points to the project surface area being controlled.

	where: $MSE_r = \sqrt{(MSE_x^2 + MSE_y^2)}$ if $MSE_x \neq MSE_y$ y $0.6 < \frac{MSE_{min}}{MSE_{max}} < 1.0$ $NSSDA_H = 2.4477 \cdot 0.5 \cdot (MSE_x + MSE_y)$ 5. Obtain the vertical NSSDA$_Z$ value according to the following expression: $NSSDA_z = 1.9600 \cdot MSE_z$
Source	FGDC (1998). *FGDC-STD-007: Geospatial Positioning Accuracy Standards, Part 3. National Standard for Spatial Data Accuracy*. Federal Geographic Data Committee, Reston, USA.

METHODOLOGY FOR POSITIONAL ACCURACY ASSESSMENT OF GEOGRAPHIC INFORMATION (UNE 148002)

The Spanish UNE 148002:2016 standard (Methodology for positional accuracy assessment of geographic information) was developed by UNE (Spanish Standardization Body) and emerged as an alternative to the problems associated with a lack of normality in positional errors and the need for sequential (i.e., lot-by-lot) control processes. The statistical and procedural basis of the UNE 148002 standard is formed by the international standards ISO 2859-1 (ISO 1999) and ISO 2859-2 (ISO 1985). These standards are widely applied in the industry and services sectors, offering extensive practical experience, and provide a common framework for working with the industry. Moreover, the use of these standards provides a common framework for expressing quality levels for both quantitative and qualitative elements, which is advantageous for reporting positional quality and other components of spatial data quality, such as thematic quality and completeness. For positional accuracy, the UNE 148002 standard is somewhat related to the NMAS standard. The existence of lots must also be considered for its application. In addition to what has already been discussed on the statistical method, the UNE 148002 standard follows the ISO 19157 framework, and provides recommendations on the scope of control, the sample of control elements, conditions of the elements used as a reference, and how to consider the metaquality of the process. Table 6 and Table 7 summarize the steps of the process based on the international standards ISO 2859-1 and ISO 2859-2, respectively.

Table 6. UNE 148002 standard - 'Lot-by-lot' process

Comparison method	With sources having higher accuracy.
Positional component	Horizontal and vertical, separately or together.
Element type	Well-defined points.
Accuracy standard	Quality is expressed by the AQL (acceptable quality level) index and a metric tolerance, which must have been indicated as conformity levels. **Report**: *Lot ID=____ of size N has been accepted/rejected for an AQL=__% and a metric tolerance D$_{tol}$=___[m].*
Description	The application of the lot-by-lot control (based on ISO 2859-1) is only suitable if there is a sequence of 10 or more lots. If the lot sequence is smaller, the isolated lot process should be applied. The quality is rated by the AQL index. This index represents the lowest process quality level that the user can consider acceptable, on average. This index is linked to a tolerance metric (e.g., D$_{tol}$ = 2.0 m). There is a lot-by-lot supply of data sets to be positionally assessed, which have been generated under homogeneous conditions. The process consists of accepting or

	rejecting each of the lots presented for inspection. For a lot of size *N*, the standard defines sampling plans by the pair *{n, c}*, such that, in a sequence of lots, accept/reject decisions are made for entire lots from a random sample of size *n* taken from each of the lots in the sequence. The accept/reject decision is made by comparing the number of error cases *d* found in the sample, and which present a size greater than the considered metric tolerance D_{tol}. It is accepted if $d \le c$, and otherwise rejected. ISO 2859-1 defines several inspection levels and provides sampling plans with different severities, which are adopted according to the progress of the process.
Procedure	1. Establish the AQL and D_{TOL} to be applied. 2. Establish the general inspection level to be applied to the entire sequence (by default, general level II). 3. Determine the size *N* of each lot to be submitted. 4. Use Table 1 of ISO 2859-1 and determine the code letter that corresponds to the lot size and general inspection level to be applied. 5. Using the series of Table 2 of ISO 2859-1, determine the sample size *n* and the acceptance and rejection values {Ac, Re} for the AQL and code letter under consideration. 6. Select a sample size *n*. 7. Calculate the error for each point in each component: $e_{x_i} = \|x_{p_i} - x_i\|$ $e_{y_i} = \|y_{p_i} - y_i\|$ $e_{z_i} = \|z_{p_i} - z_i\|$ where: x_i, y_i, z_i are the coordinates in the reference (RDS). $x_{p_i}, y_{p_i}, z_{p_i}$ are the coordinates in the product (ADS). 8. Establish for each error whether it is in or out of specification. $e_i \le D_{Tol} \rightarrow In; \quad e_i > D_{Tol} \rightarrow Out$ 9. Count the number of out-of-tolerance cases in the lot and compare with the {Ac, Re} values extracted from the table to make the accept/reject decision for the lot. 10. The rules of change proposed by the standard are applied to ensure the user's and producer's adopted risks levels (α and β).
Source	UNE (2016). UNE 148002:2016 *Methodology for positional accuracy assessment of geographic information*. UNE, Madrid. ISO (1999). ISO 2859-1:1999 *Sampling procedures for inspection by attributes – Part 1: Sampling schemes indexed by acceptance quality limit (NCA) for lot-by-lot inspection*.

Table 7. UNE 148002 standard- 'Isolated lot' process

Comparison method	With sources having higher accuracy.
Positional component	Horizontal and vertical, separately or together.
Element type	Well-defined points.
Accuracy standard	Quality is expressed by the AQL index and a metric tolerance, which must have been indicated as conformity levels. **Report**: *Lot ID=____ of size N has been accepted/rejected for an AQL=__% and a metric tolerance* D_{tol}*=___[m].*
Description	The control of isolated lots (ISO 2859-2) is the counterpart of ISO 2859-1, applicable to the case of single lots or sequences of less than 10 lots. The quality is rated by the AQL index. This index represents the worst process quality level that the user can consider acceptable, on average. This index is linked to a metric tolerance (e.g., $D_{Tol} = 2.0\ m$). As the AQL is a mean value for a sequence, it should include the case of an isolated lot, referred here as the quality limit (QL). The QL is a quality level that, for the purpose of sampling inspection, has a low probability

	of acceptance. The conversion between QL and AQL is: $QL \approx 3 \times AQL$. This summary refers to 'Procedure A' of ISO 2859-2, which is used when supplier and consumer consider an isolated lot. For a lot of size *N*, the standard defines sampling plans by the pair $\{n, c\}$, such that an acceptance/rejection decision is made for the entire lot from a random sample of size *n*. The acceptance/rejection decision is made by comparing the number of error cases *d* found in the sample, and which present a size larger than the considered metric tolerance D_{Tol}. It is accepted if $d \leq c$, and otherwise rejected. ISO 2859-2 presents numerous sampling plans in its Table A.
Procedure	1. Establish the AQL and D_{TOL} to be applied. 2. Convert the AQL to QL. 3. Determine the size *N* of the lot to be submitted. 4. Use Table A of ISO 2859-2 and determine the sample size *n* and the acceptance value *c*. 5. Select a sample of size *n*. 6. Calculate the error of each point in each component: $e_{x_i} = \lvert x_{p_i} - x_i \rvert \qquad e_{y_i} = \lvert y_{p_i} - y_i \rvert \qquad e_{z_i} = \lvert z_{p_i} - z_i \rvert$ where: x_i, y_i, z_i are the coordinates in the reference (RDS). $x_{p_i}, y_{p_i}, z_{p_i}$ are the coordinates in the product (ADS). 11. Establish for each error whether it is in or out of specification. $e_i \leq D_{Tol} \rightarrow In; \quad e_i > D_{Tol} \rightarrow Out$ 12. Count the number of out-of-tolerance cases in the lot and compare with the value of *c* extracted from the table to make the decision to accept/reject the lot.
Source	UNE (2016). UNE 148002:2016 *Methodology for positional accuracy assessment of geographic information*. UNE, Madrid. ISO 2859-2:1985. *Sampling procedures for inspection by attributes — Part 2: Sampling plans indexed by limiting quality (LQ) for isolated lot inspection*. International organization for Standardization

GENERAL METHOD FOR POSITIONAL ACCURACY ASSESSMENT

Considering that positional accuracy assessments are carried out by obtaining a sample composed of an ADS and a RDS, in which the latter derives from a source with greater accuracy (e.g., from a field observation), and generated by an external direct method according to the classification of ISO 19157, the following general process is proposed:

- Definition of the positional accuracy assessment process. If there are suitable data product specifications, the key aspects of the assessment should be extracted from those. Otherwise, they must be defined prior to the assessment. The aspects that define the assessment are:
  - The quality element. This can be any of those established by ISO 19157 for positional accuracy (e.g., absolute or relative positional accuracy). Whenever possible, it is recommended to work with absolute positional accuracy[21].
  - The scope of the quality assessment. The thematic, geographic, and temporal aspects, among others, that adequately delimit the set of geographic objects to be assessed (e.g., all constructions in a given geographic region) shall be established.
  - Data quality unit. This is the conjunction of the two previous aspects.
  - Conformity level. In the case of a quality control, there should be one or more conformity levels to make the decision to accept/reject the product (e.g., no more than 10% of positional errors greater than a given tolerance).

---

[21] In the past, cartography typically attained very high relative accuracy but failed in absolute accuracy. Nowadays, thanks to GNSS systems and their reference frameworks, high absolute positional accuracy is obtained. The latter circumstance also ensures high relative accuracy, whereas the reverse situation does not.

  - Assessment method. This should follow the specification of the assessment procedure. It is not sufficient to indicate the application of a positional accuracy assessment standard (e.g., NMAS or EMAS). The assessment method should cover all relevant aspects of the assessment, including the required accuracy, size and generation of the control sample, fieldwork guidelines (e.g., survey techniques and special cases), and statistical data analysis (e.g., verification of basic hypotheses and treatment of outliers).
  - Quality measure. The quality measure(s) should be specified (e.g., measure ID 45 from the list of Annex D of the ISO 19157 standard).
- Assurance of the positional accuracy assessment process. Once all the elements that compose the definition of the positional accuracy assessment process are available, the process should be properly applied. For this, the following aspects should be considered:
  - Having qualified personnel. The personnel who will carry out the field and office work must have adequate training and qualification.
  - Having appropriate instruments (e.g., GNSS equipment). The equipment must be suitable for the selected method and must be properly calibrated. The instruments should be combined with a suitable method to ensure that the RDS has the required accuracy to be used as a reference against the ADS.
  - Random sample[22]. A sample of assessment points[23] should be randomly generated and adequately satisfy the representativeness conditions. As a reserve, a limited set of alternative assessment points should be proposed (Table 8).
  - Field work planning. The work routes, dates, and times should be planned as usual for engineering projects.
  - Office work planning. The office processing work should be planned as usual for engineering projects.

Table 8. Replacement, movement, and displacement of assessment points (Source: JRC 2012)

Operation	Assessment points
Replacement	2
Movement Displacement	2
Note: Replacement refers to the substitution of a control point that is not accessible for one that is. Movement refers to points that are accessible but have a problem (e.g., obstacles that prevent good GNSS signal reception). In this case, another point is selected in the vicinity of the initially identified point. Displacement refers to giving eccentric positioning when the problem detected in the case of movement exists and no other suitable point can be found.	

- Fieldwork. Fieldwork consisting of obtaining accurate observations on field objects that form the RDS of the assessment sample is executed. Relevant aspects of this process are:
  - Unambiguous identification of the objects and their characteristics. There can be no error in the field identification of the objects and characteristics that are part of the assessment sample (e.g., the object is a construction and its selected characteristic for measurement is its south corner). It is important that there are no identification errors, and proof of correct identification should be generated (e.g., sketches, photographs. etc.).

---

[22] Although this aspect is methodological, it has been included in the control section as it is crucial to ensure the representativeness of the sample.

[23] Assessment points are traditionally termed 'control points', but since the term 'control' has a restricted meaning in this guide, they are referred to as 'assessment points', in such a manner that they serve both to estimate and to control the product.

  - Observation. The selected survey method (e.g., GNSS applying a fast static method) should be executed (e.g., acquisition time, horizon mask, and measures to avoid multipath effect and electromagnetic interference) following a procedure previously standardized by the organization in such a way that ensures the accuracy of the method.
- Office work. The office processing work consisting of obtaining the coordinates of the field/ product objects (RDS and ADS) that compose the assessment sample is executed. Relevant aspects of this process are:
  - Observation processing. The field observations should be processed by a calculation tool and following a procedure previously standardized by the organization, in such a way that ensures the accuracy of the method.
  - Unambiguous identification in the data set. There can be no errors in the identification of the objects and characteristics homologous to those of the field in the data set. It must be ensured that there are no identification errors.
  - Verification of cartographic prerequisites. It should be verified that the coordinates of the ADS and RDS sets correspond to the same reference system and map projection.
  - Verification of statistical assumptions on positional error. It should be verified that the statistical assumptions required as a basis for the application of the statistical method used in the analysis (e.g., normality, homoscedasticity, outliers, etc.) are met. This should be conducted following a procedure previously standardized by the organization.
  - Application of one or more PAAM for positional accuracy assessment (e.g., NMAS, EMAS, and NSSDA). This should follow a procedure previously standardized by the organization.
  - Analysis and reporting. The numerical results should be analyzed and graphic outputs generated (e.g., an error distribution map). A procedure previously standardized by the organization (e.g., through the use of a results template) should be followed.

CONSIDERATIONS FOR EXECUTION

This section considers a series of aspects related to execution that are particularly relevant to ensure the assessment quality (metaquality).

