Chapter II

Geographic Information Systems: Data Modeling

80 %

"80 % of all data throughout the world have some geographic background"

Application Domains

- Urban planning
- Land use planning
- Rural and forestry planning
- Sea and Marine Applications
- Transports
- Natural resources (mines) planning and management
- Earth sciences
- Archaeology
- Big real estate planning and management
- etc

GIS: Data Modeling

2.1 – Geographic Data Modeling

- Discrete Objects
  - Generally modeled by their boundaries
  - What models to use for points, lines, areas and volumes?
- Attribute modeling
- Continuous phenomena (ex. Temperature)
  - Modeled as continuous fields

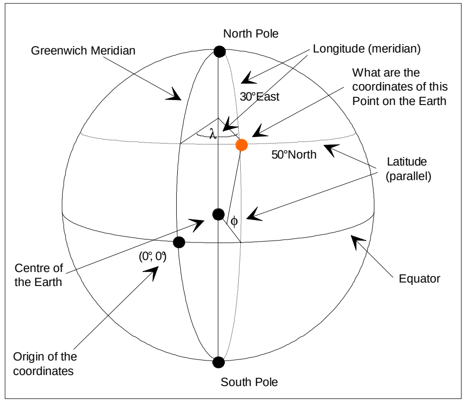
Earth Positioning

- Geodesy
- Coordinates
- Projections of the sphere

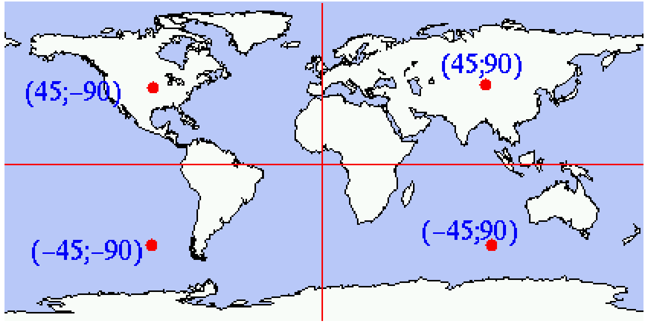
Geodesy

- The Earth is not exactly a sphere
- ellipsoid
