


Auditing Spatial Data Suitability for Specific Applications : Professional and Technological Issues

**EuroSDR, EuroGeographics & the AGILE Working Group
Quality Assurance in Geographic Data Production.**



Marc Gervais^{1,2}, Yvan Bédard^{1,2}, Rodolphe Devillers^{1,2,3}
Suzie Larrivée^{1,2}, Nicholas Chrisman¹, Johann
Lévesque^{1,2}

¹ Laval University, Quebec City, Canada

² Centre for Research in Geomatics (CRG), Quebec City, Canada

³ Memorial University of Newfoundland, St-John's, Canada

Marne-La-Vallée, France

February 14th, 2006

```
021e05.met - WordPad
File Edit View Insert Format Help

!
BEGIN          POLYGON_SECTION
NB_POLYGONS    3
BEGIN          POLYGON
ID_POLYGON     0001
COORDINATES    303820 5013763 284203 5014402 264585 5015101 265621 5042875
COORDINATES    285152 5042175 304683 5041536 303820 5013763
ENTITIES       P 1-1849,1856-2047
ENTITIES       L 1-1809,1825-2047
ENTITIES       S 1-1829,1834-2047
SOURCE_TYPE    CT+PHA (Aerial photography with field completion)
SOURCE_NAME    PHOTOS.RNCAN
VALID_DATE     1985/-1
PLAN_ACCU_QUAL E (Estimated)
PLAN_ACCURACY  10
ALTI_ACCU_QUAL E (Estimated)
ALTI_ACCURACY  5
ACTION         ACQ.COMP.STER (Complete acquisition of entities by stereo)
IMPACT_ACT_C   S (Systematic)
IMPACT_ACT_P   N (No)
POL_ED_VER     3.00
COMMENT
END            POLYGON
!
BEGIN          POLYGON
ID_POLYGON     0002
```



For Help, press F1

Introduction

- **Yesterday :**

- Measurement scientists (photogrammetrists, geodesists, surveyors, etc.) according to user's needs.
- Highly technical skills and knowledge necessary to produce “quality maps”.
- Dataset designed for single-purpose (e.g. cadastre).

- **Nowadays :**

- Almost everyone can acquire precise data.
- Possibility to modify data and integrate them with other data.
- Data perceived and marketed as multi-purpose.



A simple question...

If a user comes to you with a given dataset and a well-defined need, looking for the advice of an expert with professional liability, and asks you :

« Is the quality of these data adequate for my needs for the whole area ? »

What would you do ? How would you proceed ?

No simple answer...



Table of content

- **Context**
 - Quality issues
 - Auditing as a professional act
- **Research towards Spatial Data Quality Audit (SDQA) :**
 - Elements of a Quality Report (QR)
 - Practice guide for QR
 - DSS for auditing and QR
- **Conclusion**



Quality Issues

- Quality evaluation is very complex.
- With new markets, emergence of new legal issues.
- No algorithm can do all the job.
- Too costly to evaluate all data quality aspects.
- Focus must be placed on user's specific needs.
- Need to develop a new approach for quality evaluation.



Spatial Data Quality Audit (SDQA)

Auditing as a professional act (Resulting in a Quality Report)

Data Quality Checkup

Data Quality Check-Up

For companies seeking to improve data quality,

ensure accurate reporting, reduce excess costs

due to dirty data, and increase end user acceptance, Blue Hammock offers the Data

Quality Checkup. Working with your staff, our expert consultants will assist in the evaluation of your data, from data collection to reporting, processes to architecture. Based on our findings, we'll identify the detailed next steps involved in improving your data with quick short-term improvements, as well as more strategic, long-term enhancements.

The goal of the Data Quality Checkup is geared towards identifying the key areas of improvement available to any current initiative hurting from bad data. The Data Quality Checkup proposes to identify the amount of data corruption you are facing, understand the root cause of incorrect information, trace back the sources of bad

Approach

Analyze Current



Requirements
Definition

Gather New Requirements
End-User Usage
Goals and

The Data Quality Checkup initiative into a well-form

- Location Certificate QLS. Capacity for a mortgage loan.
- Examination Report. House quality.
- Psychometric Evaluation. Capacity to fulfill a position.
- Inquiry. Cares' quality.
- Titles Report. Real estate rights quality.
- ...
- *Data Quality Check Up.* Data quality.