REFERENCE DATA FOR THE ASSESSMENT

For a valid assessment and ensure confidence in the results, it is necessary to have an external and independent source (external direct assessment method) to generate the RDS and for minimum requirements to be met in the execution of the work. The reference can derive from an *ad hoc* acquisition of the real world, through field work, or any other data set that meets the following requirements:

1. **Independence.** The reference (RDS) must not share common processes with the product under assessment (ADS). The independent processes are, for example, those carried out on different dates, by different work groups, and with different instruments. In case of doubt, statistical techniques can be applied to demonstrate and prove independence.
2. **Greater accuracy.** The accuracy of the reference must be at least three-times greater than the accuracy of the ADS. To ensure this requirement, it is necessary to have an estimate of the accuracy of the ADS and to design and execute a RDS acquisition method that ensures this requirement is met.

3. **Completeness.** The assessment sample should be consistent with the indicated scope of the quality assessment, adequately covering all aspects that define that scope (e.g., spatial, temporal, and thematic).

It is recommended that the RDS is derived from *ad hoc* designs, generated by more accurate methods executed with more cautious processes and from a more complete reality than the product. These aspects are discussed in the following subsections.

ACCURACY OF THE ASSESSMENT TASKS

The results of the positional accuracy assessment should reflect the accuracy of the data of interest (the ADS) and should not be affected by the assessment method and the RDS. Therefore, to minimize the effects that the RDS could have on the assessment, sources with greater accuracy than the ADS under assessment should always be used. To ensure this, a general rule is to use a reference with coordinates derived from processes that are independent of those generating the data to be assessed and that are at least three-times more accurate (i.e., with three-times lower uncertainty). This rule (independence and greater accuracy) has a statistical explanation. The independence of the control and production processes is necessary as, if it did not exist, it would not be possible to detect the existence of bias. On the other hand, the requirement of greater accuracy (lower uncertainty) in the RDS allows that, from the perspective of the composition or propagation of uncertainties, the practical impact of the accuracy of the RDS is limited (approximately 5% of the estimated value in the case of the 1:3 ratio indicated above, at least three-times more accurate).

It is essential, therefore, to determine the theoretical accuracy of the product to be assessed, as this affects the accuracy of the reference to be used for the assessment and the method for obtaining the coordinates. For example, for an ADS with a scale of 1:5,000 with a maximum theoretical accuracy set according to the perception limit of the human eye, if this limit is set at ¼ of a millimeter with a probability of 95% (i.e., in 5% of cases the limit may be exceeded), a circle of expanded uncertainty to 95% will have a radius of 0.00025 x 5000 = 1.25 m. Therefore, for the RDS to be at least three-times more accurate, a standard uncertainty of $1/3 \times 1.25/k = 0.17$ m (or less) is required, where k is the coverage factor obtained from the circular normal distribution[24]. From this knowledge, the GNNSS observation and calculation method that guarantees this level of accuracy can be determined (e.g., RTK, Stop & Go, and fast static).

Another aspect to consider is the type of data; the assessments for altimetry, planimetry, a mosaic, vector elements, point-clouds, or DTM are not identical. Therefore, each requirement must be specifically considered, as the field work method, types of control elements, and type of sampling are not necessarily identical.

NUMBER AND DISTRIBUTION OF ASSESSMENT POINTS

Both the estimation and control processes are based on representative samples. Therefore, it is crucial to have clear criteria for sample size and other qualitative aspects to ensure representativeness. As previously mentioned, those methods with estimation and control perspectives require different sample sizes given that their purposes are different.

Quantity in estimation. In estimation, the sample size should be related to the population size, the variability of the characteristic of interest in the population, the intended precision of the estimation, and a level of significance. In the case of positional accuracy assessment, the characteristics of

---

[24] The coverage factor k has a value of 2.4477 for a 95% probability in the circular normal distribution. For the one dimensional case k = 1,96 for the same 95%.

interest in the population may be the error mean and the standard deviation. For the case of the PAAM based on counts, it is also of interest to estimate the proportion of errors greater than a given tolerance(s). For the particular case of positional accuracy, most standards consider a minimum size of 20. There are several studies indicating that this quantity is not adequate for positional accuracy estimation processes. Below are formulas for determining the sample sizes for each of the above cases:

- Mean estimation. This option is suitable for estimating bias (see Example 1). The determination of the sample size results directly from Equation 3.

$$n = \frac{N\hat{\sigma}^2}{\hat{\sigma}^2 + N\frac{e^2}{t^2_{\alpha/2}}}; \qquad \text{if } N \to \infty \qquad n = \frac{t^2_{\alpha/2}\hat{\sigma}^2}{e^2} \qquad \text{Equation 3}$$

- Proportion estimation. This option is suitable for estimating the proportion of cases that exceed a given tolerance (see Example 2). The determination of the sample size results directly from Equation 4.

$$n = \frac{N\cdot P\cdot Q}{P\cdot Q + \frac{(N-1)\cdot e^2}{z^2_{\alpha/2}}}; \qquad \text{if } N \to \infty \qquad n = \frac{z^2_{\alpha/2}\cdot P\cdot Q}{e^2} \qquad \text{Equation 4}$$

- Standard deviation estimation. This option is suitable for estimating precision (see Example 3). The determination of the sample size is more complex as the Chi-square distribution modeling the variance behavior is not symmetric and its shape depends on the degrees of freedom (the sample size). This case is solved iteratively between Equation 5 and Equation 6 (parts a and b), which can be resolved by fixing two of the three parameters (υ, n, α).

$$\alpha = p_1 + p_2 \qquad \text{Equation 5}$$

$$p_1 = P\left[\frac{(n-1)s^2}{\sigma^2} > (1+u)^2(n-1)\right] = P[\chi^2_{n-1} > (1+u)^2(n-1)\,] \qquad \text{Equation 6a}$$

$$p_2 = P\left[\frac{(n-1)s^2}{\sigma^2} < (1-u)^2(n-1)\right] = P[\chi^2_{n-1} < (1-u)^2(n-1)\,] \qquad \text{Equation 6b}$$

where:

P	*Probability of correct assignment in the binomial model.*
Q = 1 -P	*Probability of incorrect assignment in the binomial model.*
N	*Population size.*
n	*Sample size.*
α	*Intended significance level.*
e	*Allowed error in the estimation by sampling in units of the parameter being estimated.*
υ	*Allowed error in the estimation by sampling as a fraction of the deviation.*
$Z_{\alpha/2}$	*Statistic corresponding to the normal distribution with significance α.*
$t\alpha/2$	*Statistic corresponding to the Student's t-distribution with significance α.*
χ^2_{n-1}	*Statistic corresponding to the Chi-square distribution with n-1 degrees of freedom.*

Example 1: Mean estimation. This example considers the case of a grid-type digital terrain elevation model representing a large area and bare ground surface from which the mean error in the Z coordinate is to be estimated. According to ASPRS (2015), the normality of the altimetric errors can be assumed. Moreover, the population size is considered to be infinite ($N \to \infty$). In this case, the formula presented in Equation 3 can be applied. To determine a sample size, the following input values are required: i) an estimate for the standard deviation of the population, ii) the permitted error in the estimate, and iii) the significance level. To illustrate the importance of the estimation precision, Table 9 presents results for different values of this parameter whilst holding the others constant

(*ceteris paribus*). As can be seen, for this case, the ratio between the highest and lowest precision values is 1 to 4 (0.5 m vs. 2 m), while the ratio of the corresponding sample sizes is 16 to 1 (753 vs. 47), demonstrating the impact of this parameter on the estimation of the mean value.

Table 9. Sample size to estimate the mean error of a terrain elevation model as a function of the intended accuracy for N → ∞ and σ=7 m

N	σ (m)	$Z_{\alpha/2}$ $\alpha = 0.05$	Estimation precision (± m)	Sample size
∞	7	1.96	0.5	753
∞	7	1.96	0.75	335
∞	7	1.96	0.9	232
∞	7	1.96	1	188
∞	7	1.96	1.1	156
∞	7	1.96	1.2	131
∞	7	1.96	1.3	111
∞	7	1.96	1.4	96
∞	7	1.96	1.5	84
∞	7	1.96	1.6	74
∞	7	1.96	1.7	65
∞	7	1.96	1.8	58
∞	7	1.96	1.9	52
∞	7	1.96	2	47

Example 2: Proportion estimation. Considering the previous case of a grid-type digital terrain elevation model representing a large area, but focusing on the area that does contain bare ground, normality cannot be considered. The interest in this case is to determine the percentage of points that are below/above a threshold determined by a metric tolerance t. As in the example 1, the population size is assumed infinite ($N \rightarrow \infty$). In this case, the formula presented in Equation 4 can be applied. Without prior knowledge of the proportion, the condition $p = q = 0.5$ can be assumed as the worst case, that is, this requires the largest sample size. The results are presented in Table 10. For an estimation with 95% confidence and an estimation precision of plus or minus ten percent (10%), for example, a sample size of 96 control elements is required. With this sample, the percentage of cases exceeding the tolerance would be determined by counting, and this would constitute the estimation.

Table 10. Sample size to estimate a proportion of cases exceeding a tolerance as a function of the intended precision of the estimation and for N → ∞ and p = q

Confidence level	Estimation precision							
	±0.30	±0.25	±0.20	±0.15	±0.10	±0.05	±0.025	±0.01
0.8	5	7	10	18	41	164	657	4,106
0.9	8	11	17	30	68	271	1,082	6,764
0.95	11	15	24	43	96	384	1,537	9,604
0.975	14	20	31	56	126	502	2,010	12,560
0.99	18	27	41	74	166	663	2,654	16,587

Example 3: Standard deviation estimation. For better convenience, Figure 5 provides a graphical representation of Equation 6a and Equation 6b. Suppose such a case where a control sample size is required for the mean standard deviation to be within ± 15% of its true value, with 90% confidence. Figure 5 indicates that a sample size (n) of 60 is required, which is the intersection of the vertical line that crosses 0.9 on the horizontal axis and the dotted line representing the 15% hypothesis. Figure 5

can be used in a different manner. For example, in such a case of n = 20 control points, if 95% confidence is intended for this sample size, the graph indicates that the width of the interval is ± 31% of its true value.

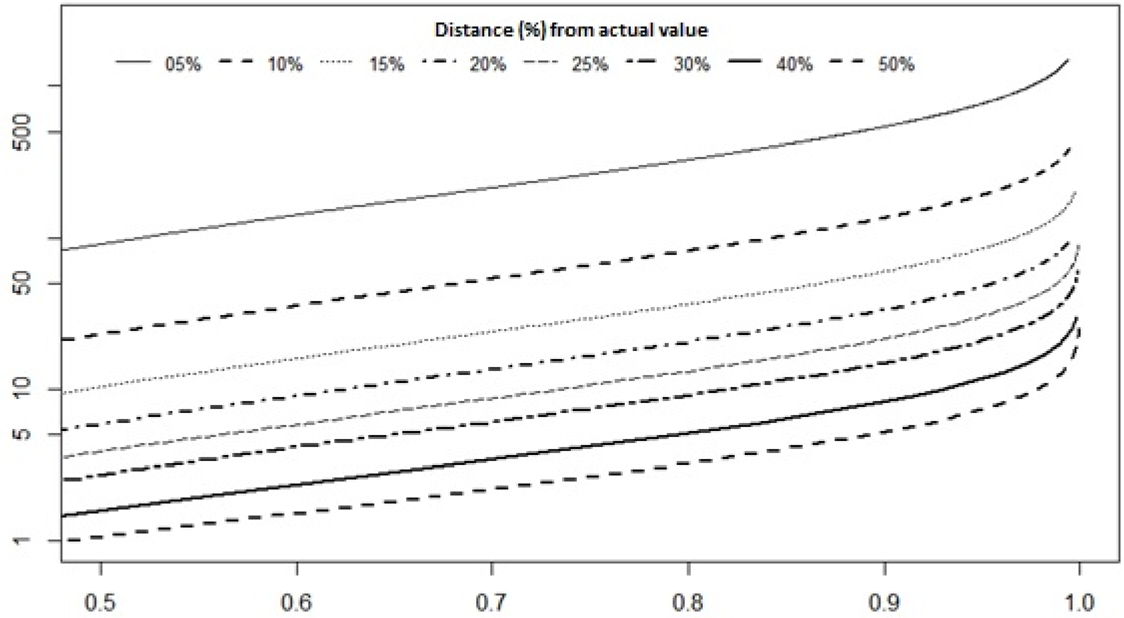


Figure 5. Relationship between the confidence coefficient [0.5, 1.0] and sample size [1, 500] required to estimate the standard deviation for different distances of the estimation from the actual value

Quantity in control. In the case of control by statistical hypothesis testing, any sample size ensures the desired level of the type I errors (producer's risk). In control processes, sample size is more related to type II errors (user's risk). Therefore, larger sample sizes are required if fewer type II errors are intended. As indicated above, for the particular case of positional accuracy, existing standards consider a minimum size of 20. Whilst this may be adequate for quality control that intends to ensure the desired level of type I errors, this is not sufficient to ensure the desired level of type II errors. In the case of PAAM, it has only been recently suggested that sample sizes should be linked to the project size to be controlled, and the ASPRS method (2015) includes a look-up table for the required number of control elements according to the number of hectares, as presented in Table 11. The UNE 148002 (UNE 2016) standard also specifies sample sizes according to the size of the population to be assessed (the lot), supported by the International Standard ISO 2859 parts 1 (ISO 1999) and 2 (ISO 1985) for lot-by-lot and isolated-lot inspection processes, respectively.

