# Australian Map and Spatial Data
# Horizontal Accuracy Standard

2009

Subcommittee for
Permanent Committee on Topographic Information
Intergovernmental Committee on Surveying and Mapping

# Table of Contents

# 1 FOREWORD

## 1.1 Purpose of standard

(a) The purpose of this standard is to provide a common standard for calculating and reporting accuracy to facilitate the appropriate use of maps and spatial data for geographic and geoscientific applications.

(b) This standard will allow users of maps and spatial data which comply with this standard to judge whether those maps and data are sufficiently accurate for their applications.

## 1.2 Superseded documents

This standard supersedes horizontal accuracy statement(s) in the following documents:

- 1975, *Standards of Map Accuracy (2nd edition)*, National Mapping Council of Australia; and
- 1953, *Standards of Map Accuracy,* National Mapping Council of Australia.

## 1.3 Acknowledgement

This standard has in parts been based on the *US National Standard for Spatial Data Accuracy* (NSSDA), published by the US Federal Geographic Data Committee (FGDC) in 1998.

## 1.4 ISO Standard compatability

This standard is compatible with ISO TC211 standards relating to data quality, including:

- ISO/TC211 19105 Geographic information - Conformance and testing
- ISO/TC211 19113 Geographic information - Quality principles
- ISO/TC211 19114 Geographic information - Quality evaluation procedures
- ISO/TC211 19115 Geographic information - Metadata
- ISO/TC211 19138 Geographic information - Data quality measures.

# 2 SCOPE

(a) This standard

(i) specifies absolute horizontal positional accuracy calculation and reporting requirements for Australian maps and digital spatial datasets,

(ii) applies to all feature types of real world geographic and geoscientific phenomena contained in maps and digital spatial data, and

(iii) applies to maps and spatial data in raster, vector or hard copy format, and derived from sources such as aerial photographs, satellite imagery, or ground surveys.

(b) This standard does not apply to abstract features such as cadastral boundaries, survey networks, or geodetically surveyed points.

# 3 INTENDED USE OF STANDARD

(a) Accuracy statements for new or revised topographic maps or spatial datasets claiming conformance to this standard must satisfy the reporting requirements in this standard.

(b) Accuracy of existing or legacy spatial data and maps may be reported according to this standard or the accuracy standard by which they were originally evaluated.

**COMMENTARY**

**Non conformance**

If the requirements of this standard cannot be met, it is recommended that information be provided to enable users to evaluate how the data fit their applications. This information may include descriptions of:

- the source material from which the data were compiled,
- accuracy of ground surveys associated with compilation,
- digitizing procedures,
- equipment, and
- quality control procedures used in production

# 4 TERMS AND DEFINITIONS

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

| Term | Definition |
| --- | --- |
| AMSDAS | Australian Map and Spatial Data Accuracy Standard |
| accuracy | in the context of this standard, is absolute positional accuracy, as defined in the international standard ISO/TC211 19113 Geographic information - Quality principles, as opposed to relative accuracy |
| ICSM | Intergovernmental Committee on Surveying and Mapping |
| ISO TC/211 | International Organization for Standardization – Technical Committee 211 (Geographic Information/Geomatics) |
| NSSDA | National Standard for Spatial Data Accuracy |

## 5 ACCURACY REPORTING

(a) Accuracy must be reported at a threshold level of 95 % and the methodology by which the reported value was determined must also be reported.

> **COMMENTARY**
>
> **Explanation of 95 % threshold**
>
> - Accuracy reported at the 95 % threshold level means that 95 % of the positions in the dataset or on the map must have an error with respect to true ground position that is equal to or smaller than the reported accuracy value.
> - 95 % is often considered to be double the standard deviation. It is more closely determined as being 1.96 times the standard deviation.
> - The reported accuracy value must reflect all uncertainties, including those introduced by map production, data conversion, and data manipulation.

(b) The reported accuracy must apply to all features, including point, line, and polygon features.

(c) Where a dataset contains themes or geographic areas that have different accuracies and these accuracies can be identified separately in the dataset, these accuracies must be reported separately.

(d) Accuracy must be reported for features in their entirety and must not be limited to well defined point locations.

(e) Accuracy values must be reported in ground distances in metric units. Accuracy reporting in ground distances allows users to directly compare datasets of differing scales or resolutions.

(f) For digital spatial data, the accuracy must be reported in digital metadata.

(g) For hard copy maps, the accuracy must be reported on the map surround. The accuracy must be stated using words such as: “Horizontal accuracy xx metres at 95 % threshold level”. A description of how the accuracy was determined may be included.

# 6 ACCURACY TESTING

## 6.1 Introduction

This standard recommends that accuracy be tested in one or more of three ways:

(a) testing against an independent source of higher accuracy. Refer to paragraph 6.2,

(b) testing by deductive estimate. Refer to paragraph 6.3, and

(c) testing by inference. Refer to paragraph 6.4.

## 6.2 Testing against an independent source of higher accuracy

### 6.2.1 General

Using the method of testing against an independent source of higher accuracy, accuracy must be tested by comparing the planimetric coordinates of locations in the test dataset with coordinates of locations that can be assumed to be the same in the independent source of higher accuracy.