Because that's the way the others do...

Auditing as a professional act (Resulting in a Quality Report)

- **Definition :**
 - “Systematic method from which one defines, collects and analyzes information on a geographic dataset mainly attached to the customer’s/user’s needs in order to make an objective judgment and/or a decision concerning the use of this data set”.
- **Auditing process can be view as :**
 - A professional act.
 - A certain form of diagnosis.
 - State of dataset (*internal quality*)
 - Tendency to satisfy user’s need for a given usage (*external quality*).



Auditing as a professional act

(Examples of questions to be addressed)

- Is the ground precision sufficient for his needs ?
- Do objects' classes fit his needs ?
- Do they share the same semantic definitions?
- If they don't, what is the impact on the queries ?
- Can database structure support his queries ?
- Is the quality heterogeneous ?
- If it is, do this situation will have an impact on the outcomes ? In which way ?



Auditing as a professional act

- The professional must :
 - Certify data quality (risk absorption).
 - Analyze data acquisition / treatment / management / diffusion methods.
 - Describe data quality.
 - Recommend consistent uses (or use limitations) with regards to data quality (*internal quality* to *external quality*)
 - Etc.



Auditing as a professional act

- Outcomes for the user :
 - Can act based on advices received.
 - Uncertainty is decreased.
 - A large part of the risk is transferred to the professional (liability).




Table of content

- Context
 - Quality issues
 - Auditing as a professional act
- **Research towards Spatial Data Quality Audit (SDQA) :**
 - Elements of a Quality Report (QR)
 - Practice guide for QR
 - DSS for auditing and QR
- Conclusion



Elements of a Quality Report

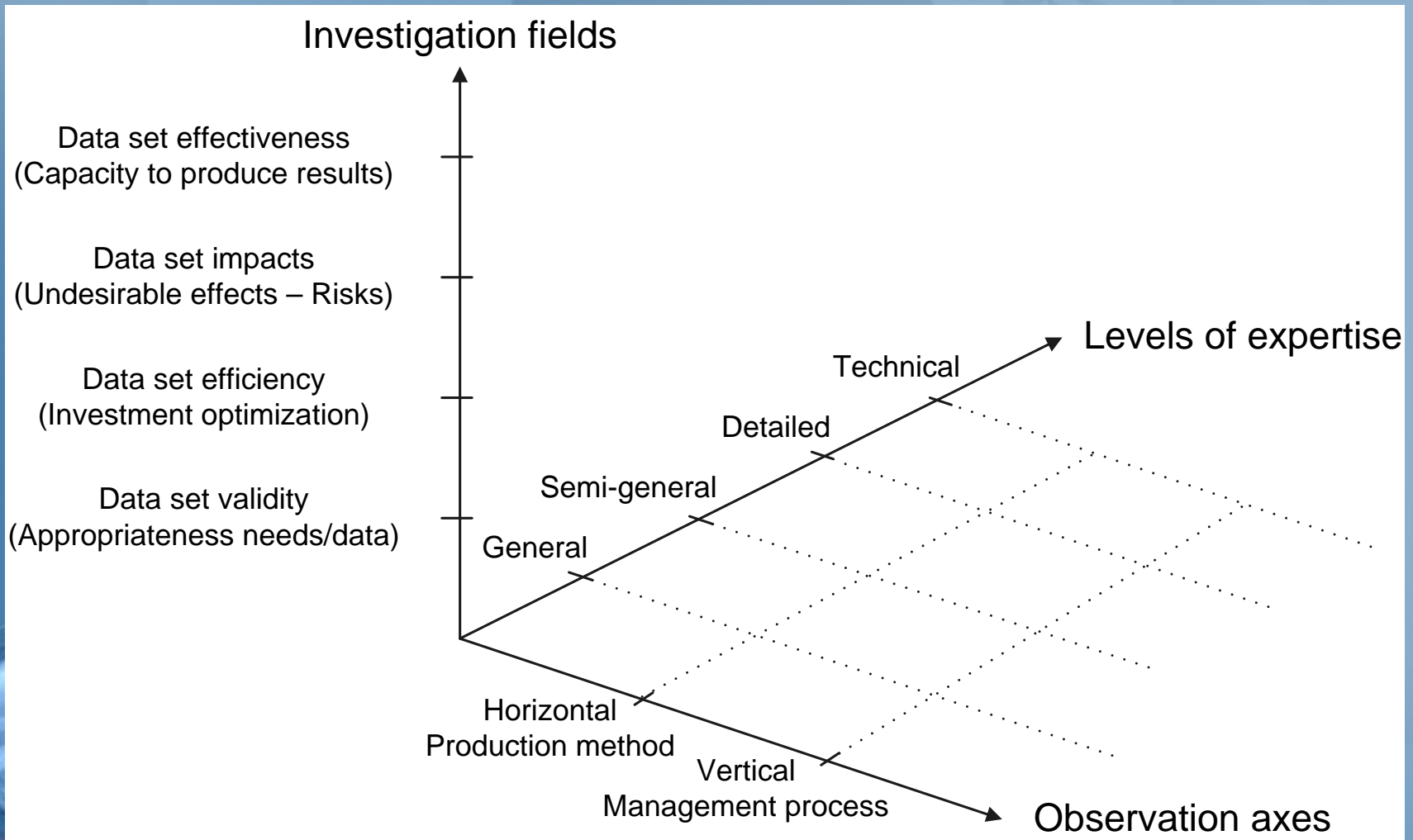
- Mandate / needs.
- Product description.
- General advices
- Uses :
 - Recommended.
 - Not Recommended.
- Statement report :
 - Favorable.
 - Unfavorable.
- Warnings (risks).
- Recommendations.
- Guarantees (liability).
- Licensing.
- Etc.