Table 11. Recommended sample sizes based on project surface area (Source: ASPRS 2015)

Project area (km^2)	Horizontal accuracy testing of orthoimagery and planimetrics	Vertical and horizontal accuracy testing of elevation data sets		
	Total number of static 2D/3D checkpoints (well-defined points)	Number of static 3D checkpoints in non-vegetated terrain	Number of static 3D checkpoints in vegetated terrain	Total number of static 3D checkpoints
≤500	20	20	5	25
501–750	25	20	10	30
751–1,000	30	25	15	40
1,001–1,250	35	30	20	50
1,251–1,500	40	35	25	60
1,501–1,750	45	40	30	70
1,751–2,000	50	45	35	80

2,001–2,250	55	50	40	90
2,251–2,500	60	55	45	100

Distribution. The selection of the assessment points should be in accordance with the scope defined in the specifications for the positional accuracy assessment (the data quality unit). In addition, an adequate distribution of the assessment points is required to attain representativeness, and, for this purpose, two basic requirements are established:

- **Randomness.** The assessment points should be selected randomly from the geographic objects belonging to the subpopulation defined by the scope. If the selection is not random (e.g., performed by an operator), bias is introduced into the sample. As it is not always possible to observe all the points initially determined due to several accessibility problems (e.g., owner obstruction, flooding, destruction, etc.), a list of previously determined and randomly selected alternative points should be used.
- **Representativeness.** This refers to an adequate spatial, thematic, and temporal distribution among other factors. In the case of spatial distribution, recommendations include that presented in Figure 6, which is valid in the case where the area to be controlled is uniform with respect to its elements and their uncertainty, such that a homogeneous distribution can be assumed. Figure 6 recommends a distribution by quadrants, with each quadrant containing at least 20% of the control points and a distance between these elements of the order of 1/10 of the diagonal[25]. Thus, if the positional accuracy of the data used in a linear engineering project (e.g., railroad, highway, etc.) is to be assessed, the shape of the area of interest will be elongated and a more appropriate spatial distribution criterion should be considered. To ensure that the area of interest is fully covered and there are no extrapolation problems at the borders, the JRC (2012) indicates that the thematic assessment points should cover the area of interest as well as a buffer (25% of the area), and the distribution of points following these proportions: 80% in the area of interest and 20% in the buffer area. This consideration can be equally valid for positional accuracy assessment.

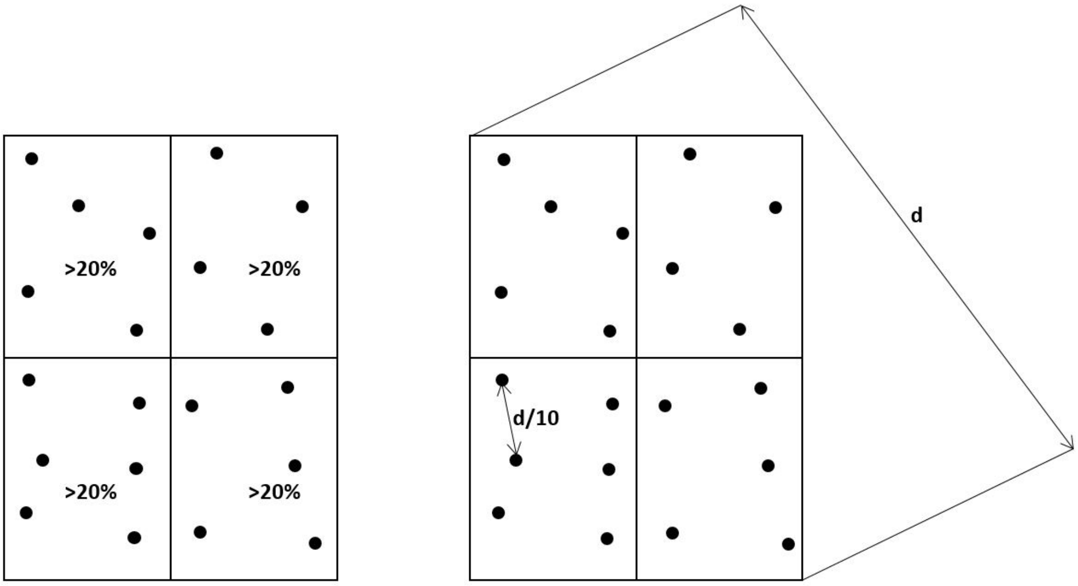


Figure 6. Recommendation for the spatial distribution of the assessment points (Based on MPLMIC 1999)

---

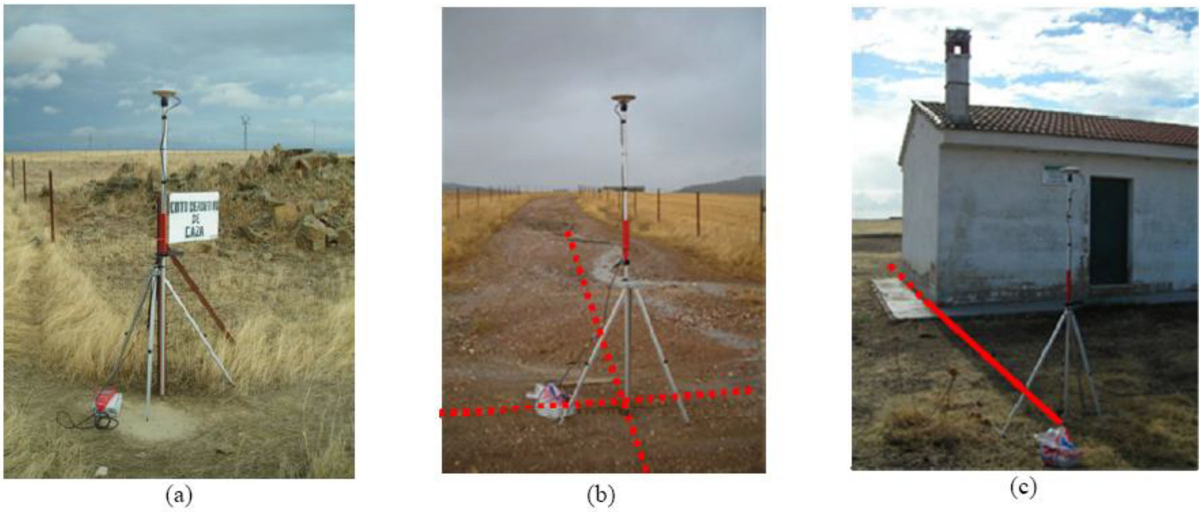
[25] The JRC 2012 indicates that this distance should not be less than 1/7 of the diagonal.

IDENTIFICATION AND OBSERVATION OF ASSESSMENT POINTS IN THE FIELD

The correct execution of field work is a key aspect to ensure an adequate reference is obtained and, therefore, a well-executed positional accuracy assessment. This step of the process must be performed with the greatest of care as malpractice would invalidate all the assessment process. Given that field work is very expensive, it must be correctly executed on the first attempt to prevent deviations from the plans. Some guidelines for the execution of field work are as follows:

- **Unambiguous identification.** Only planned assessment points that can be unambiguously identified in the field should be observed. In the case of any doubt, a reserve point should be considered.
- **Observation.** The observation of the coordinates of the planned assessment point(s), for example, using GNSS techniques, should be carried out provided there are optimal conditions for this (e.g., for GNSS methods: sufficient observation horizon, no antennas or radio-electric obstructions, etc.). If a direct observation of the position is not possible, an eccentric acquisition can be carried out provided that the auxiliary methods do not degrade the accuracy of the final coordinate assigned to that element with respect to the assessment requirements. In all cases, the selected position(s) should be clearly indicated in accompanying sketches and reviews. Three cases are outlined below that illustrate different possible situations during field work (Figure 7):
  - Case A. An isolated break in a wire fence forming an angle of almost 90°. This made identification in the field and in the ADS is very easy and provides low positional uncertainty and a low possibility of error. Moreover, the flat and unobstructed terrain allow GNSS observation.
  - Case B. A road intersection. Here, the measured point has no real representation on the ground, and it is, therefore, necessary to estimate the intersection of the axes, for which auxiliary measurements are required. This type of point is less accurate than that in Case A.
  - Case C. A point located at the corner of a house that cannot be measured due to antenna obstruction of the observation horizon. Here, the observation point can be moved to a position with a suitable distance for GNSS observations. This is an eccentric measurement that can be supported by topographic techniques and COGO[26] tools.

In cases B and C, the auxiliary measurements, if not carefully taken, can reduce the positional accuracy achieved with GNSS observations. The final positional accuracy should remain valid for the purposes of the assessment (i.e., "at least-three times better than the product").

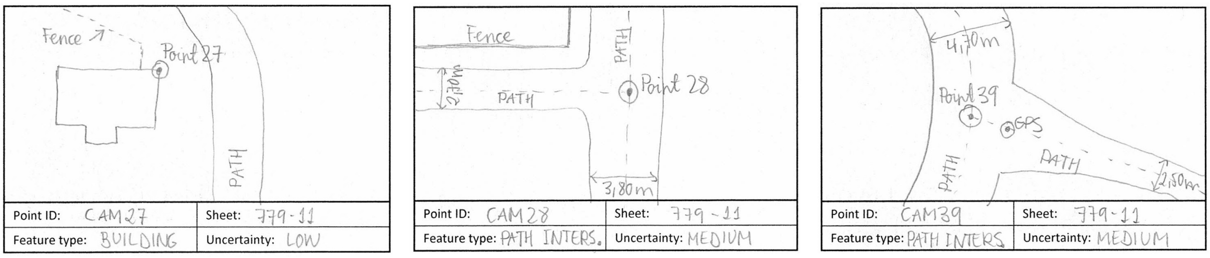

(a) (b) (c)

---

[26] http://wiki.gis.com/wiki/index.php/COGO

Figure 7. Examples of situations that can occur during field work to obtain the coordinates of assessment points (Source: Fundamentals of Geographic Information Quality Assessment; Ariza-López Ed. 2013)

- **Review of the assessment point.** Whenever an assessment point is surveyed, photographs are taken and a sketch is created with the level of detail appropriate to the requirements of the assessment and the type of product. All this information is included in the review of each of the assessment points so that it can be consulted and used in the production stage as well as for audit processes and metaquality assessment. Figure 8 presents examples of assessment point sketches.



Point ID: CAM27 | Sheet: 779-11
Feature type: BUILDING | Uncertainty: LOW

Point ID: CAM28 | Sheet: 779-11
Feature type: PATH INTERS. | Uncertainty: MEDIUM

Point ID: CAM39 | Sheet: 779-11
Feature type: PATH INTERS. | Uncertainty: MEDIUM

(a) (b) (c)

Figure 8. Examples of sketches for a review (Source: Ariza-López Ed. 2013)

IDENTIFICATION AND OBSERVATION OF ELEVATION IN THE FIELD

Positional accuracy assessment can be carried out on altimetry products such as digital elevation models (DEM). In these cases, it is customary to assume that planimetric positional accuracy is adequate and to assess the vertical component independently.

In grid-type DEMs, which are the most common, no well-defined or easily identifiable points are available, requiring a specific method to overcome this limitation. This can be done based on two elements, namely constant slope planes and an interpolation process. In the latter case, given that there are many interpolation methods that can provide very different values when applied to the same data set, methods that limit this constraint should be considered. One option consists of limiting the interpolation to the most elementary method that best fits the reality of the terrain. For this, the reference points are selected from terrain surfaces considered as a plane, that is, they have a constant slope. These planes can be easily determined with prior GIS analysis[27] in the office. This can also be visually confirmed in the field with sufficient approximation and safety. These planes should be represented at the resolution of the model itself, that is, at least twice the grid spacing (side ≈ 2 × grid spacing). Hence, the points belong to a plane and the appropriate interpolation method is linear. An additional constraint is imposed on the previous condition of the plane by limiting its slope. This is intended to limit the influence of the accuracy (inaccuracy) of the planimetric component on the assessment of the altimetric component[28]. Therefore, the slopes of these planes should be as low as possible. All these conditions can be adequately managed in the office using GIS capabilities for the design of the control sample.

When executing the observation, once a constant slope plane is located, four points (quartet) should be surveyed using a GNSS methodology appropriate to the accuracy required for the reference. The relative horizontal spatial distribution of this quartet of points should be a quadrangle with a side

---

[27] Geographic Information System, understood as a software tool with analysis capabilities.

[28] If the planes had a slope of zero (flat terrain), the horizontal uncertainty of a position would not affect the altimetry of that position.

equal to the grid spacing of the DEM being assessed, which guarantees that there is a point of the product grid in its interior. Once the coordinates of the quartet are available, a plane can be adjusted by least squares using all of the points and, in that plane, the altimetric coordinate of the internal point of the grid can be obtained using linear interpolation. This scheme is depicted in Figure 9, where the dashed line represents a zone that can be assumed as a plane, crosses represent the points of the grid, and circles represent the quadrangle containing the point of the grid with an altimetric value obtained by linear interpolation of the plane defined by the quartet.

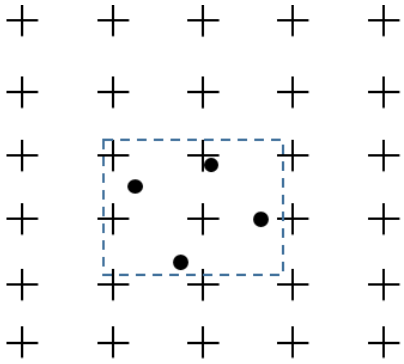


Figure 9. Example of a quartet configuration

IDENTIFICATION AND OBSERVATION OF ASSESSMENT POINTS ON IMAGES

Images are often used in positional accuracy assessment. In some cases, images can be used as a reference and, in many others, are related to the positional accuracy assessment of this type of product (e.g., an orthophotomosaic). In all cases, the use of images to extract coordinates should consider all aspects that directly affect the extraction of coordinates (e.g., data model, coordinate extraction method, operator, etc.). For the data model, the pixel size or spatial resolution of an image acts as a threshold of uncertainty for the extracted positions[29]. For the extraction of coordinates from the image, human operators usually perform this task —typically a single operator— and the confidence in the goodness of the extracted coordinates is based solely and exclusively on the experience and skill of that operator. This situation is unfavorable; the extraction of coordinates should be based on the means of different extractions[30]. In addition, during the extraction of coordinates from images, it may be convenient to use auxiliary methods to determine a position more precisely, such as the use of vector alignments to help locate the position of an intersection (Figure 9).