- geoid
- altitude

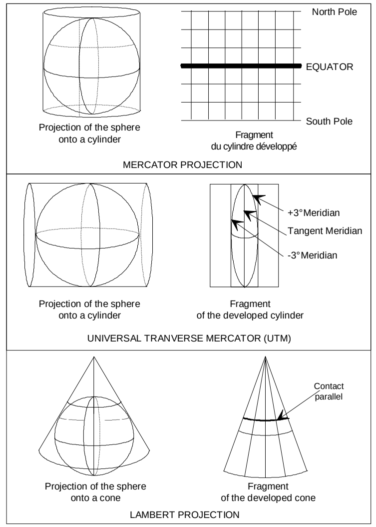
Coordinates



Coordinate center

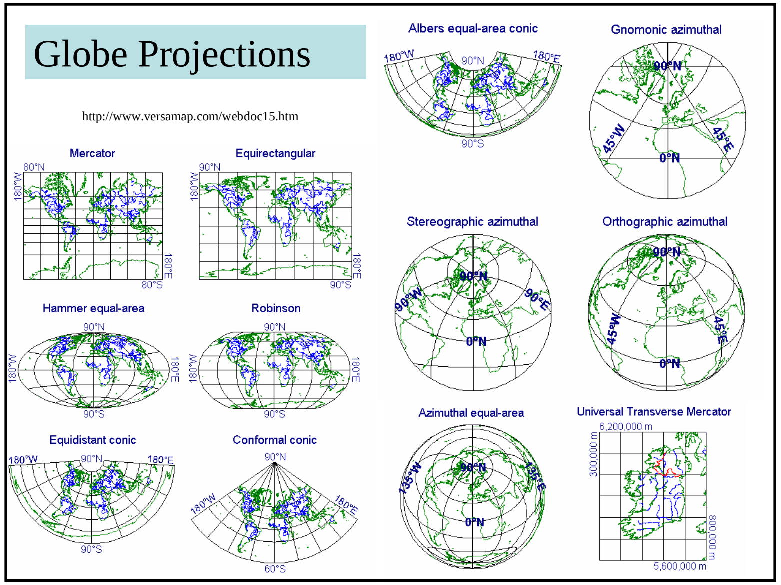


Projections of a sphere

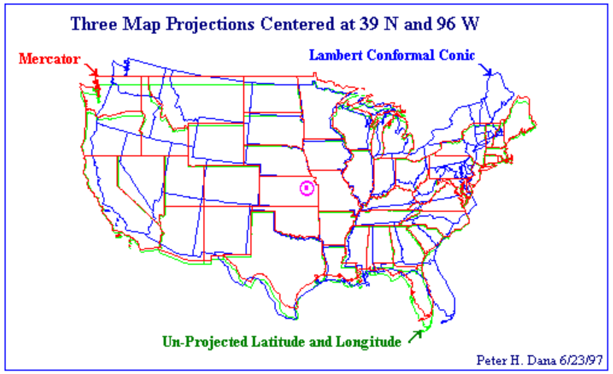


Globe Projections

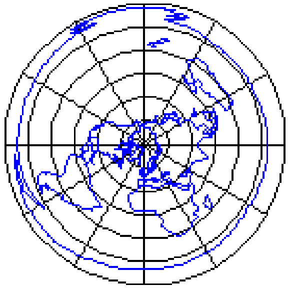
http://www.versamap.com/webdoc15.htm



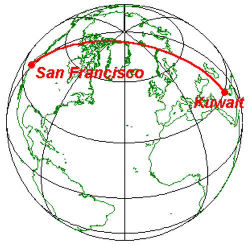
Different projections



Deformations according to projections



Minimum path



Global curvilinear distance between two points

$P_1(LO_1, LA_1)$

$P_2(LO_2, LA_2)$

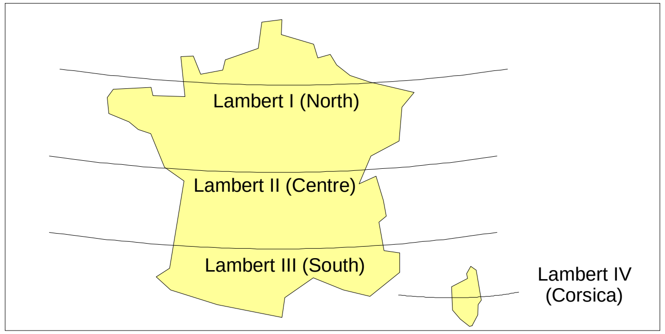
$$d(P_1, P_2) = R \times \arccos(\sin(LA_1) \times \sin(LA_2) + \cos(LA_1) \times \cos(LA_2) \times \cos(LO_2 - LO_1))$$

$R = Mean\ Earth\ Radius$

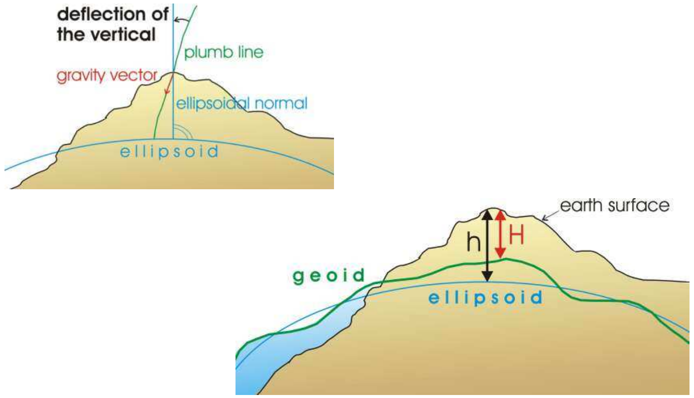
$R = 6{,}378.135\ Km\ \ (Equatorial\ \ Radius)$

$R = 6{,}356.766\ Km\ \ (Polar\ Radius)$

For instance in France (Lambert System)