Symbols	
	Warning
	Problem to correct
	Information
	Urgent repairing
	Expertise recommended
	Potential danger
	Monitoring recommended
	Limited inspection
	Suggestion

Contents coming from legal rules (Gervais, 2004)

Practice guide for Quality Report

- Contextual and multidimensional approach :



Practice guide for Quality Report

- **Four (4) possible levels of expertise :**
 - Level 1 (General) : Analysis restricted or limited to the documentation transmitted with the dataset.
(metadata, etc.)
 - Level 2 (Semi-general) : Level 1 + dataset visualization
(spatial/descriptive).
 - Level 3 (Detailed) : Level 2 + analysis of GIS's functions
(data integrity and queries' feasibility).
 - Level 4 (Technical) : Level 3 + global quality evaluation based on other data sources
(field surveys, GPS, aerial photographs, satellite imagery, etc.).



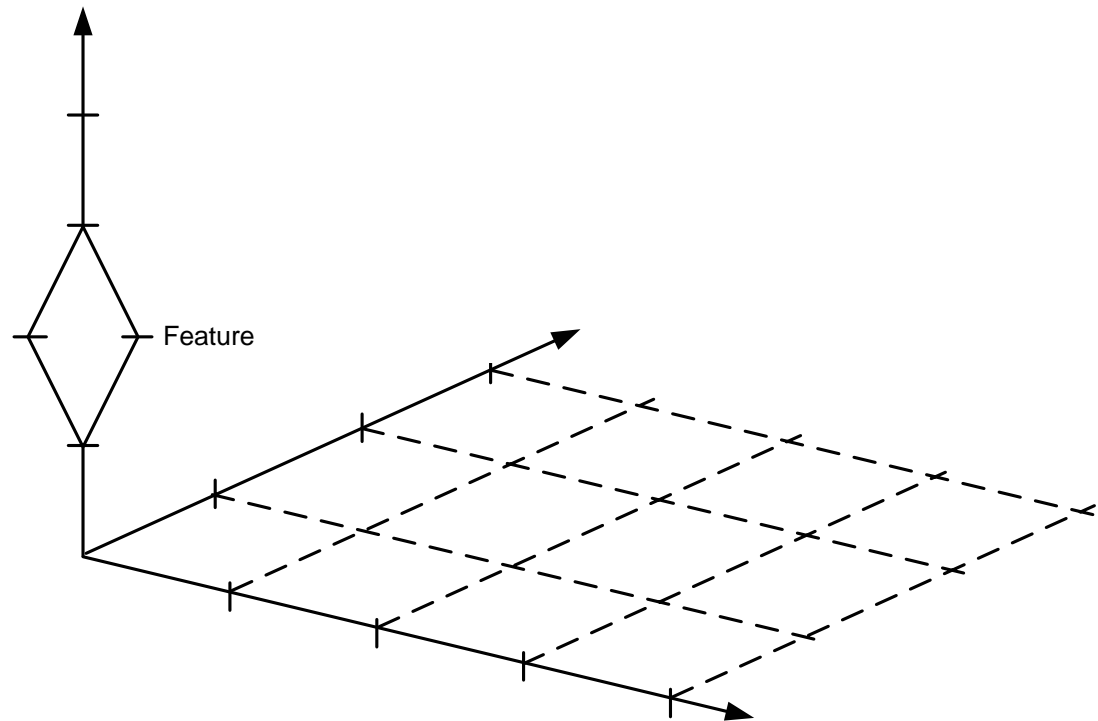
DSS for auditing and QR

- **Objective**: Designing a tool that would support professionals in the assessment of quality for a given application
- **Approach**:
 - Combining GIS and advanced database techniques (OLAP) to support data granularity = **SOLAP**
 - Simplify and organise the large and complex volume of information about data quality using « **indicators** » following ISO (e.g. 19113) or not



DSS for auditing and QR

- **Structure**: Information is organised hierarchically and some aggregation can be done to provide aggregated view of the quality