---

[29] If an image was perfectly georeferenced, the positions within a cell would be assimilated to that of the cell as there are no identifiable characteristics in its interior.

[30] Consider that in GNSS acquisition methods different epochs are observed and they are averaged to give a solution to a positioning.

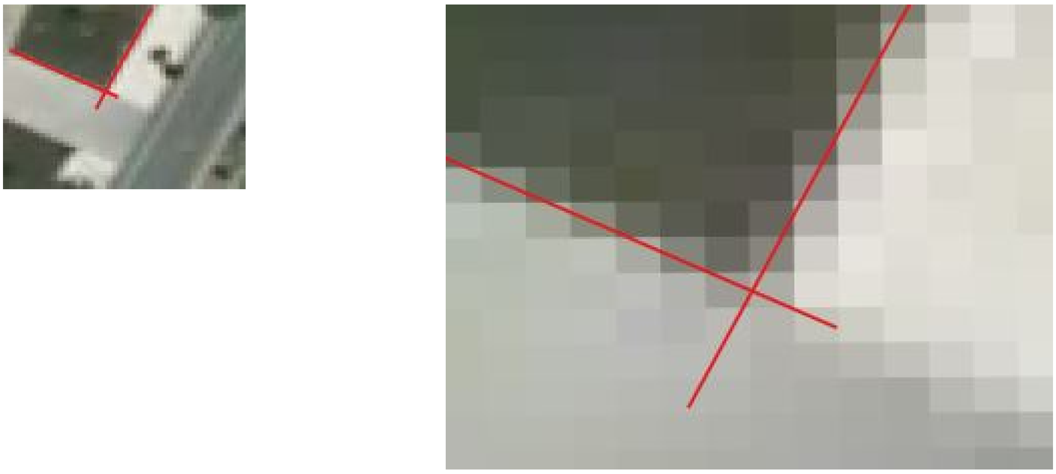


Figure 10. Example of a corner point in an image that is extracted as an intersection of alignments

CONSIDERATIONS FOR STATISTICAL ANALYSIS

Prior to the application of any positional accuracy assessment method (e.g., EMAS or NSSDA), it should be confirmed that the statistical assumptions on which the method is based are met, whether they are explicit or implicit or underlying assumptions. For example, Table 12 presents the relationships between the statistical aspects that should be considered and the positional accuracy assessment methods that have been presented.

Table 12. Suggested relationships between hypothesis testing and assessment methods

	Control of statistical assumptions					
Method	Randomness	Outliers	Normality	Bias	Independence	Homoscedasticity
NMAS	N					
EMAS	N	N	N		N	R
NSSDA	N	N	N	N	N	
UNE 148002	N					
N = necessary, R = recommended						

For example, in the cases of the NSSDA and EMAS where positional errors are supposed to follow a normal statistical distribution function, the former uses the normal distribution to propose a confidence level, and the latter supports the decision to accept or reject it. Therefore, if this assumption is not fulfilled, all of the resulting conclusions (the statistical inferences made) are questionable. Another important aspect to control is the presence of outliers, the quantities of which are linked to the distributional assumption, with their control recommended for methods assuming data normality. The EMAS method controls the presence of bias, and the method itself controls for this. On the other hand, the NSSDA indicates that bias should have been appropriately treated, and it is thus recommended to apply such a control before applying the NSSDA method. The NSSDA indicates that $0.6 < (MSE_{min} / MSE_{max}) < 1.0$ must be met, giving some control over homoscedasticity. In the case of the EMAS, the deviations in each of the components can be within the tolerance limits but very disparate, in which case a homoscedasticity test is recommended.

Based on these considerations, although this document is not intended to provide a thorough discussion of statistical analyses or standardize them, it is worthwhile to indicate their importance and the need to include them in well-documented processes. Below is a brief summary of these statistical aspects[31]:

- **Randomness of the sample.** Randomness is a quality that positional errors must have. Compliance is common in this regard, but it is important to verify randomness in the sample as a lack of randomness is an indicator of some degree of data manipulation (intentional or unintentional). Randomness ensures a stochastic (non-deterministic) process. There are several statistical tests to verify the randomness of a data set (errors, in this case), including the Wald-Wolfowitz test based on the presence of runs.
- **Normality of the sample.** The normal or Gaussian distribution function is the model generally assumed by most of the positional accuracy assessment methods. The lack of normality of some errors can be more or less serious; if mild, non-normality will not have a great impact on the analyses, but if pronounced (e.g., a bimodal distribution with many extreme values), the application of the normal model can lead to invalid results. In general, a lack of normality results from a range of causes and circumstances including: i) the presence of too many outliers; ii) the overlap of more than one process; iii) insufficient discrimination in the data (e.g., rounding, poor resolution, etc.); iv) removal of data from the sample; v) distribution of values close to zero or another natural limit; and vi) data following another distribution (e.g., Gamma, Weibull, etc.). In the case of positional error data, the most common situations are the first and second. When data are received from third parties, the fourth cause is also common. The verification of this assumptions can be carried out using graphical tests (e.g., QQ-Plot graphs) and appropriate statistical tests, such as the Shapiro-Wilk and Lilliefors tests. The positional accuracy assessment does not require strict normality, and it can be approximate, so the more permissive Kolmogorov-Smirnov test is recommended. It is known that LiDAR altimetry data from covered terrain do not follow a normal distribution (Maune 2007; Zandbergen 2008, 2011), and in these cases, error percentiles or proportions (count of cases) are used.
- **Presence of bias.** Bias or systematisms are trends introduced in the data due to, for example, equipment miscalibration, processes with systematic error, or operator involvement[32]. The statistical significance of a bias can be analyzed by a hypothesis test on a mean error value. In this case, it is generally assumed that the mean error is zero. If bias is present, an assignable cause must be determined. After identifying the cause and verifying that it is the one that truly generates the bias, it can be eliminated from the data by the appropriate mathematical operation.
- **Outlier detection.** Outliers are extremely small or large values that occur in reality with a low probability (they are thus 'untypical' because are out-of the normality). If there is a high percentage of outliers, they are no longer outliers and signify the presence of special circumstances and, in general, arise due to a range of processes. In such cases, the possible cause should be determined to improve the processes. Outliers are far from the true mean and greatly affect the calculated means and deviations as they have a strong leverage effect on the calculations assuming a normal distribution. Therefore, the presence of outliers results in poor estimates (due to overestimation) of means and deviations and of all their derivatives. There are several methods for the detection of outliers, such as applying a coverage factor *k* on the standard deviation, graphical comparisons, and the Generalized Extreme Studentized Deviate (GESD) method. The first approach is simplest, where *k* is a coefficient and the larger it is, the greater the degree of atypicality of the values to be eliminated.

---

[31] Supplementary material and R package codes that implement these analyses can be found at https://coello.ujaen.es/investigacion/web_giic/SubWeb_IPGH2016/resultados.html. More information on these aspects and how to control them can be found in Ariza-López (2013).

[32] An operator is generally positioned in a certain manner when making an observation and can introduce systematic errors as a result (MSHA, 2001).

- **Homoscedasticity of the components.** Homoscedasticity is the uniformity in the variational behavior of the errors in the components that are analyzed together. For example, for errors in X and Y in the case of planimetry, it is generally considered that $\sigma_x \sim \sigma_y$. This is a common assumption and is motivated by a certain logic in the processes (i.e., the components should work more or less the same) but also by the simplicity of the calculations (formulas). Altimetry often operates independently of planimetry, typically with differing uncertainty. If the assumption of homoscedasticity of the uncertainties in X and Y is not fulfilled, heteroscedasticity exists, indicating that the results of their combined analysis will not be valid. Therefore, for example, the NSSDA establishes a maximum range of difference between σ_x and σ_y, so that its formulas are applicable. There are several statistical tests to verify this condition, including the F-test, the Bartlett's test, and the Levene's test.
- **Non-correlation of the components.** Correlation indicates the proportion of the behavior of a variable that can be described by another variable, so that a certain dependence can be assumed to exist. The non-existence of correlation is a logical hypothesis, although it is also important to verify this as positional accuracy assessment methods rely on its absence to provide simpler calculation methods. If correlation does exist, it can be analytically modeled, but this makes the calculations more complex. Metrics that allow detection and quantification of correlation include the Pearson, Spearman, and Kendall coefficients.

METAQUALITY OF POSITIONAL ACCURACY ASSESSMENT

Metaquality refers to the quality of a quality result. That is, given a quality result related, for example, to the positional accuracy assessment of spatial data, how good is the quality of that result? The result may provide a higher or lower value but it cannot be truly considered if the quality of the process is not known and, therefore, the value cannot be trusted. Metaquality is proposed in the international standard ISO 19157 and includes the following elements:

- **Confidence.** Accuracy of a data quality result.
- **Homogeneity.** Expected or tested uniformity of the results obtained within the scope.
- **Representativeness.** Degree to which the sample has produced a result that is representative of the data within the scope.

To date, only the Spanish UNE 148002 (UNE 2016) standard considers these elements in relation to positional accuracy, as follows:

- **Confidence.** Qualitative and quantitative aspects should be considered with regard to confidence. Qualitative aspects relate to the rigor in the application of the methods and the participation of experts, who are the main guarantors of rigor. Quantitative aspects are related to the quantities in the application of the methods, such as sample sizes, the degree of statistical independence, and the relationships between the accuracies of the ADS and the RDS.
- **Homogeneity.** Aspects related to both the ADS and the RDS must be considered. The ADS may be more or less homogeneous due to its genesis. This is critical for an ADS in which many individuals or organizations have been actively involved with differing backgrounds, knowledges, and skills, or where different working methods have been applied (e.g., OpenStreetMap). The assessment process can also affect consistency. In assessment processes that are lengthy in space or time, appropriate quality management measures should be adopted to ensure uniformity at all times. The key elements to ensure homogeneity include, among others, the availability of written procedures, the establishment of standards in the training and qualification of the personnel, and the inclusion of verification mechanisms to ensure homogeneous processes.

- **Representativeness.** This must be assessed with a multiple perspective that includes, among other things, spatial (geographic areas), thematic (by typology of geographic objects of the data product), and temporal (by dates) aspects. Representativeness should be analyzed, as far as possible, by statistical techniques as sampling techniques are used in the assessment. In this regard, among other options, the following techniques can be applied: visual comparisons of histograms or distribution functions of the sample and the population; and adherence contrasts between the curves representing the distribution functions of the sample and the population (e.g., using the Kolmogorov-Smirnov test for continuous cases and the Chi-square test for discrete cases).

  Figure 11 shows an example of an analysis of the representativeness of samples based on altitude for the assessment of two DEM-type products. A comparison of two histograms is shown, in which the frequency distribution of each of the RDSs is shown in red and the distribution of the assessed ADSs is shown in blue. In the left-hand plot, the sample does not reflect the full range of variation of the population (e.g., there is no sample above 1,000 m), whereas in the right-hand plot, the two histograms almost completely overlap. Therefore, the figure on the right indicates better representativeness.

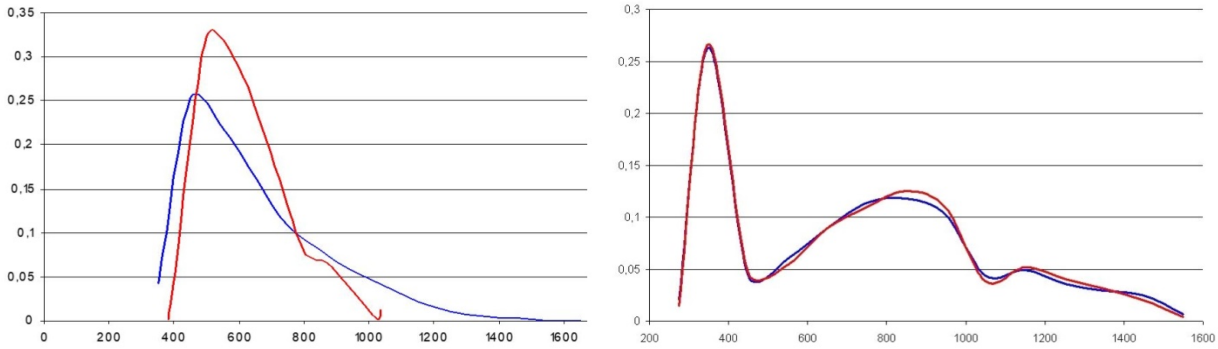


Figure 11. Examples of histograms from two different DEMs used to analyze the representativeness of the samples based on altitude (horizontal axis elevation, vertical axis frequencies)

ASSESSMENT REPORT

This final section considered an important aspect of any assessment—the reporting of the results. In the presentation of the methods in the section 'Positional accuracy standards' (see Tables 3, 4, 5, and 6), the legends that should be included in reporting according to that standards were indicated. However, this is insufficient as none of these standards establish well-defined processes for all aspects of a positional accuracy assessment.

With the aim of creating standalone quality reports mentioned in ISO 19157 (ISO 2013), and based on the Spanish UNE 148002 (UNE 2016) standard, Table 13 presents an outline of a detailed and adequate report. It should include, at the very least, the aspects considered necessary to obtain conformity with respect to the designed and applied assessment method. Each of the proposed steps is explained below, and the complete outline is developed further in Annex 1.

- **Identification of the ADS to be assessed.** This step identifies the ADS and describes the most relevant aspects (e.g., content, purpose, specifications, etc.) as well as the specifications specific to positional accuracy.

