Metadata and quality

Moe Myint[*] Claude Collet[†] Dominique Schneuwyl[‡]

June 21, 2010

[*]UNIFR moemynth@bluewin.ch, Overall
[†]UNIFR claude.collet@unifr.ch, Concept
[‡]UNIFR dominique.schneuwly@unifr.ch, Specials

Contents

Metadata and quality

In this lesson you will

- Be introduced to the definition, components of, acquiring, importing, editing and exporting of Metadata.
- Learn how to quantify the positional accuracy of data and evaluate its quality.
- Learn how to manage the errors in order to produce quality results.

Learning Objectives

- You will learn about documenting data about data or metadata. You have to create metadata for the data you've created. You have to acquire metadata if you acquired data from someone else. You will learn about elements of metadata, metadata standards and tools used to create metadata.
- Positional accuracy is discussed from the perspective of metadata documenting and quality assessment if you acquired the data from a data provider or you created the data. You will learn that accurate data (accuracy information from the metadata) does not mean the data is precisely measured. Moreover, precise measurements (precision information from the metadata) do not necessarily indicate accurate data. In addition, you will learn that these are scale dependent and depending on the accuracy requirements of your particular project. No single universal accuracy standard is correct for every application GIS Project.
- Attribute accuracy is discussed from the perspective of metadata documenting and quality assessment if you acquired the data from a data provider or you created the data. It is recommended to test the attribute accuracy of the dataset based on enough sample points to validate the accuracy before using the dataset, if you acquired it or before distributing the dataset if you created it. You will learn a quantitative method to calculate the sample size for testing and a method for testing the attribute accuracy. The results would be documented in the metadata.
- You will learn how to live with errors because we cannot eliminate all the errors in GIS. However, the error can be managed; a checklist for data quality is provided. Moreover, the concepts of acceptable

and unacceptable known positions are provided to validate precision and accuracy of your data. In addition, error propagation and map overlay errors are discussed to control data quality throughout the analysis process. You will also learn sensitivity analyses to control the quality of data and results.

Metadata Acquisition

Objectives

This unit will introduce the definition, components of, acquiring or importing, editing and exporting of metadata. At the end of the unit, you will have understood why you need to acquire metadata when you receive the data or, create metadata when you create data and export the metadata when you distribute the data.

Metadata Acquisition

Metadata Acquisition

Download the PDF version of the page here.: www.gitta.info/MetaDataQual/en/multimedia/MetaDataAcqua.pdf

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 1: Click on the buttons to view the complete animation.

Creating Metadata

1. Browse and create the ISO 19115 standard document using the free metadata tool, M3CAT.

Download the M3Cat tool from the intelec website

M3Cat tool here.: www.intelec.ca/technologie_a.html

Read the installation guide first. Then install necessary software components such as IIS (Internet Information Server) or PWS (Personal or Peer Web Server) first. Then install the M3Cat. It is important to note, that to create the virtual directory that points to the M3Cat <Web> installation directory in IIS or PWS. Moreover, a system ODBC has to be created because M3Cat use the ODBC link to communicate to the database. The default user name is **user** and password is **eng** , if you install the English version of M3Cat.

Follow the Guide Tour of M3Cat. Then create an ISO 19115 document.

2. Optional exercise

Creating Metadata usingArcCatalog.

If you are using ArcGIS, download the optional exercise to generate metadata using ArcCatalog.

Optional exercise here: www.gitta.info/MetaDataQual/en/multimedia/MetaDataAcquaArcTool.pdf

Exercise Question

1. What is metadata?
2. What information about data should be documented in metadata?
3. Create ISO Standard Metadata for a selected SWISS Map using the M3CAT tool.

Positional Accuracy

Learning Objectives

This unit will explain the basic concept of Positional Accuracy of feature geometry. Moreover, the unit will explain the quantitative method to quantify positional accuracy. Other topics such as resolution and scale of map, which affect the positional accuracy, will also be discussed. You will learn how to quantify positional accuracy and evaluate positional accuracy quality of data that you want to use in your future projects.