Map Quality

Select indicator to map

Map

Risk Tolerance

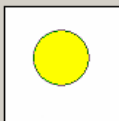
Average



QUALITY DASHBOARD

Global quality indicator

870



Completeness



Logical Consist



Positional Accu



Temporal Accur



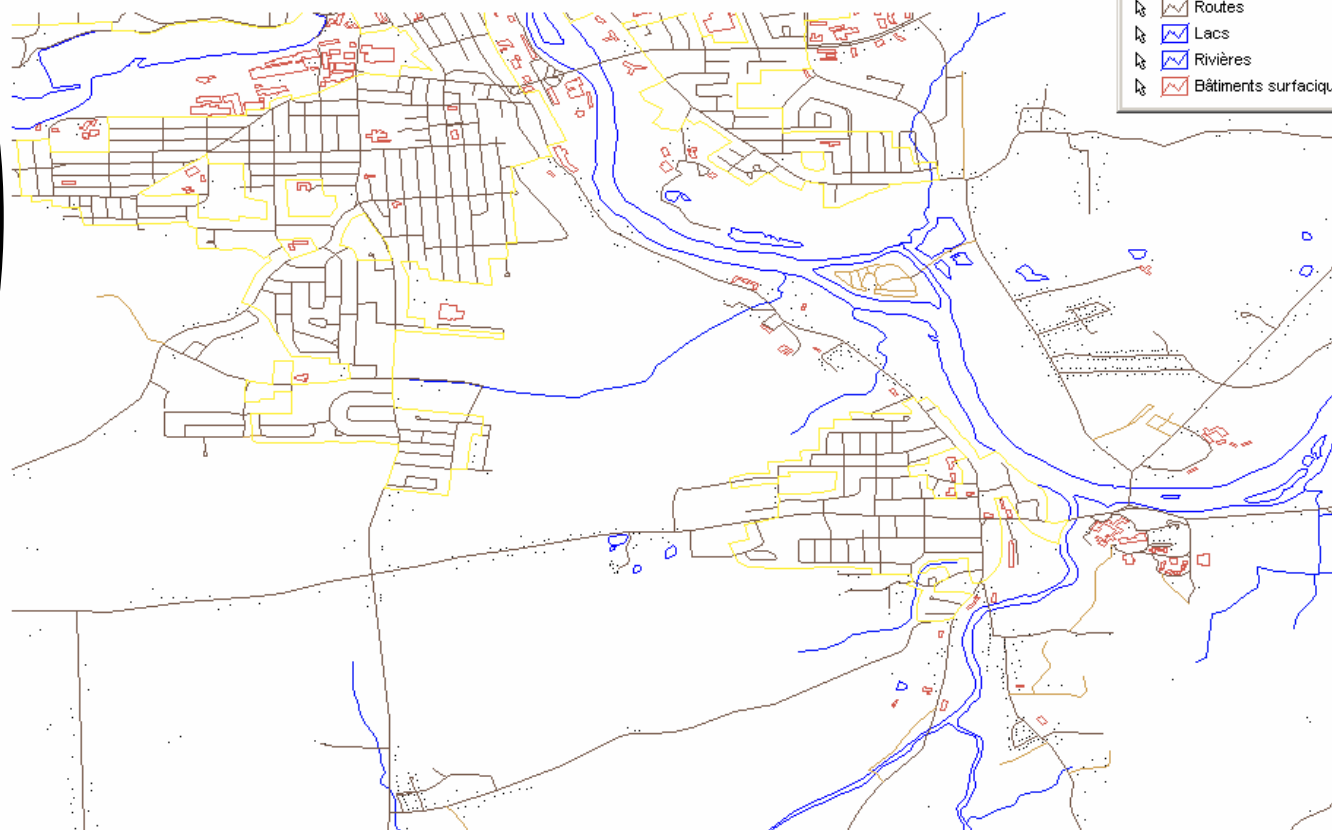
Thematic Accur



Up to date

Close

↑
Indicators





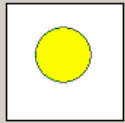
Map Quality
 Up to date

Risk Tolerance
 Average

QUALITY DASHBOARD

Global quality indicator

1870

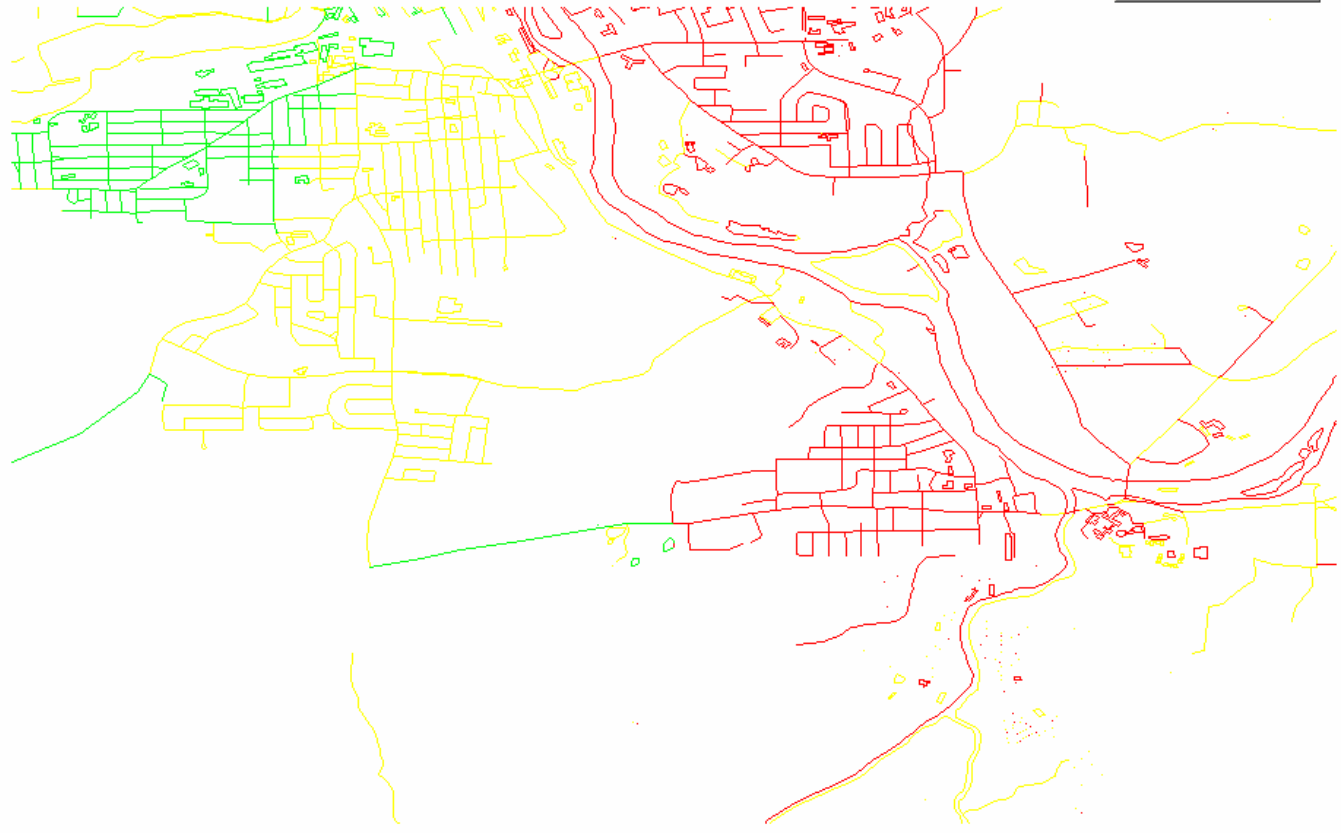


Completeness	Logical Consist	Positional Accu
Temporal Accur	Thematic Accur	Up to date

Explore

Legend

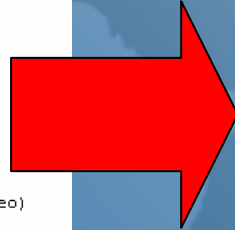
- Map Quality
- Poor
- Average



For a more extensive demo come to see me (on my laptop)

DSS for auditing and QR

```
021e05.met - WordPad
File Edit View Insert Format Help
!
BEGIN          POLYGON_SECTION
NB_POLYGONS    3
BEGIN          POLYGON
ID_POLYGON     0001