- **General aspects of the assessment.** The data quality unit (DQU) is established to clearly identify and communicate the quality element to be assessed. Moreover, the measure(s) and assessment method should be identified in such a manner that no relevant aspect remains ambiguous (as indicated in the different sections of this guide). For this, in order not to make the report too long, it is advisable to include the documents describing these aspects as external links.
- **Highest accuracy source (RDS) and list of coordinates.** This is a critical part of any assessment with regard to metaquality and should include content that allows the quality of the RDS to be assessed. Information can be provided on the reference source, whether it is planimetric and/or altimetric, its accuracy, the population coverage (sample size), thematic coverage, spatial coverage, whether interoperability has been verified (that the ADS and the RDS are in the same reference system), and how the sample was generated.
- **Statistical hypotheses testing on errors.** After presenting the list of errors, this part of the report focuses on the evidence of the verification that all statistical assumptions required by the applied PAAM(s) are met, whether implicit or explicit.
- **Results.** After presenting the final list of errors (which will differ from the initial list in cases where some data points have been removed, e.g., due to classification as outliers), this step presents the final results based on the applied method(s) (compliant/non-compliant or an estimated value) as well as the results of the measurements and their conformity, if applicable. The error distributions should also be included in the relevant ways (e.g., spatial, histogram, etc.) as well as the basic statistical parameters. If bias is present and it has been assigned, this should be explained. Finally, a brief interpretation of all the results together is desirable.
- **Metaquality of results and processes**. This part of the report should develop supporting explanations related to the metaquality elements based on the objective facts presented in the previous sections.
- **Date and signature of the responsible person.** All assessments must be dated and signed by the responsible technician.

Table 13. Outline of contents of an standalone quality report for positional accuracy assessment.

1) **Identification of the ADS to be assessed**
    - Name.
    - ID.
    - Producer.
    - Qualitative description.
    - Purpose.
    - Specifications.
    - Design accuracy (theoretical).
2) **General aspects of the assessment**
    - Data quality unit.
    - Assessed components.
    - Quality measures.
    - Assessment method.
3) **Highest accuracy source and list of coordinates**
    - Reference source.
    - Dimension.
    - Positional accuracy of the reference.
    - Coverage.
    - Interoperability
    - Other aspects (related to the acquisition method).
    - List of coordinates.
4) **Statistical hypothesis testing**
    - List of errors
    - Randomness.
    - Outliers.
    - Normality.
    - Bias.
    - Independence.
    - Homoscedasticity.
    - Interpretation of statistical hypotheses.
5) **Results**
    - Final list of errors.
    - Basic statistical parameters.
    - Graphs.
    - Bias assignment.
    - Results of the quality measures.
    - Results of the PAAM.
    - Interpretation of the results.
6) **Metaquality of results and processes**
    - Confidence.
    - Homogeneity.
    - Representativeness.
7) **Date and signature**
    - Date of report.
    - Place.
    - Responsible person.

FREQUENTLY ASKED QUESTIONS

A set of frequently asked questions are listed in this section with corresponding answers. Although all these aspects have been covered in the document, they are included here as a quick reference resource.

- **How can I generate a random and well-distributed sample of control elements?** The sample should be generated by a software tool that considers both criteria.
- **What is the minimum number of control points one should use in a positional accuracy assessment?** This depends on whether an estimation of the positional uncertainty of the ADS is intended or a simple control of its positional accuracy. In the former case, the number of points must be established based on the sampling theory and may be high. In the latter case, the number is not linked to the population size. A lower number is required to ensure type I errors (producer's risk) but a certain sample size is required to control for type II errors (user's risk).
- **Is homoscedasticity really present in the data?** Yes, there are many processes in which homoscedasticity occurs, such as in coordinates obtained by GNSS systems when the reference stations do not have an equivalent East-West and North-South distribution, in digitization on tablets and screens, and in scanning systems.
- **Is the normality of errors really necessary?** This depends on what the hypothesis of normality is used for, either in the applied PAAMs or for subsequent use of the assessment results. For example, if normality is used to propagate uncertainties, it is relevant. Similarly, the expansion coefficients to determine confidence intervals are based on the normal distribution. That is, any statistical inference based on normality will be questionable if the data are non-normal.
- **Why is accuracy often used as a synonym of precision?** This document assumes the definition of the ISO 5725-1 and JCGM200 (International Vocabulary of Metrology, VIM), which is not known or applied by all. In this standard, and for the VIM, there is a clear conceptual difference between accuracy and precision. However, in practical terms, when there is no bias, accuracy is equal to precision. For this reason, bias can be eliminated if it has an assignable cause, which is the most common situation (e.g., consider a calibration for which the device offset is corrected), and accuracy and precision can be confused as being the same. In any case, this practice generates confusion and care must be taken to use both terms correctly.
- **Why should the sample of assessment points be random?** The sample must be random so that no bias is introduced through the actions of the operators (personnel) involved in the process. If an operator selects the assessment points that compose the sample, the selection of these is biased in one way or another based on their experience (e.g., their considerations of what may be a more comfortable job, or their personal interests when visiting areas, etc.).
- **How many outliers are 'normal'?** Consider the following example: assumption of errors in altimetry (1D) that behave as normal. If the value of k in the normal distribution is 1.96 (a confidence interval of 95%), one should expect approximately 5% of the sample to be outliers. If the proportion of detected outliers far exceeds this, the existence of a mixture of distributions should be considered.
- **What should I do if I know that my data is not homogeneous in terms of positional accuracy?** This depends on your role and the assessment method. In the case of a user, it may not be relevant for accuracy control since it is only important that the data set meets the expected quality level. For a producer, if estimation is required, there will be a mixture of error populations, so the result will not strictly follow a normal distribution. Furthermore, in this case, it is of interest

to know the causes that have generated such a situation. In general, it is advisable to analyze each homogeneous set independently and to report the results of each independently.

- **Is it more important for me to estimate positional accuracy or to control it?** This depends on your role. If you are a user, control is sufficient. If you are a producer, you are interested in having your processes characterized and estimation is required for that. The producer is also interested in control.
- **Which positional accuracy assessment method is the most appropriate?** Each method (e.g., NMAS, EMAS, or NSSDA) has a different perspective, so you should consider which is the most relevant for your role before deciding to apply one or the other. In any case, as all methods[33] present different perspectives on the same reality, given that the cost of applying them with current software tools/capabilities is zero, several can be applied, giving a richer, more complete and complementary view of what is occurring with respect to positional accuracy.
- **What can I do if the error data I am working does not follow a normal distribution?** The best option in this case is to work with the empirically observed distribution using percentiles. In this case, the methods based on tolerances, such as the NMAS, the UNE 148002 standard or the method proposed by Ariza-López et al. (2018), which allows the use of several tolerances and error proportions. Another possibility is the use of a mixture of normal distributions.
- **Is metaquality really important in a positional accuracy assessment?** Of course. Without having knowledge of this the result of an assessment cannot be trusted. For example, should an assessment carried out with 5 points be trusted the same as one with 500? Should an assessment in which the assessment points are well distributed be trusted as much as one in which they are concentrated in one part of the project?
- **Is a positional accuracy assessment by points valid for other types of elements of the product, such as surfaces or lines?** The positional accuracy assessment by points is only valid for elements that have the same characteristics as those points, that is, that are well defined and easily identifiable, in which case they should be included in the assessment sample. Elements that do not meet these characteristics cannot be assigned these results as their own. Therefore, linear elements and fuzzy boundaries (e.g., some land covers) that do not have well-defined points can and should be assessed by other techniques (e.g., by using line-based methods). In a strict sense, the results of an evaluation are only valid for the typologies of the elements that make up the DQU.
- **What should I do with the distribution of the assessment points if my area of interest is not rectangular as shown in Figure 6?** Apply common sense. The intention of the distribution presented in Figure 6 is an example where the entire zone of the project is well covered, more or less, in a homogeneous manner. The same should be applied to any work area (e.g., an assessment of a roadway, which will be predominantly elongated).

---

[33] The three methods outlined here are not the only options; there are many others (see Ariza-López and Atkinson-Gordo 2008).

REFERENCES

Ariza-López FJ (2002). Calidad en la producción cartográfica. RA-MA, Madrid. ISBN 978-84-7897-524-2.

Ariza-López FJ (Editor) (2013). Fundamentos de evaluación de la calidad de la Información Geográfica. Universidad de Jaén. ISBN 978-84-8439-813-4.

Ariza-López FJ (2017). Documento P2.1.1: Observaciones sobre aspectos de evaluación de la calidad (especificaciones de un elemento de la calidad). Proyecto de Modernización de la Administración de Tierras en Colombia. Agencia de Implementación.

Ariza-López FJ, Atkinson AD (2008a). Analysis Of Some Positional Accuracy Assessment Methodologies. Surveying Engineering, 134(vol 2), pp. 404, 407. https://doi.org/10.1061/(ASCE)0733-9453(2008)134:2(45)

Ariza-López FJ, Atkinson AD (2008b). Variability of NSSDA estimations. Surveying Engineering, 134(vol 2), pp. 404-407. https://doi.org/10.1061/(ASCE)0733-9453(2008)134:2(39)

Ariza-López FJ, Atkinson AD, García-Balboa JL, Rodríguez-Avi J (2010). Analysis of User and Producer Risk when Applying the ASPRS Standards for Large Scale Maps. Photogrammetric Engineering & Remote Sensing, 5(8): 625-632(8). https://doi.org/10.14358/PERS.76.5.625

Ariza-López FJ, Atkinson AD, Rodríguez-Avi J (2008). Acceptance curves for the positional control of geographic data bases. Surveying Engineering. 134(vol 1), pp. 26,32. https://doi.org/10.1061/(ASCE)0733-9453(2008)134:1(26)

Ariza-López FJ, Rodríguez-Avi J (2014). A statistical model inspired by the National Map Accuracy Standard. Photogrammetric Engineering and Remote Sensing. 80 (3):271-281. https://doi.org/10.14358/PERS.80.3.271

Ariza-López FJ, Rodríguez-Avi J (2015). Using International Standards to Control the Positional Quality of Spatial Data. Photogrammetric Engineering & Remote Sensing. 81(8) https://doi.org/10.14358/PERS.81.8.657

Ariza-López FJ, Rodríguez-Avi J, Alba-Fernández V (2018). A Positional Quality Control Test Based on Proportions. In: Mansourian A., Pilesjö P., Harrie L., van Lammeren R. (eds) Geospatial Technologies for All. AGILE 2018. Lecture Notes in Geoinformation and Cartography. Springer, Cham. https://doi.org/10.1007/978-3-319-78208-9_18

Ariza-López FJ, Xavier E, Chicaiza EG, Buenaño X (2017). Métodos de evaluación de la calidad posicional en Hispanoamérica: análisis de la situación. Revista Cartográfica 94, 65-88. http://comisiones.ipgh.org/CARTOGRAFIA/rca/RCA94_Digital.pdf

ASCE (1983). Map Uses, scales and accuracies for engineering and associated purposes. American Society of Civil Engineers, Committee on Cartographic Surveying, Surveying and Mapping Division, New York, USA.

ASPRS (1990). Accuracy standards for large scale maps. PE&RS, vol. 56, nº7, 1068-1070.

ASPRS (2015). ASPRS Positional accuracy standards for digital geospatial data, Photogrammetric Engineering & Remote Sensing, Volume 81(3), 53 p. http://www.asprs.org/a/society/divisions/pad/Accuracy/Draft_ASPRS_Accuracy_Standards_for_Digital_Geospatial_Data_PE&RS.pdf

FGDC (1998). FGDC-STD-007: Geospatial Positioning Accuracy Standards, Part 3. National Standard for Spatial Data Accuracy. Federal Geographic Data Committee, Reston, USA. https://www.fgdc.gov/standards/projects/accuracy/part3/chapter3

IPGH (1978). Instituto Panamericano de Geografia e Historia: Especificaciones para mapas topográficos. Panamá: Instituto Panamericano de Geografía e Historia – IPGH, Panamá, 1978.

ISO (1985). ISO 2859-2:1985 Sampling procedures for inspection by attributes -- Part 2: Sampling plans indexed by limiting quality (LQ) for isolated lot inspection.

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

ISO (1999). ISO 2859-1:1999 Sampling procedures for inspection by attributes -- Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection.

ISO (2006a). ISO 3534-1:2006 Statistics -- Vocabulary and symbols -- Part 1: General statistical terms and terms used in probability

ISO (2006b). ISO 3534-2:2006 Statistics -- Vocabulary and symbols -- Part 2: Applied statistics.

ISO (2007). ISO 19131:2007 Geographic information -- Data product specifications.

ISO (2012). ISO 19158:2012 Geographic information -- Quality assurance of data supply

ISO (2013). ISO 19157:2013 Geographic information -- Data quality.

ISO (2014). ISO 19115-1:2014 Geographic information -- Metadata -- Part 1: Fundamentals.

ISO (2015a). ISO 19135-1:2015 Geographic information -- Procedures for item registration -- Part 1: Fundamentals.

ISO (2015b). ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary

Joint Committee for Guides in Metrology (2012). JCGM 200:2012. International vocabulary of metrology – Basic and general concepts and associated terms (VIM), 3rd edition. Working Group 2 of the Joint Committee for Guides in Metrology (JCGM/WG2).

JRC (2012) GMES Initial Operations emergency mapping validation service. Tender Specifications JRC/IPR/2012/G02/007/OC

Maune DF (Editor) (2007) Digital Elevation Model Technologies and Applications: The Dem User's Manual. American Society for Photogrammetry and Remote Sensing, Bethesda, ISBN 978-1-57083-082-2.

MPLMIC (1999). Positional Accuracy Handbook: Using the National Standard for Spatial Data Accuracy to measure and report geographic data quality. Minnesota Planning Land Management Information Center. https://www.mngeo.state.mn.us/pdf/1999/lmic/nssda_o.pdf

MSHA (2001). Survey field procedures manual. The Division of Plats and Surveys. Maryland State Highway Administration.

RAE (2021). *Diccionario de la lengua española*, 23.ª ed., REAL ACADEMIA ESPAÑOLA [versión 23.4 en línea]. <https://dle.rae.es> [Fecha de la consulta].

Sevilla MJ (1991). Criterios de precisión cartográfica. Instituto de Astronomía y Geodesia. 1991, nº 182.