Positional Accuracy and Cartographic Representation of Spatial Features

Positional Accuracy and Cartographic Quality

Download the PDF version of the page here.: www.gitta.info/MetaDataQual/en/multimedia/PositAccuracy.pdf

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 2: positaccuracy.swf

Graphical Presentation of Quantitative Method to Measure Positional Accuracy

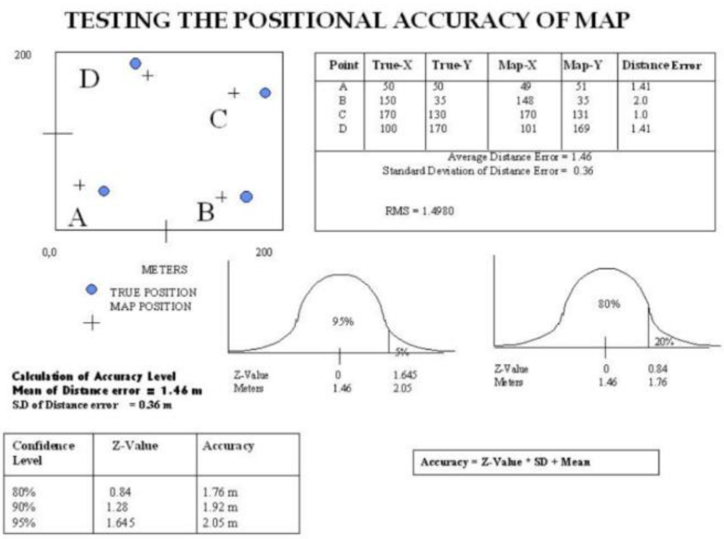
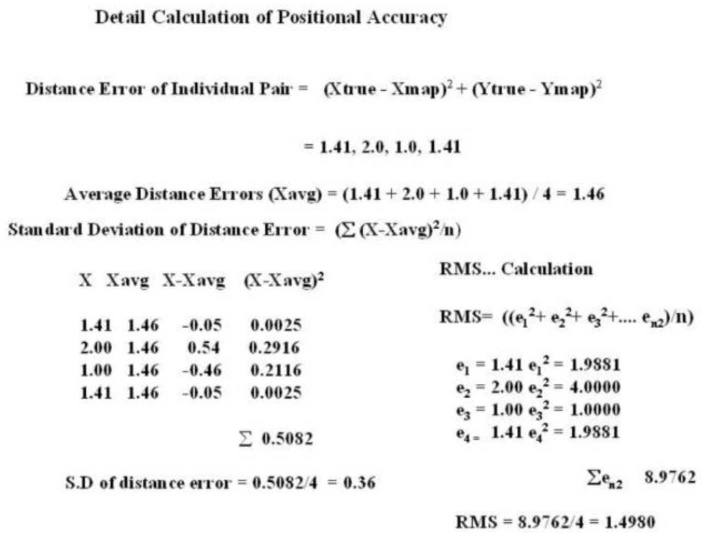


Figure 3: Graphical Presentation of Position Accuracy Calculation (Click on it)

Animation to Calculate Acceptable Error and Allowable RMS

The following animation can be used to calculate the allowable RMS.

Graphical Presentation of Resolution and Scale of the Map



Detail Calculation of Positional Accuracy

Distance Error of Individual Pair = $(Xtrue - Xmap)^2 + (Ytrue - Ymap)^2$

= 1.41, 2.0, 1.0, 1.41

Average Distance Errors (Xavg) = (1.41 + 2.0 + 1.0 + 1.41) / 4 = 1.46

Standard Deviation of Distance Error = $(\Sigma (X-Xavg)^2/n)$

X	Xavg	X-Xavg	$(X-Xavg)^2$
1.41	1.46	-0.05	0.0025
2.00	1.46	0.54	0.2916
1.00	1.46	-0.46	0.2116
1.41	1.46	-0.05	0.0025
		Σ	0.5082

S.D of distance error = 0.5082/4 = 0.36

RMS... Calculation

RMS= $((e_1^2 + e_2^2 + e_3^2 + ... e_{n2})/n)$

e_1 = 1.41 e_1^2 = 1.9881
e_2 = 2.00 e_2^2 = 4.0000
e_3 = 1.00 e_3^2 = 1.0000
e_4 = 1.41 e_4^2 = 1.9881

Σe_{n2} 8.9762

RMS = 8.9762/4 = 1.4980

Figure 4: Detailed Method of Position Error Calculation (Click on it)

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 5: Calculation of allowable RMS (Klick on it)

Calculate the Positional Accuracy

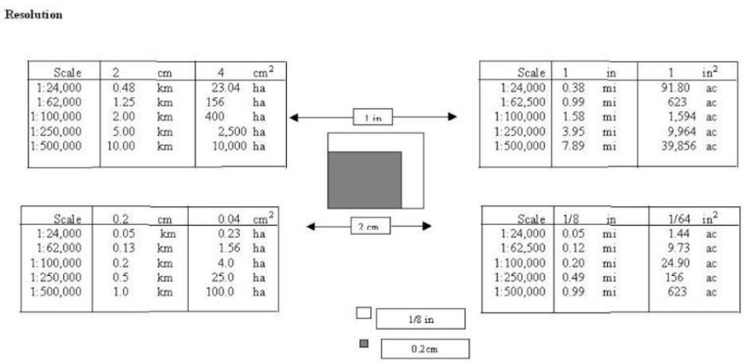
Calculate Average Distance Error, Standard Deviation of Distance Error, RMS, Positional accuracy at 80%, 90% and 95% confidence level based on the following dataset.