### 6.2.2 Independent source of higher accuracy

(a) The independent source of higher accuracy referred to in 6.2.1 above must be the highest accuracy practicable to evaluate the accuracy of the dataset and must be acquired from sources separate from the test dataset.

(b) Possible sources for higher accuracy information are:

  (i) geodetic ground surveys,

  (ii) Global Positioning System ground surveys,

  (iii) photogrammetric surveys, and

  (iv) and spatial databases of substantially higher accuracy.

(c) A minimum of 20 locations must be tested and distributed to reflect the geographic area of interest and the distribution of error in the dataset.

### 6.2.3 Accuracy of linear features

#### 6.2.3.1 General

There are various methods of measuring the accuracy of linear features against higher resolution representations of the same phenomena. This standard recommends accuracy testing for linear features using equally spaced perpendicular offsets along the tested feature to their intersection with the independent source of higher accuracy.

#### 6.2.3.2 Method does not match locations

In some circumstances, such as with a linear feature with irregular bends, this method does not logically match locations on tested features with their corresponding location on an independent source. In these cases the offsets to the independent source should be adjusted so that a logical matching is achieved.

#### 6.2.3.3 Calculating the accuracy

(a) When direct offset measurements are taken, 95 % threshold accuracy can be calculated through a root-mean-square-error (RMSE) determination (see Appendix 1):

$$95\% Accuracy = 1.7308 \times RMSE$$

(b) RMSE is the square root of the average of the set of squared differences between dataset coordinate values and coordinate values from an independent source of higher accuracy for identical points. Using an RMSE determination it is assumed that systematic errors have been eliminated and that error is normally distributed.

(c) When a large number of offsets are taken, the 95 % threshold accuracy value is equivalent to the 95$^{th}$ percentile offset value. Using this method, no assumption is made that systematic errors have been eliminated or that the error is normally distributed.

## COMMENTARY

### Example of positional accuracy

The following is an example of an estimate of positional accuracy undertaken by a government mapping authority using testing against a source of higher accuracy.

### Example

1. Testing against a source of higher accuracy has been undertaken for DATASET_A to determine its positional accuracy.
2. DATASET_A has a scale of 1:250 000 and the source of higher accuracy is at 1:25 000 scale.
3. Offsets from the linear features in DATASET_A to their intersection with matching features in the reference dataset have been measured. 125 000 regularly spaced perpendicular offsets have been measured.
4. Offsets have been inspected, and where necessary adjusted, to ensure that locations in the DATASET_A match those in the reference source in a geographically sensible way.
5. Offsets for a broad range of feature types in DATASET_A and across the full extent of DATASET_A have been measured.
6. In ground distances it has been found that 95 % of the offsets are less than 140 metres long and 5 % of the offsets are longer than 140 metres.
7. DATASET_A therefore has a horizontal accuracy of 140 metres at 95 % threshold level.

## 6.3 Testing by deductive estimate

(a) Modelling and propagation of errors through the data/map production processes is undertaken using testing by deductive estimate. For instance, for a vector dataset derived from a paper map, the accuracy may be deduced by accumulating the errors of the original data capture, including those due to photogrammetry and cartography, with those due to data conversion and manipulation.

(b) In the example in (a) above, where there are errors due to map source material, data conversion and data manipulation the 95 % threshold accuracy can be calculated as follows:

$$95\% Accuracy = 1.96 \times S = 1.96 \times \sqrt{(S_m)^2 + (S_l)^2 + (S_{man})^2}$$

where $S$ is the standard deviation and the subscripts $m$, $l$ and $man$ refer to source material, conversion processes, and manipulation processes respectively.

**COMMENTARY**

**Example of deductive estimate**

- The following is an example of a deductive estimate of positional accuracy undertaken by a government mapping authority.
- The example is for a 1:250 000 digital dataset derived from original map repromat, and subject to a small degree of manipulation, such as coordinate rounding and format conversion.
- Errors are given in millimetres on the map, prior to conversion to ground distances.

**EXAMPLE**

The planimetric accuracy attainable in the 1:250 000 DATASET_B will be composed of errors from three sources:

- the positional accuracy of the source material,
- errors due to the conversion processes, and
- errors due to the manipulation processes.

## The positional accuracy of the source material

(a) There is an expectation that the source data complies with the following statement: "*Not more than 10 % of locations will be in error by more than 0.5 mm measured on the source material".*

(b) Statistically, this relates to a standard deviation on the map $(S_m)$ of 0.31 mm.

## Errors due to the conversion processes

(a) The errors due to the conversion/capture process depend on the accuracy of the variety of factors dependant on the type of activity being undertaken. Some of these factors include requested digitising accuracy off imagery/aerial photography, method of raster to vector conversion, digitising table set-up or the scanner resolution, systematic errors in the equipment, errors due to software, and errors specific to the operator.