UNE (2016). UNE 148002:2016 Metodología de evaluación de la exactitud posicional de la información geográfica. UNE, Madrid.

USBB (1947). United States National Map Accuracy Standards. U.S. Bureau of the Budget. Washington, USA.

Zandbergen PA (2011). Characterizing the error distribution of Lidar elevation data for North Carolina. International Journal of Remote Sensing 32(2):409-430. https://doi.org/10.1080/01431160903474939

Zandbergen PA (2008). Positional Accuracy of Spatial Data: Non-Normal Distributions and a Critique of the National Standard for Spatial Data Accuracy. Transactions in GIS 12(1):103–130. https://doi.org/10.1111/j.1467-9671.2008.01088.x

IN MEMORIAN

Remaining in our memory, our gratitude to Edison Rojas, our colleague and friend, who requested this project and who has left us without seeing its fruit, which is also due to his effort and leadership.

ACKNOWLEDGEMENT

To the Pan American Institute of Geography and History, especially to the Cartography and Geography commissions for partially financing this project.
To the Research Plan of the University of Jaén and the Andalusian Research Plan, for financing the activities of the TEP 164 Research Group.

ANNEX 1. EXAMPLE OF PLANIMETRIC POSITIONAL ACCURACY ASSESSMENT STANDALONE REPORT OF AN ORTOPHOTOGRAPHY

This annex presents an example of a standalone report about the positional accuracy assessment of an ortophotography. This dataset has been produced by the Servicio Aerofotogramétrico de la Fuerza Aérea de Chile (Chilean Air Force Aerophotogrammetric Service) for the «Ilustre Municipalidad de Quilicura-Santiago de Chile» (municipality of Quilicura-Santiago de Chile). It is an extensive document in the line of the standalone quality report proposed in the ISO 19157 standard. As indicated in this standard, the standalone report never replaces the obligation to report in the form of metadata.

Assessment results are presented for the standards NMAS, EMAS and NSSDA. It is not usual to present the results for several standards, but given that this is an example report, it has been considered appropriate to offer a broader view. Furthermore, the different standards represent different perspectives and their computing cost is zero. Therefore, it is also a good option.

The contents and structure of the report follow what is presented in Table 13 of this document. Remember that the data set to be assessed is denoted by DSA and that the reference data set, with higher accuracy, is denoted by RDS.

Some descriptive comments are included in italics. They help to understand how this example has been developed.

This annex is centered on the report of the assessment process and it aims to be exhaustive and rigorous. But to apply the assessment process itself, well-described procedures are required in relation to critical processes. In this report the organization is considered to have implemented the following procedures:

- Proc1. Procedure for fieldwork with GNSS techniques.
- Proc2. Procedure for coordinate capture of assessment points over images.

Further, it is considered that the following software tools are available:

- ST1. Tool for generating random and well distributed positions in a geographic scope.

Finally, it should be noted that all statistical test are carried out with a significance level of 5%.

Given the formative character that we want to give to this example, we have included an addendum at the end which includes some comments which analyze this real case.

1. IDENTIFICATION OF THE DSA TO BE ASSESSED

This section must give information to identify the DSA as well as its main features such as name, identifier, date or version, producer, a qualitative description, specifications (i.e. with parameters such as scale, resolution, CRS), and design accuracy (for each component, x, y, z or joint component xy).

It is suggested to present this information in a clear and concise form in a tabular format, as in Table A1.1.

Table A1.1 Identification of the DSA to be assessed

Name	Orthorectified mosaic of the commune of Quilicura, Metropolitan Region of Santiago			
ID	SAF-OFM-001			
Producer	Servicio Aerofotogramétrico de la Fuerza Aérea de Chile			
Qualitative description	Orthorectified mosaic, obtained from an aerophotogrammetric survey from the community of Quilicura (Metropolitan Region of Santiago), performed in 2010.			
Purpose				
Specifications	Resolution (GSD, *Ground Sample Distance*): 15 cm Scale equivalent to 1:2000 Coordinate Reference System (CRS): EPSG 5361			
Design (theoretical) accuracy	☒ XY (m) MSE: 0.71 m μ = 0 m, σ = 0.5 m	☐ X (m) MSE: 0.5 m μ = 0 m, σ = 0.5 m	☐ Y (m) MSE: 0.5 m μ = 0 m, σ = 0.5 m	☐ Z(m) MSE: ______m μ = ____m, σ = ___m

2. GENERAL ASPECTS OF THE ASESSMENT

In the framework of the ISO 19157 standard, the data quality unit (DQU) should first be specified. The DQU is composed of the data quality element and the scope of the quality assessment. Information should be also included about the data quality measures to be used, the conformity levels (if applicable) and the assessment method (a summary).

It is suggested to present this information in a clear and concise form in a tabular format, as is shown in Table A1.2.

For convenience, Note that in Table A1.2 a list has been included with all eligible measures from the ISO 19157 standard. An ordinal identifier has been assigned to each measure in order to be cited in later sections of this report. The list includes all measures related to positional accuracy from annex D of the ISO 19157 standard, indicating their identifier and name. As an example, the measures with ordinals 1, 2, 16, 19 and 21 have been marked. The first four are merely informative since no conformity levels have been established. A conformity level has been established for the last one.

It has been also considered to apply the specific positional quality standards NMAS, EMAS and NSSDA. Following what is indicated by the ISO 19157 standard in relation with user-defined data quality measures, it is recommended to perform an external document which specifies measures based on these standards. In that case they could be added to the list of eligible measures.

Tabla A1.2 General aspects of the assessment

Data quality unit	**Data quality element**	☐ Absolute or external accuracy	☐ Relative or internal accuracy

(DQU)	**Data quality scope**	Spatial: zone defined by the administrative boundary of the community of Quilicura. Thematic: not applicable			
Data quality measures	**Ordinal**	**Source**	**Identifier**	**Name**	**Conformity level**
	☒ 01	ISO 19157	28	Mean value of positional uncertainties X ☐ Y ☐ Z ☐ XY ☒ XYZ ☐	-
	☒ 02	ISO 19157	128	Bias of positions X ☐ Y ☐ Z ☐ XY ☒ XYZ ☐	-
	☐ 03	ISO 19157	29	Mean value of positional uncertainties excluding outliers X ☐ Y ☐ Z ☐ XY ☐ XYZ ☐	-
	☐ 04	ISO 19157	30	X ☐ Y ☐ Z ☐ XY ☐ XYZ ☐	-
	☐ 05	ISO 19157	31	Rate of positional uncertainties above a given threshold X ☐ Y ☐ Z ☐ XY ☐ XYZ ☐	-
	☐ 06	ISO 19157	32	Covariance matrix X ☐ Y ☐ Z ☐ XY ☐ XYZ ☐	-
	☐ 07	ISO 19157	33	Z: Linear error probable	-
	☐ 08	ISO 19157	34	Z: Standard linear error	-
	☐ 09	ISO 19157	35	Z: Linear map accuracy at 90% significance level	-
	☐ 10	ISO 19157	36	Z: Linear map accuracy at 95% significance level	-
	☐ 11	ISO 19157	37	Z: Linear map accuracy at 99% significance level	-
	☐ 12	ISO 19157	38	Z: Near certainty linear error	-
	☐ 13	ISO 19157	39	Z: Root mean square error	-
	☐ 14	ISO 19157	40	Z: Absolute linear error at 90% significance level of biased vertical data (alternative 1)	-
	☐ 15	ISO 19157	41	Z: Absolute linear error at 90% significance level of biased vertical data (alternative 2)	-
	☒ 16	ISO 19157	42	XY: Circular standard deviation	-
	☐ 17	ISO 19157	43	XY: Circular error probable	-
	☐ 18	ISO 19157	44	XY: Circular error at 90% significance level	-
	☒ 19	ISO 19157	45	XY: Circular error at 95% significance level	-
	☐ 20	ISO 19157	46	XY: Circular near certainty error	-
	☒ 21	ISO 19157	47	XY: Root mean square error of planimetry	≤ 0.25 m
	☐ 22	ISO 19157	48	XY: Absolute circular error at 90% significance level of biased data	-
	☐ 23	ISO 19157	49	XY: Absolute circular error at 90% significance level of biased data	-
	☐ 24	ISO 19157	50	XY: Uncertainty ellipse	-
	☐ 25	ISO 19157	51	XY: Confidence ellipse	-
	☐ 26	ISO 19157	52	Z: Relative vertical error	-
	☐ 27	ISO 19157	53	XY: Relative horizontal error	-

Data quality evaluation method

Type of method: direct external

Inspection method: ☐ full inspection ☒ sampling

Description:

A random and suitably-sized sample of homologous points between the DSA and the CRS is generated. These points are well defined and clearly identifiable, therefore their discrepancy values result in a sample of planimetric positional errors in X and Y.

For the capture of the coordinates X and Y of the sample the guidelines for fieldwork specified in the document *Proc1* have to be taken into account as well as the guidelines for digitizing in the laboratory specified in the document *Proc2*.

Regarding the sample of errors:

- The basic statistical hypotheses are verified: randomness, outlier values treatment, normality, bias, correlation and homoscedasticity.
- Basic statistics are computed for each component X, Y.
- The chosen measures are computed.
- The following positional quality standards are applied:

  ☒ NMAS ☒ EMAS ☒ NSSDA

Note: more information about the sampling is offered in the section "highest accuracy source and list of coordinates"

3. HIGHEST ACCURACY SOURCE AND LIST OF COORDINATES

This part of the report is very important from the perspective of metaquality. Thus, contents should be included which allow us to appreciate this metaquality. It is therefore a critical section within the report.

Details about the establishment of the sample of assessment points should be provided. Users must be informed about: the source of higher accuracy, the sample size, if the sample is 1D, 2D or 3D, the positional accuracy of the RDS, the types of features generally used to locate the assessment points, how the randomness of the sample is guaranteed, what the method is to select the specific location of an assessment point, the capture method of the field coordinates, etc. It would be desirable to have external documents which specify the guidelines and recommendations to follow and to cite them in this section. For example, documents about choice of sampling points, execution of the fieldwork, etc.

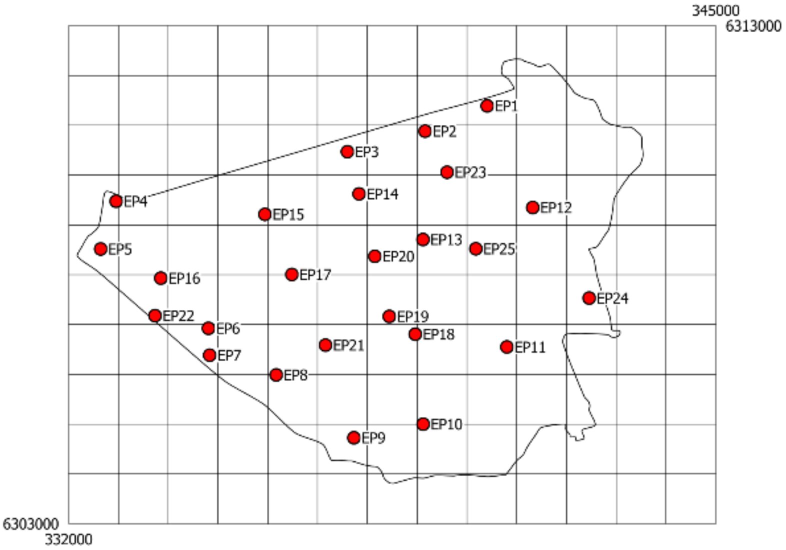
It is suggested to present this information in a clear and concise form in a tabular format, as in Table A1.3.

In this section a list of coordinates of the DSA and the RDS is also included, with which the calculation and subsequent analysis are going to be performed. All data clean of any mistakes will be included and before carrying out any filtering process, such as the elimination of outliers. Also observations about the points could be included (i.e., the feature type which corresponds to each point).

Table A1.3 Highest accuracy source and list of coordinates

Reference source	Fieldwork performed with GNSS techniques		
Dimension	☒ Planimetry (2D: XY)		☐ Altimetry (1D: Z)
Reference accuracy	$MSE_X = MSE_Y = 0.05$ m ☐ ×2 ☐ ×3 ☐ ×4 ☒ ×5 or better		$MSE_Z =$ ___ ☐ ×2 ☐ ×3 ☐ ×4 ☐ ×5 or better
Coverage	**Population**	Sample size:	25
	Thematic	Feature types used for the assessment: • Corners in paintings over the asphalt. • Corners in grass (garden areas). • Corners in sidewalks. • Sidewalks poles. • Corners in manhole covers. • Other geometric shapes (circles, squares) in the image, with small size and well-contrasted. Note: as the product to be assessed is an image, there are no thematic layers	

Spatial



Interoperability

☒ It has been checked that the DSA and the RDS use the same CRS.

Other aspects

Assurance of sample randomness:

A random sample of planimetric locations (points) was generated by means of the tool ST1, within the scope defined by the DQU.

Method for the selection of the assessment points:

Over the DSA, which is an image in this case, the operator searches for clearly identifiable and well-defined characteristics in objects which appear in the ortophotography in the immediate surrounding of each point.

This clearly identified point is chosen taking into account that it also must be clearly identified at fieldwork and that its location must be measured with GNSS equipment.

The guidelines for digitizing in the laboratory specified in the document *Proc2* include recommendations for the selection of assessment points.

Aspects of the coordinate capture in the laboratory:

In order to mitigate the influence of the interpretation over the image the coordinates of each point have been digitized twice, each time by a different operator. The coordinates are averaged after checking that the discrepancy in the coordinates has not exceeded a preset threshold. Further information is provided in the document *Proc2*.

Aspects of the coordinate capture in the field:

Once the assessment point is identified in the field, the appropriate measurements are taken with GNSS equipment by means of the rapid static method. Coordinates are later determined in the laboratory.

It was always ensured that the CDS was the same as that of the DSA.

For coordinate capture in the field the guidelines specified in the document *Proc1* must be taken into account.