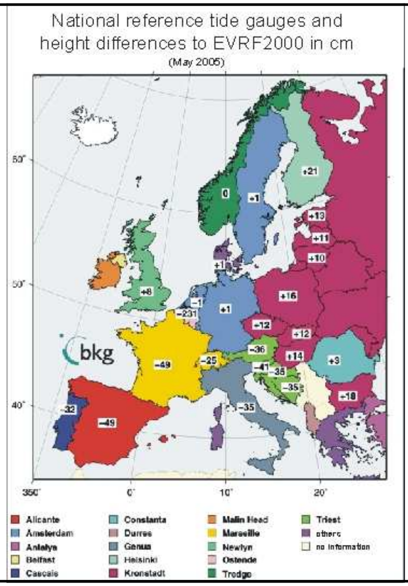


Definitions of altitudes

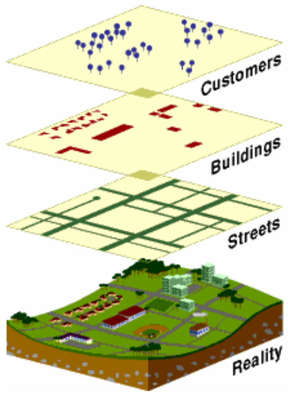


Consequences

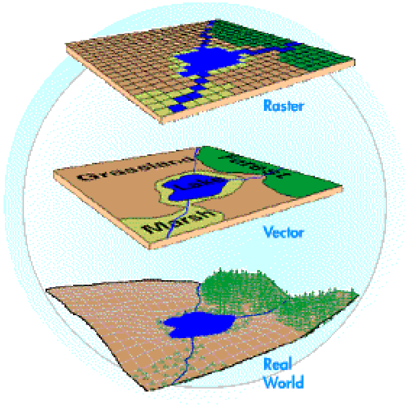
Fragment of the Dutch topographic map showing the border of Belgium and the Netherlands. The Mean Sea Level of Belgium differ -2.34m from the MSL of The Netherlands. As a result, contour lines are abruptly ending at the border.



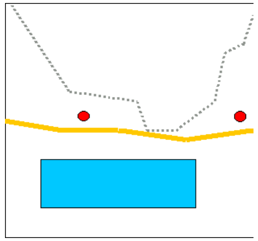
Geographic data layers



Real world models



Geographic Formats



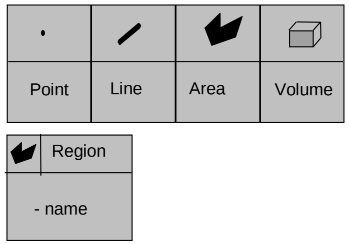
Vector Format

R	L	L	L	L	L	L	R
L	R	L	L	L	L	R	R
L	L	R	R	R	L	R	L
L	L	H	L	L	R	L	H
B	B	B	B	B	B	B	B
S	C	C	C	C	C	S	S
S	C	C	C	C	C	S	S
S	S	S	S	S	S	S	S

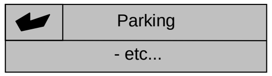
Raster Format

Formalism with spatial pictograms and icons

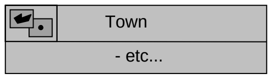
- Small drawing (graphic symbol) representing some geometric type
- Spatial
  Pictograms
- Usage



Other examples

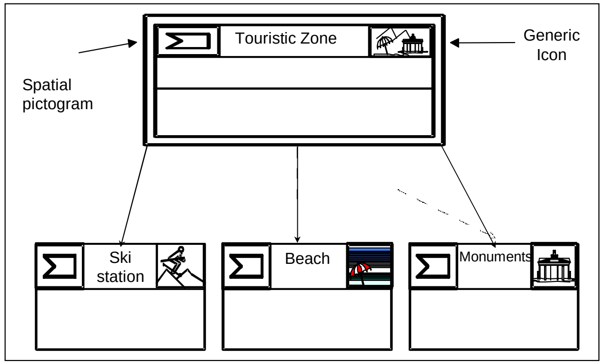


Simple Pictogram

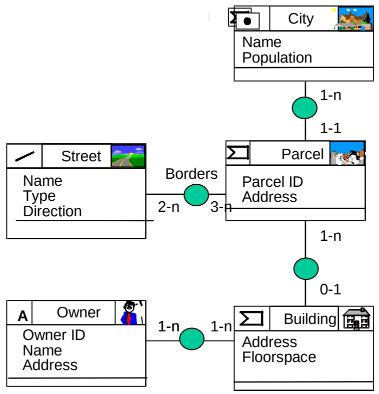


Alternative Pictogram

Pictogram and icons



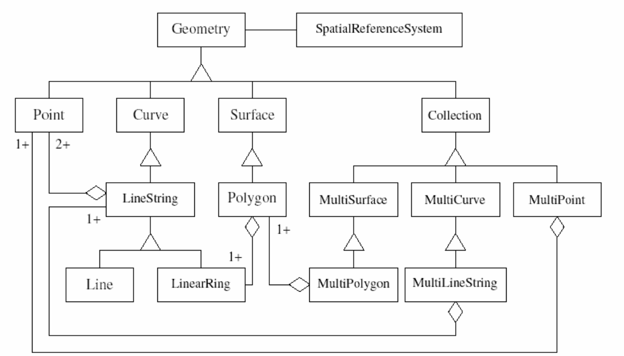
Conceptual model with pictograms and icones