Point	TrueX	TrueY	MapX	MapY
A	53	58	54	60
B	48	86	51	84
C	30	36	28	36
D	70	172	68	174

Table 1: Legend missing

E	20	25	20	25
F	174	203	171	204

Table 1: Legend missing



Scale	2 cm		4 cm^2	
1:24,000	0.48	km	23.04	ha
1:62,000	1.25	km	156	ha
1:100,000	2.00	km	400	ha
1:250,000	5.00	km	2,500	ha
1:500,000	10.00	km	10,000	ha

Scale	1 in		1 in^2	
1:24,000	0.38	mi	91.80	ac
1:62,500	0.99	mi	623	ac
1:100,000	1.58	mi	1,594	ac
1:250,000	3.95	mi	9,964	ac
1:500,000	7.89	mi	39,856	ac

Scale	0.2 cm		0.04 cm^2	
1:24,000	0.05	km	0.23	ha
1:62,000	0.13	km	1.56	ha
1:100,000	0.2	km	4.0	ha
1:250,000	0.5	km	25.0	ha
1:500,000	1.0	km	100.0	ha

Scale	1/8 in		1/64 in^2	
1:24,000	0.05	mi	1.44	ac
1:62,500	0.12	mi	9.73	ac
1:100,000	0.20	mi	24.90	ac
1:250,000	0.49	mi	156	ac
1:500,000	0.99	mi	623	ac

Figure 6: Relationship of scale and minimum mapping unit (Click on it)

Questions

1. Propose a minimum mapping unit for 1:25000 scale maps.
2. Describe the map resolution with examples.

Attribute Accuracy

Learning Objectives

This basic concept of attribute accuracy has been introduced in the unit of quality and coverage. This unit will teach a basic technique to quantify the attribute accuracy. At the end of the unit, you will be able to assess the attribute accuracy based on simple a error matrix.

Attribute Accuracy

Attribute Accuracy

Download the PDF version of the page here.: `www.gitta.info/MetaDataQual/en/multimedia/attributeaccuc.pdf`

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 7: attributeaccuracy.swf

Attributes types and values

1. Nominal

It is any name or number. Numbers merely establish identity, allowing logical operations, classification and identification.

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 8: bank.swf

2. Ordinal

Ordinal attribute numbers establish order only. Functions on logical and ranking operations and comparison of magnitude operations are allowed.

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 9: ordinal.swf

3. Interval

On the interval scales, the different intervals between numbers are meaningful.

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 10: interval.swf

4. Ratio

On a ratio scale, measurement has an absolute zero, the difference between numbers is significant, and division makes sense. A population density map is an example of ratio scale measurement.

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 11: ratio.swf

5. Boolean

Boolean attribute has 0 and 1 to indicate presence and absence or Yes and No. It is useful for logical and indicator operations such as truth versus falsehood.

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 12: binary.swf

6. Multiscale Representation

A phenomenon, characteristic, or attribute of a real world objects represented in different ways at the different scale.