Coordinate list [m]

Id	**RDS**			**DSA**			**Observations**
	X_C	**Y_C**	**Z_C**	**X_P**	**Y_P**	**Z_P**	
EP1	340408,214	6311389,779		340408,133	6311389,518		vertex of pathway
EP2	339160,180	6310882,149		339160,120	6310882,029		vertex of pathway
EP3	337599,246	6310470,578		337599,066	6310470,450		corner of white square
EP4	332953,628	6309475,489		332953,635	6309475,536		center of black spot
EP5	332643,438	6308517,124		332643,510	6308517,006		outer limits of white spot
EP6	334808,520	6306924,530		334808,462	6306924,604		center of black spot
EP7	334833,602	6306384,681		334833,649	6306384,762		vertex of pathway

	EP8	336170,053	6305989,103		336170,163	6305989,343		vertex of grass
	EP9	337730,216	6304726,637		337730,083	6304726,500		ending of white line
	EP10	339121,724	6304999,681		339121,561	6304999,505		vertex of black manhole cover
	EP11	340803,068	6306550,788		340803,028	6306550,487		vertex of white line
	EP12	341322,568	6309348,265		341322,460	6309348,020		vertex of manhole cover
	EP13	339117,805	6308706,981		339117,777	6308707,731		vertex of white line
	EP14	337833,715	6309623,811		337833,519	6309623,851		ending of white line
	EP15	335942,644	6309212,095		335942,607	6309212,270		center of white spot
	EP16	333851,880	6307931,556		333851,612	6307931,552		center of rock
	EP17	336484,559	6308005,303		336484,414	6308005,200		vertex of sidewalk
	EP18	338963,237	6306806,313		338963,090	6306806,178		ending of white line
	EP19	338441,394	6307163,523		338441,460	6307163,409		vertex of concrete
	EP20	338151,037	6308369,663		338150,882	6308369,558		lower vertex of letter “L”
	EP21	337159,639	6306586,164		337159,566	6306586,055		vertex of grass
	EP22	333735,148	6307177,903		333735,214	6307178,112		center of white spot
	EP23	339601,413	6310058,623		339601,228	6310058,442		vertex of grass
	EP24	342456,920	6307531,681		342456,662	6307531,515		inner vertex of white lines
	EP25	340180,126	6308521,032		340179,975	6308520,752		ending of white line

4. STATISTICAL HYPOTHESIS TESTING

This section of the report focuses on evidence that the hypotheses required by the statistical analysis method are satisfied, whether they are implicit or explicit. These checks can be carried out using any appropriate tool, such as a general statistical software (i.e. SPSS), specific developments for application in positional quality (i.e. R codes), spreadsheets with special-purpose scripts, etc. The list of positional errors on which the checks will be carried out must be included.

It is suggested to present this information in a clear and concise form in a tabular format, as shown in Table A1.4. *Although they have not been included in the example, graphics could be included for better clarity of the checks carried out. At the end of the table a section has been added for result interpretation and to conclude whether or not to continue with the quality assessment process.*

Table A1.4 Statistical hypothesis testing

Error list **[m]**	**Id**	**E_X**	**E_Y**	**E_Z**	**Outliers**				
					X	**Y**	**Z**	**Point***	**Possible cause**
	EP1	-0,080526	-0,26086		☐	☐	☐	☐	
	EP2	-0,060268	-0,11973		☐	☐	☐	☐	
	EP3	-0,180206	-0,12778		☐	☐	☐	☐	
	EP4	0,006575	0,04734		☐	☐	☐	☐	
	EP5	0,071553	-0,11807		☐	☐	☐	☐	
	EP6	-0,057508	0,07409		☐	☐	☐	☐	
	EP7	0,047297	0,08075		☐	☐	☐	☐	
	EP8	0,109683	0,23957		☐	☐	☐	☐	
	EP9	-0,133466	-0,13742		☐	☐	☐	☐	
	EP10	-0,162899	-0,17565		☐	☐	☐	☐	
	EP11	-0,039788	-0,30098		☐	☐	☐	☐	
	EP12	-0,107757	-0,2448		☐	☐	☐	☐	
	EP13	-0,027503	0,7504		☐	☒	☐	☒	
	EP14	-0,195867	0,04046		☐	☐	☐	☐	
	EP15	-0,037015	0,17523		☐	☐	☐	☐	
	EP16	-0,268448	-0,00376		☐	☐	☐	☐	

	EP17	-0,14543	-0,10263		☐	☐	☐	☐	
	EP18	-0,147036	-0,13499		☐	☐	☐	☐	
	EP19	0,06567	-0,11428		☐	☐	☐	☐	
	EP20	-0,155278	-0,10457		☐	☐	☐	☐	
	EP21	-0,073324	-0,10916		☐	☐	☐	☐	
	EP22	0,066459	0,20853		☐	☐	☐	☐	
	EP23	-0,185416	-0,18078		☐	☐	☐	☐	
	EP24	-0,258211	-0,1659		☐	☐	☐	☐	
	EP25	-0,151007	-0,27981		☐	☐	☐	☐	

* A point is considered an outlier when any of the components is evaluated as an outlier.

Randomness

Test: Wald–Wolfowitz runs test

Level of significance: 5 %

Null hypothesis: the sequence of errors is random
Note: In the error list the sequential order of the random sample of planimetric positions generated by the tool ST1 has not been modified.

X	p-value: 0,6764	☒ Not rejected	☐ Rejected	☐ Not checked
Y	p-value: 0,0950	☒ Not rejected	☐ Rejected	☐ Not checked
Z	p-value: ____	☐ Not rejected	☐ Rejected	☒ Not checked

Outliers

Test: Outlier detection in a normal distribution. Coverage factor: k = 3

X	☒ Checked	☐ Not checked	Outlier points: EP13
Y	☒ Checked	☐ Not checked	
Z	☐ Checked	☒ Not checked	

Comment:

The outlier points should be analyzed in detail to determine the underlying cause. In any case, they must be removed from the remaining statistical hypothesis tests based on normality.

The outlier points are included in the figure of spatial distribution of error in section 5, for possible analysis.

Normality

Test: Kolmogorov–Smirnov test (table of critical values)

Level of significance: 5 %

Null hypothesis: the population is normally distributed

X	p-value: ____	☒ Not rejected	☐ Rejected	☐ Not checked
Y	p-value: ____	☒ Not rejected	☐ Rejected	☐ Not checked
Z	p-value: ____	☐ Not rejected	☐ Rejected	☒ Not checked

Bias

Test: one-sample t-test for a mean value μ

Level of significance: 5%

Null hypothesis: the mean of the population is equal to μ.

X μ = 0	p-value: ____	☐ Not rejected	☒ Rejected	☐ Not checked
Y μ = 0	p-value: ____	☐ Not rejected	☒ Rejected	☐ Not checked
Z μ =	p-value: ____	☐ Not rejected	☐ Rejected	☒ Not checked

Independence

Test: Spearman's rank correlation coefficient

Level of significance: 5%

Null hypothesis: both populations are independent one another

X - Y	p-value: 0,042	☐ Not rejected	☒ Rejected	☐ Not checked

	Pearson correlation coefficient: 0,454	
	X - Y	Interpretation: moderate positive correlation
Homoscedasticity	Test: Bartlett's test for homogeneity of Variances Level of significance: 5 % Null hypothesis: both population variances are equal	
	X - Y	p-value: 0,098 ☒ Not rejected ☐ Rejected ☐ Not checked
Interpretation of the tests	In general the sample complies with the statistical hypotheses of randomness, normality and homoscedasticity. As for independence there is a moderate correlation. Components X and Y do not pass the bias test. This test is influenced by the standard deviation values, much lower than the design accuracy. Therefore these results concerning bias are not considered as relevant. Of the total 25 points in the sample, 1 is considered an outlier. This represents 4% of the total sample, which is a significant proportion. Nevertheless, the small sample size limits the obtaining of sound conclusions. Conclusion: ☒ DO proceed to the results section ☐ DO NOT proceed to the results section	

5. RESULTS

Firstly, the definitive error list should be included (once the outliers have been removed, if applicable). All subsequent results will be computed from this list. It is recommended to attach to it some descriptive statistics. Next some charts can be included to help with the interpretation of the spatial and statistical distribution of the errors, such as: circular diagram of distribution of the planimetric error components (X, Y), box plot of each error component, histogram of each error component, chart with the spatial distribution of the errors, etc. If there is bias, and it has been assigned, it should be explained.

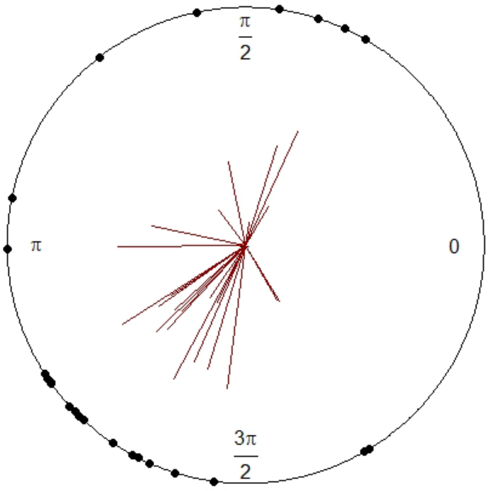
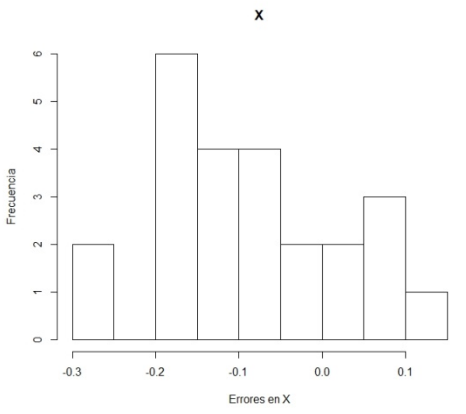
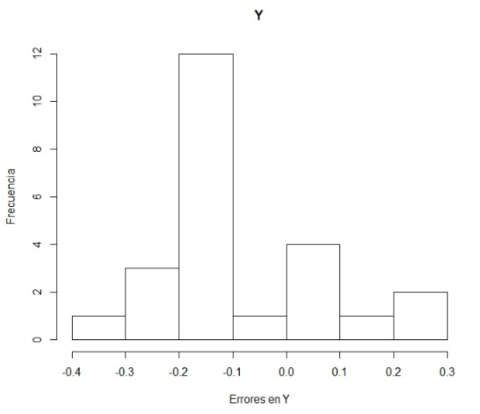
Then all results for the measures selected in section 2 should be presented. If a conformity level has been established for any measure, it should be indicated whether the measure result is compliant or not. Next all results for specific standards concerning positional quality (NMAS, EMAS and NSSDA) should be included.

Finally, it is convenient to introduce a brief interpretation of all results together.

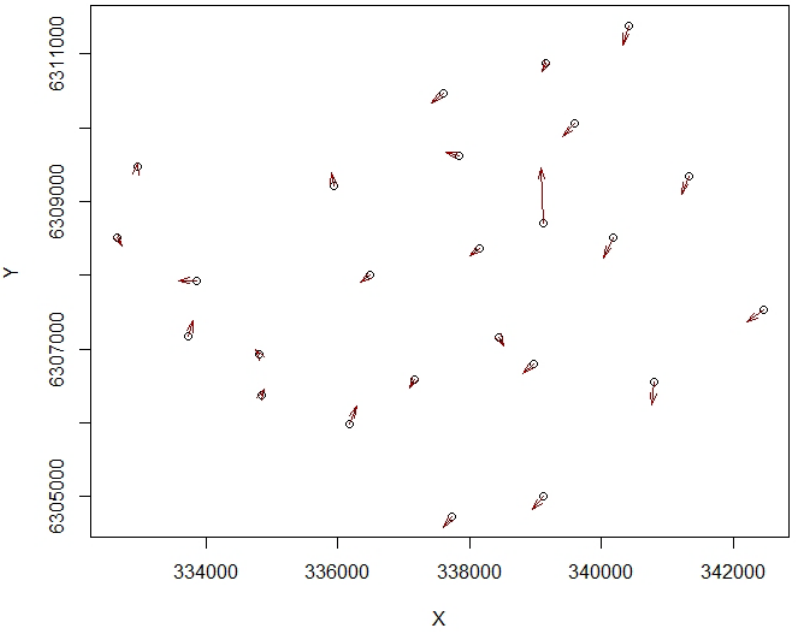
It is suggested to present this information in a clear and concise form in a tabular format, as shown in Table A1.5.

Table A1.5 Results

Definitive error list [m]	**Id**	**E_X**	**E_Y**	**E_Z**	**Circular diagram of distribution of the errors X,Y**
	EP1	-0,081	-0,261		
	EP2	-0,060	-0,120		
	EP3	-0,180	-0,128		
	EP4	0,007	0,047		
	EP5	0,072	-0,118		
	EP6	-0,058	0,074		
	EP7	0,047	0,081		
	EP8	0,110	0,240		
	EP9	-0,133	-0,137		

	EP10	-0,163	-0,176		
	EP11	-0,040	-0,301		
	EP12	-0,108	-0,245		
	EP14	-0,196	0,040		
	EP15	-0,037	0,175		
	EP16	-0,268	-0,004		
	EP17	-0,145	-0,103		
	EP18	-0,147	-0,135		
	EP19	0,066	-0,114		
	EP20	-0,155	-0,105		
	EP21	-0,073	-0,109		
	EP22	0,066	0,209		
	EP23	-0,185	-0,181		
	EP24	-0,258	-0,166		
	EP25	-0,151	-0,280		
Descriptive statistics	**Mean**	-0,086	-0,076		**Diagram of the distribution of errors Z**
	Stand. dev.	0,106	0,151		
	MSE	0,134	0,166		
	Minimum	-0,268	-0,301		
	Maximum	0,110	0240		
	Median	0,071	-0,116		
	Perc 95 (abs)	0,249	0,277		
Histogram of errors					

Spatial distribution of errors (errors field)



Note 1: This chart includes outliers removed in section 4.