COORDINATES    303820 5013763 284203 5014402 264585 5015101 265621 5042875
COORDINATES    285152 5042175 304683 5041536 303820 5013763
ENTITIES       P 1-1849,1856-2047
ENTITIES       L 1-1809,1825-2047
ENTITIES       S 1-1829,1834-2047
SOURCE_TYPE    CT+PHA (Aerial photography with field completion)
SOURCE_NAME    PHOTOS.RNCAN
VALID_DATE     1985/-1
PLAN_ACCU_QUAL E (Estimated)
PLAN_ACCURACY  10
ALTI_ACCU_QUAL E (Estimated)
ALTI_ACCURACY  5
ACTION         ACQ.COMP.STER (Complete acquisition of entities by stereo)
IMPACT_ACT_C   S (Systematic)
IMPACT_ACT_P   N (No)
POL_ED_VER     3.00
COMMENT
END            POLYGON
!
BEGIN          POLYGON
ID_POLYGON     0002
```



Map Quality: Up to date | Map | Risk Tolerance: Average

QUALITY DASHBOARD

Global quality indicator: 1870

Completeness	Logical Consist	Positional Accu
Temporal Accu	Thematic Accu	Up to date

Explore

Metadata

Quality
Decision-Support System



DSS for auditing QR

(Example)

- **User general need :**
 - Evaluate if a given dataset can be used to automatic routing generation between an origin and a destination (optimal path)
 - Based on Geomatics Canada topographic maps (NTDB) + ArcGIS ArcView 8.3



DSS for auditing QR

(Example)



- **User specific needs :**
 1. To be 10 meters accurate.
 2. To be complete and up-to-date.
 3. To be error-free.
 4. To generate a network that could identify shortest paths.
 5. To identify origin/destination points.
 6. To have object classes that could be used as control points.
 7. To know cities limits

Etc.



DSS for auditing QR

(Example – tests performed)

- **1. Select bridges and road segments, to make sure that bridges were digitized correctly.**
 - **Result :**
 - No overlapping between bridges and road segments.
 - Correct inclusion of topological relationships between these two object classes.
-   Increase the likelihood of supporting network analysis.