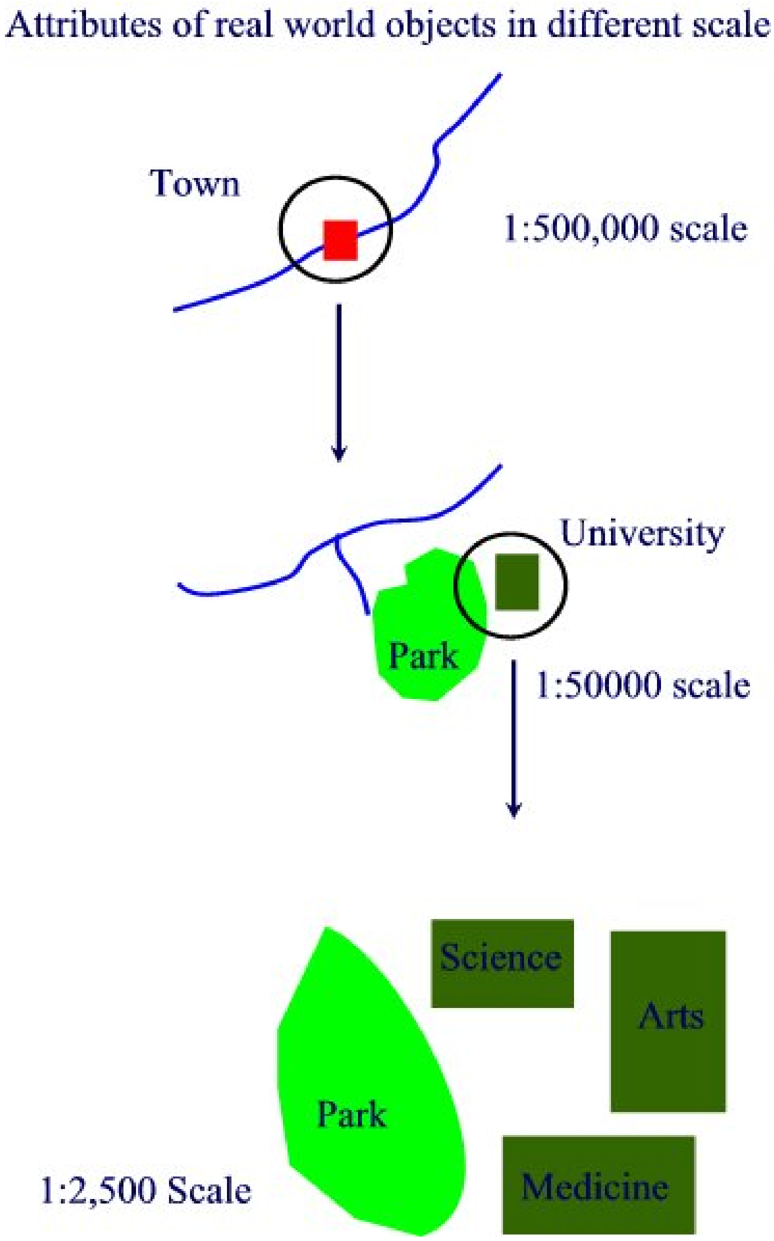


Figure 13: Multiscale representation (Click on it)

Exercises

1. Review the facts that influence the attribute accuracy of real world objects and phenomenon.
2. The 1990 land use map is 1:250000 scale. The 2000 land use map is 1:50000 scale. Is it logical to overlay the two maps to analyze the land use changes? Explain your opinion.
3. Calculate the sample size to evaluate the attribute accuracy for a map that has 15 categories.
4. Calculate the overall, producer and user attribute accuracy of the following dataset.

Land cover	**A**	**B**	**C**	**D**	**E**	**F**
A. Mixed Forest	8	0	1	0	2	0
B. Paddy	0	10	0	2	0	1
C. Plantation	3	0	21	0	4	0
D. Corn	0	0	0	9	0	3
E. Pine	4	0	5	0	23	0
F. Sugarcane	0	2	0	3	0	14

Table 2: Legend missing

Error Management in Data Capture

Learning Objectives

There is no perfect data for GIS for a particular application. The bad news is that, you cannot eliminate the errors of source data. The good news is that, you can manage the errors. You will learn how to manage the errors in order to produce quality results.

Managing Errors in Data Capture

Download the PDF version of the page here.: `www.gitta.info/MetaDataQual/en/multimedia/errormanage.pdf`

Please note:
Only pictures can be viewed in the PDF version! For Flash etc. see online version. Only screenshots of animations will be displayed.

Figure 14: errormanage.swf

Data Quality Report

Optional Link:

Data Quality Report of DLG (Digital Line Graph): `eros.usgs.gov/guides/dlg.html`

Exercises

1. What should be considered when checking the quality of data?
2. Accuracy of the measurement is the nearness of that measurement to truth. Yes/No
3. Precision is defined as the spread (or dispersion) of the measured values of a quantity. Yes/No
4. Precision provides an indication of the quality of data but does not provide an indication of the accuracy of data. Yes/No
5. High precision does not indicate high accuracy. High accuracy data does not indicate high precision. Yes/No
6. Absolute accuracy requires measured data to be referenced in terms of known or existing data. Yes/No
7. Relative accuracy does not need to be referenced to other data in order for measurements to be correct. This is usually the case where data has been collected as a completely self-contained survey. Yes/No
8. The digitizing staff must be well informed of the allowable RMS of the project in order to maintain the positional accuracy requirement of the project. Yes/No