Assignment of bias

Data quality measures

Ordinal	Result	Compliant	Ordinal	Result	Compliant	Ordinal	Result	Compliant
1	0,202 m	-	11			21	0,214 m	Yes
2	0,115 m	-	12			22		
3			13			23		
4			14			24		
5			15			25		
6			16	0,129 m	-	26		
7			17			27		
8			18					
9			19	0,315 m	-			
10			20					

NMAS

Horizontal

Tol_hz = 1,693 m — % points > Tol_hz = 0% — ☐ Not checked

☒ Pass ☐ Fail

Vertical

Tol_vert = — % points > Tol_vert = — ☒ Not checked

☐ Pass ☐ Fail

EMAS

X

Bias

Mean limit: 0 m — α = 5 %

p-value: ______ — t_x = -3,974 — $t_{n-1,\alpha/2}$= 2,080

☐ Pass ☒ Fail

Disper.

Standard deviation limit: 0,5 m — α = 5 %

p-value: ______ — χ^2_x = 1,042 — $\chi^2_{n-1,\alpha}$ = 32,671

☒ Pass ☐ Fail

☐ Not checked

Y

Bias

mean limit: 0 m — α = 5 %

p-value: ______ — t_y = -2,450 — $t_{n-1,\alpha/2}$= 2,080

☐ Pass ☒ Fail

Disper.

Standard deviation limit: 0,5 m — α = 5 %

p-value: ______ — χ^2_y = 2,105 — $\chi^2_{n-1,\alpha}$ = 32,671

☒ Pass ☐ Fail

☐ Not checked

	Z	Bias	Mean limit: ____ α = p-value: _____ t_z = $t_{n-1,\alpha/2}$ = ☐ Pass ☐ Fail	☒ Not checked
		Disper.	Standard deviation limit: α = p-value: _____ χ_Z^2 = $\chi_{n-1,\alpha}^2$ = ☐ Pass ☐ Fail	
	TOTAL		☐ Pass ☒ Fail	☐ Not checked
NSSDA	Horizontal		ECM_X = 0,135 m ECM_Y = 0,167 m ECM_{min} / ECM_{max} = 0,81 $NSSDA_H$ = 0,369 m Tested 0,369 meters horizontal accuracy at 95% confidence level.	
	Vertical		ECM_Z = $NSSDA_Z$ = Tested _____ meters vertical accuracy at 95% confidence level.	
	Note: Error data is not bias-free, therefore the NSSDA results should be considered with some caution.			
Interpretation of the results	**Theoretical quality.** a circular standard deviation is assumed of 0,5 m (sigma-c = sigma-x = sigma-y = 0,5 m) and without any bias in any component. **Statistical hypotheses.** The sample of errors passes most statistical hypotheses, which gives us confidence in the work performed. Bias is detected for both components, which is influenced by the low standard deviation values of the errors. One outlier is removed, so the sample size is reduced from 25 to 24 points. The cause of the outlier is unknown. **Descriptive statistics.** The proximity between the central tendency (mean and median) and the dispersion (standard deviation) statistics seems to point to the presence of some bias. Nevertheless, the maximum values (0,25 m in Y), minimum values (-0,30 m en Y) and the percentile 95 (0,249 m in X and 0,277 in Y) suggest that the errors are lower than expected for such a product with the specified theoretical quality (standard deviation in each component of 0,5 m). **Standard NMAS.** The standard is passed, since no point exceeds the threshold of 1,693 m. This is a logical result for a product with a theoretical standard deviation of 0,5 m in each component. **Standard EMAS.** The standard shows that the tests in both components X and Y are passed for variance but not for the mean. This confirms what was expected from the proximity between the mean and the standard deviation values. **Standard NSSDA.** This standard, which assumes that there is no bias, returns a value of 0,369 m at 95% confidence level. This is equivalent to a circular standard deviation of 0,369/2,4477 = 0,15 m, much lower than the theoretical quality. Therefore it can be stated than the product is better than expected. **Summary.** The product has a global quality better than the expected quality (0,5 m). which is confirmed by the standard NSSDA, which is the standard that returns a numeric and interpretable value. Nevertheless, at a more detailed level with the standard EMAS and the descriptive statistics, it is noticed that bias could exist in both the X and Y components. Bias and dispersion combined hide the problem when applying NSSDA, whose value has not been influenced by the bias. The standard NMAS is obviously passed since the threshold is 1,7 m, far above the measured errors.			

6. METAQUALITY OF RESULTS AND PROCESSES

This section of the report is dedicated to supporting explanations regarding metaquality elements. These should be based on objective facts presented in the previous sections of the report. It is suggested to follow the recommendations given in the UNE 148002 standard.

It is suggested to present this information in a clear and concise form in a tabular format, as shown in Table A1.6.

Table A1.6 Metaquality of results and processes

Confidence	**Qualitative description** Confidence in the assessment work is assured by: i) random generation of the sample by means of a random location generator (tool *ST1*); ii) sufficient sample size and somewhat larger than that usually required by the standards NMAS, EMAS and NSSDA; iii) coordinate capture in the field by means of GNSS carrier phase observable and the rapid static method; iv) GNSS observations processing by a specialist in geodetic computation; v) independence between the DSA and the RDS since they are not used in any operation common to both datasets; vi) personnel with specific training and more than 5 years of experience in positional quality assessment have participated in all phases **Quantitative data**: Sample size: >20 Highest accuracy source: >5x
Homogeneity	**Qualitative description.** There are quality management measures in place to assure the homogeneity of the assessment process: • Standards on training and qualification of the intervening personnel • Written procedures: *Proc1* with guidelines for fieldwork. *Proc2* with guidelines for digitizing in the laboratory. All the personnel involved have been trained in this type of work. The Chilean Air Force Aerophotogrammetric Service has a procedure manual and an ISO 9001 QMS (quality management system) for this type of product.
Representativity	**Spatial**. The spatial distribution is homogeneous and covers all the scope defined in the data quality unit. **Thematic**. Not applicable for an image product. **Population**. The sample size is sufficient. The hypothesis of normality of the errors in each component X, Y is not rejected.

7. DATE AND SIGNATURE

Any assessment should be assigned with a date and a responsible technician, who should sign the report.

Date of the report	28 March 2019
Signature of the responsible technician	PAC JOSELYN ROBLEDO CEBALLOS

ADENDUM: COMMENTS ABOUT THE ASSESSMENT

As indicated in the initial presentation of this annex, we would like to introduce some comments on the report itself and on the real case of assessment which has been shown.

In relation to the report, we would like to emphasize that it is only a proposal; everyone can take from it those aspects that are more convenient to them. The presented report may seem long, but it should be taken into account that results have been included for the required statistical hypotheses, different data quality measures, three positional quality standards and a part dedicated to metaquality. Furthermore, a spacious and verbose form of presentation has been adopted, all of which has notably increased its length. However, we believe that this type of report should be sufficiently comprehensive and clear, rather than synthetic and summary.

The report has also tried to be very visual, so an assessment points distribution chart with vector errors and a circular diagram of distribution of errors X, Y, histograms, etc. have been included. This is not very common in this type of report but its inclusion enriches the report and furthermoreexpands the analysis possibilities it offers.

Regarding the assessment carried out, it is a real case that presents some aspects that should be indicated since the objective of this annex is to present this type of report, not a theoretical perfect case of assessment of positional accuracy. Some relevant aspects are:

- Source of higher accuracy. The spatial distribution of the assessment points is not entirely adequate since there are areas which are not covered sufficiently. Furthermore, as indicated in the section "number and distribution of assessment points", the assessment points should cover the area of interest with a certain buffer, but this is not the case. In relation to the feature types used in the assessment, these should mostly include those from the DSA to be assessed. On the other hand, the use of paint on asphalt is only recommended if it is very recent and there is little lag between the painting time, the photogrammetric data capture and the GNSS fieldwork. Furthermore, this would not allow a stable set of assessment points to be generated for future quality assessment work. In order to highlight the rigorous nature of this study a reference to the ST1 tool and the Proc2 procedure has been introduced. These are necessary for this process to be carried out properly and with sufficient rigor.

- Statistical hypotheses checking. As can be observed, the randomness of the errors in Y is close to being rejected and its value is far from that obtained for the errors in X. This indicates that there may possibly be some special circumstance in the Y component of the errors. The normality is close to being rejected and there is some bias. But since the theoretical deviation of 0,5 m per component is higher than the compositions of bias and deviations, it is not a problem in practice to assume them. The analysis of independence points out that there is some correlation between the values of X and Y errors. In the analysis of homoscedasticity the null hypothesis is accepted with a value near to being rejected. In conclusion, the error data in X and Y of this real case do not comfortably meet the required statistical hypotheses for most of the positional quality assessment methods. The possible causes should be discerned by analyzing the processes that have intervened in obtaining the data. The producer of this assessment should analyze whether this has been a standard or anomalous process within the organization. In any case, given that we wanted to present a complete example, it is concluded to proceed with the evaluation

- Results. This section confirms the findings indicated in the previous paragraph. We want to highlight the inclusion of both a circular diagram of distribution of the errors and a chart with the spatial distribution of the errors field. Both are interesting and the information presented is complementary. It is observed in the circular diagram that there is a higher concentration of cases in the lower left quadrant (which is the aforementioned bias) and that these cases tend to be concentrated in the eastern zone (approximately from $X \geq 339000$). The rest of the map presents a more random behavior. This shows that these two graphic tools facilitate a better understanding of the reality of what is happening. From there, the possible assignable causes should be found.

- Metaquality. This section has been included with a clear informative intention and its content aims to offer an example of what can be included, since the assessment that has been presented does not allow a satisfactory metaquality report to be generated. We consider that reporting on metaquality is a moral obligation of those who perform quality evaluations because if there is no confidence in their methods, they should not have confidence in their results. Thus, in this case the sample size is small and its distribution is incomplete since it does not adequately cover the spatial scope. In addition, the existence of the ST1 tool has been assumed to create the random distribution of assessment points, but actually the distribution and selection have been carried out by an operator. Similarly, the existence of two procedures (Proc1 and Proc2) has been assumed to support the report. One aspect of confidence is that the Chilean Air Force Aerophotogrammetric Service has a QMS certified according to ISO 9001 that covers production.

186 MEMBER STATES

Argentina

Belice

Bolivia

Brasil

Chile

Colombia

Costa Rica

Ecuador

El Salvador

United states

Guatemala

Haití

Honduras

México

Nicaragua

Panamá

Paraguay

Perú

República Dominicana

Uruguay

Venezuela

187 PAN AMERICAN INSTITUTE OF GEOGRAPHY AND HISTORY

EL IPGH, SUS FUNCIONES Y ORGANIZACIÓN

The Pan American Institute of Geography and History was founded on February 7, 1928 by resolution approved at the Sixth International American Conference that took place in Havana, Cuba. In 1930, the Government of the United Mexican States built, for the use of the PAIGH, the building in the "Calle Ex Arzobispado 29", Tacubaya, in Mexico City.

In 1949, an agreement was signed between the Institute and the Council of the Organization of American States and it became the first specialized body of it.

The Statute of the PAIGH cites in its article 1 its aims:

1) Promote, coordinate and disseminate cartographic, geophysical, geographical and historical studies, and those related to related sciences of interest to America.
2) Promote and carry out studies, work and training in these disciplines.
3) Promote cooperation between the Institutes of their disciplines in America and with related international organizations.

Only the American States can be members of the PAIGH. There is also the category of Permanent Observer, currently they are under this condition: Spain, France, Israel and Jamaica.

The PAIGH is made up of the following Pan-American bodies:

1) General Assembly.
2) Board of Directors.
3) Commission of:
    Cartography (Costa Rica)
    Geography (United States of America)
    History (Mexico)
    Geophysics (Ecuador)
4) Meeting of Authorities.
5) General Secretariat (Mexico, D.F., Mexico).

In addition, in each Member State there is a National Section whose components are appointed by each government. They have their President, Vice President, National Members of Cartography, Geography, History and Geophysics.

188

METADATA

Title	**GUIDE FOR THE POSITIONAL ACCURACY ASSESSMENT OF GEOSPATIAL DATA**
Authors	Ariza López, F.J.; García Balboa, J.L.; Rodríguez Avi, J.; Robledo, J.
Subject	Geographic Information, geospatial data production, data quality, positional accuracy
Description	An applied guide is provided to allow the development of assessments of absolute positional accuracy in a correct and reliable manner. Theoretical and practical aspects are discussed together. The entire process is developed within the framework established by the ISO 19100 standards of Technical Committee 211. As a complement, a report template is proposed that summarizes the characteristics of the evaluation carried out, as well as the results obtained. This guide focuses primarily on the processes to be developed in the office, but also offers guidelines for data collection processes in the field.
Publisher	Pan American Institute of Geography and History
Collaborators	Florencia Manduca (Argentina), Emerson Magnus de Araújo Xavier (Brasil), Pablo Morales (Chile), Elena Gabriela Chicaiza (Ecuador), Xavier Buenaño (Ecuador), Héctor Gómora (México), Rosario Casanova (Uruguay), Edison Rosas (Uruguay), Hebenor Bermúdez (Uruguay)
Date	Translated in April 2021
Type	Text
Format	PFD
Identification	Occasional Publications 561
Source	PAIGH publication #557
Language	English
Relation	
Coverage	International
Rights	PAIGH (publication), Authors (contents)
Audience	Technicians using geospatial data.

The PAIGH is the continental entity that provides the scientific space to rethink America, based on its geography and history

Instituto Panamericano de Geografía e Historia (IPGH)
Ex-Arzobispado 29, Colonia Observatorio, 11860 México, D.F. secretariageneral@ipgh.org
http://www.ipgh.org