(b) When table digitising, the accepted standard is that the line accuracy should be within half a line width. As a majority of symbolised features in topographic mapping have a line width of 0.2 mm or greater, then half the line width is taken as 0.1 mm and this is interpreted as one standard deviation $S_{data}$ for the distribution of errors.

(c) The standard *deviation* of distribution errors in setting up the digitisation table is determined by the square root of the sum of all residual errors at each of the registration locations squared over the number of registration locations (minus one). For this conversion process the resultant standard deviation measurement $S_{test}$ is estimated to also be 0.1 mm (at map scale).

The error due to the conversion processes is:

$$
\begin{aligned}
S_l &= \sqrt{(S_{data})^2 + (S_{test})^2} \\
&= \sqrt{(0.1)^2 + (0.1)^2} \\
&= 0.14mm
\end{aligned}
$$

**Errors due to the manipulation processes**

(a) Errors due to the manipulation process include separating vector contours which have merged in the raster to vector process, correcting topological structure rules such as dangles and intersect errors and, as an extreme case, smoothing vegetation boundaries after conversion of raster vegetation analysis data.

(b) As a general rule, the processes used during data manipulation do not introduce an error greater than 10 % of the vector capture error $S_{data}$. $S_{man}$ is therefore estimated as 0.01 mm.

The *standard* deviation of the errors for DATASET_B, in millimetres on the map, is therefore

$$
\begin{aligned}
S &= \sqrt{(S_m)^2 + (S_l)^2 + (S_{man})^2} \\
&= \sqrt{(0.31)^2 + (0.14)^2 + (0.01)^2} \\
&= 0.34mm
\end{aligned}
$$

This represents an error of 85m on the ground for 1:250 000 data.

The 95 % threshold accuracy can then be calculated as:

$$
\begin{aligned}
95\%\,Accuracy &= 1.96 \times S \\
&= 1.96 \times 85 \\
&167
\end{aligned}
$$

DATASET_B therefore has a horizontal accuracy of 167 metres at 95 % threshold level.

## 6.4 Testing by inference

When testing by inference, accuracy is equated to that of another dataset or map of identical scale and method of production. For instance, where each map in a map series has the same scale and is subject to the same method of production it would be reasonable to assume they all have the same accuracy, and that by knowing the accuracy of one map, the accuracy of all subsequent maps in the series can be inferred.

**COMMENTARY**

**Example of an inferred estimate**

- The following is an example of an inferred estimate of positional accuracy undertaken by a government mapping authority.
- The example is for a digital dataset derived from the same map series as DATASET_B in the previous example.
- The source repromat was subject to the same procedures as for DATASET_B, and the parameters for conversion and manipulation of the data were the same as for DATASET_B.

**EXAMPLE**

1. DATASET_C has been derived from repromat from the same map series as DATASET_B, and has been subject to the same methods and parameters for data conversion and manipulation and is assumed to have the positional accuracy as DATASET_B.
2. DATASET_C therefore has a horizontal accuracy of 167 metres at 95 % threshold level.

# Appendix 1: Explanation of root-mean-square-error (RMSE)

## 1.1 Explanation

Let:

$$RMSE_E = \sqrt{\left[\frac{\sum (E_{data,i} - E_{check,i})^2}{n}\right]}$$

$$RMSE_N = \sqrt{\left[\frac{\sum (N_{data,i} - N_{check,i})^2}{n}\right]}$$

where:

$E_{data,i}$ , $N_{data,i}$ are the coordinates of the $i^{th}$ check location in the dataset

$E_{check,i}$ , $N_{check,i}$ are the coordinates of the $i^{th}$ check location in the independent source of higher accuracy

*n* is the number of check locations tested

*i* is an integer ranging from 1 to *n*

Horizontal error at location *i* is defined as $\sqrt{\left[\left(E_{data,i} - E_{check,i}\right)^2 + \left(N_{data,i} - N_{check,i}\right)^2\right]}$ .

Horizontal RMSE is:

$$RMSE_r = \sqrt{\left[\frac{\sum ((E_{data,i} - E_{check,i})^2 + (N_{data,i} - N_{check,i})^2)}{n}\right]}$$

$$= \sqrt{[RMSE_E{}^2 + RMSE_N{}^2]}$$

## 1.2 Special Case: When $RMSE_E = RMSE_N$

If $RMSE_E = RMSE_N$ then

$$RMSE_r = \sqrt{(2 \times RMSE_E^{\ 2})} = \sqrt{(2 \times RMSE_N^{\ 2})}$$

$$= 1.4142 \times RMSE_E = 1.4142 \times RMSE_N$$

where $RMSE_r$ is RMSE in any radial direction.

It is assumed that systematic errors have been eliminated as best as possible. If error is normally distributed and independent in the E and N components, the factor 2.4477 is used to compute horizontal accuracy at the 95 % threshold level. When these conditions apply, $95\% Accuracy_r$, the AMSDAS accuracy value, shall be computed by the formula:

$$95\% Accuracy_r = 2.4477 \times RMSE_E = 2.4477 \times RMSE_N$$

$$= 2.4477 \times RMSE_r / 1.4142$$

$$= 1.7308 \times RMSE_r$$