DSS for auditing QR

(Example – tests performed)

- **2. Verify the accuracy of distance measurements.**
 - Test measurement:
 - **From** the intersection between *rue de la Terrasse / avenue du Séminaire*
 - **To** the intersection *rue de la Terrasse / avenue de la Médecine*.
 - Result : 502m from NTDB dataset compared to 500m from another dataset at the 1:1000 scale .



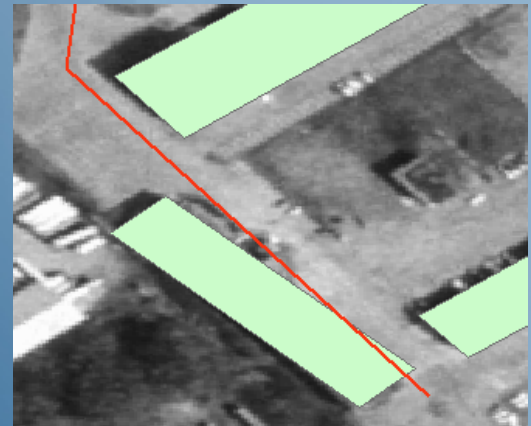
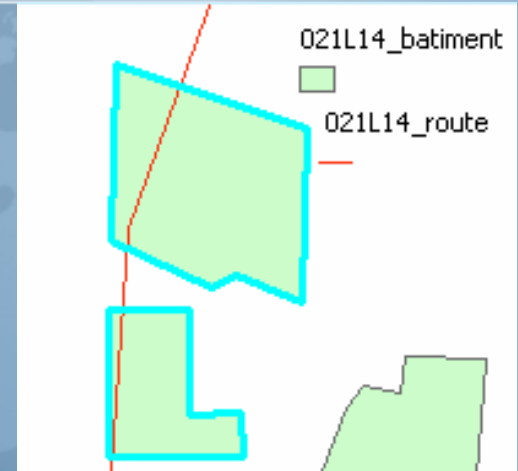
Increase the likelihood of reaching the required precision.



DSS for auditing QR

(Example – tests performed)

- **3. Verify if buildings cross road segments.**
- Result : 40/1692 buildings (2,3%) cross a road segment
(NB: some of these buildings could really overlap the road).



Some query results could be wrong.



DSS for auditing QR

(Example – tests performed)

- **4. Verify the NTDB level of resolution. It actually divide itself into two resolution classes.**
- **Results :**
 - Urban and rural areas = topographic maps to 1/50 000.
 - Isolated area = topographic maps to 1/250 000.
 - Then, resolution is variable.



Some control points located in isolated areas can be missing from the map.



DSS for auditing QR

(Example)

- **Favorable statements :**
 - NTDB data is in general of good quality for the usage and the whole region identified in your needs.
 - Etc.
- **Unfavorable statements :**
 - Dataset is not adequately structured to perform network analysis.
 - Object class “*Building*” has an incomplete spatial coverage.
 - Etc.



DSS for auditing QR

(Example)

- **Recommendations :**
 - To fix the dataset to be useful for network analysis.
 - To complete the spatial coverage of the object class “*Building*”.
 - Etc.



Table of content

- Context
 - Quality issues
 - Auditing as a professional act
- Research towards Spatial Data Quality Audit (SDQA) :
 - Elements of a Quality Report (QR)
 - Practice guide for QR
 - DSS for auditing and QR
- **Conclusion**



Conclusion

- **Research ongoing on Spatial Data Quality Audit (SDQA) :**
 - Elements of a Quality Report (QR)
 - Practice guide for QR
 - Decision Support System (DSS) for auditing and QR
- Metadata are not often used mostly because their current format is inadequate
- Metadata are expensive to produce but critically important in such approach
 - Follow standards (interoperability)
 - Document at a sufficient level of granularity



More details

- **Publications :**

- PhD thesis (Gervais + Devillers, 2005) + MSc thesis (Lévesque, Summer, 2006)
- Scientific papers:
 - *PE&RS* vol.71, n.2 (2005)
 - *Revue Internationale de Géomatique* (Devillers et al., 2004 + Gervais and al. 2006)
- Chapters in the book “Spatial Data Quality” to be published in 1-2 months by ISTE Publishing



- Thank you for your attention...

Questions ?