OpenGIS Model

- Consortium of companies, research centers and administrations
- Interoperability of geographic applications
- Propositions of standards
- http://www.opengis.org

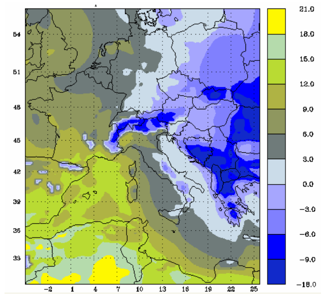
OpenGIS Model

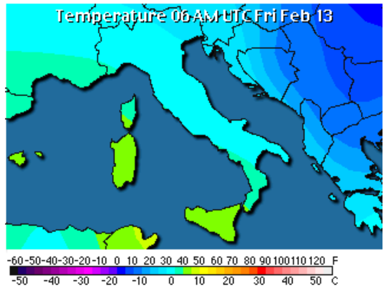


Continuous Phenomena

- Theory of continuous fields
  - Scalar fields
  - Vector fields
- Applications
  - Meteorology
  - Sea studies
  - Terrains, soils
  - Etc.

Examples





Continuous Field Modeling

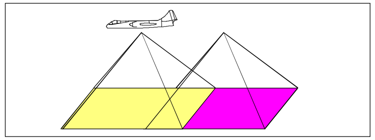
- Impossible to know the function everywhere
- Necessity of sampling points
- Necessity of interpolating functions
- Modeling (two levels)
  - Field as object (ex Temperature in a region)
  - Field as abstract data type (ex value of temperature in some point)

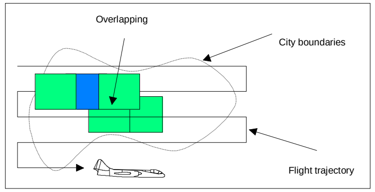
2.2 – Data Acquisition Modes

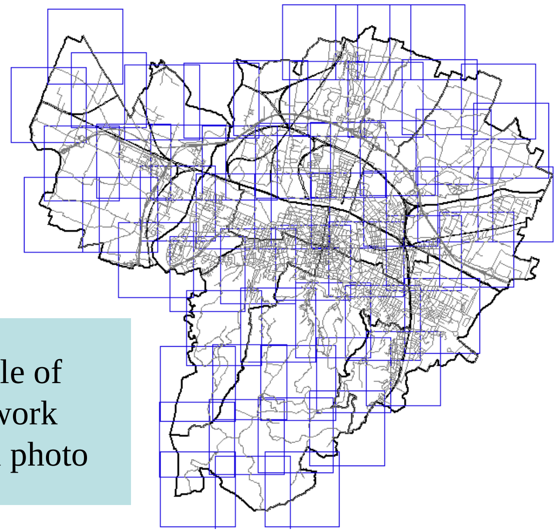
- Surveys
- Digitizing
- Aerial photos
- Satellite images
- Laser
- GPS
- Sensors

Theodolite

Aerial photos





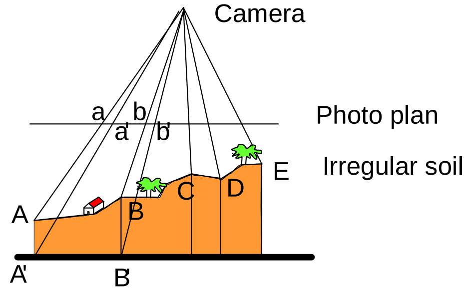


Example of framework for aerial photo

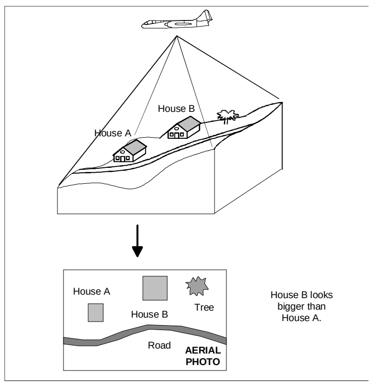
Aerial photos Characteristics

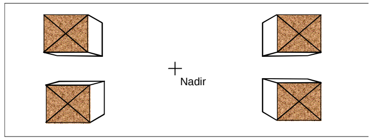
- Altitude: from 5 00 to 3,000 meters
- Format: 23 cm × 23 cm
- Scale from 1:3,000 to 1:25,000
- Photos pair ➔ relief
- Parallaxes ➔ determination of altitudes
- Photo-interpretation
- Orthophotos (mosaicking)

Distortions

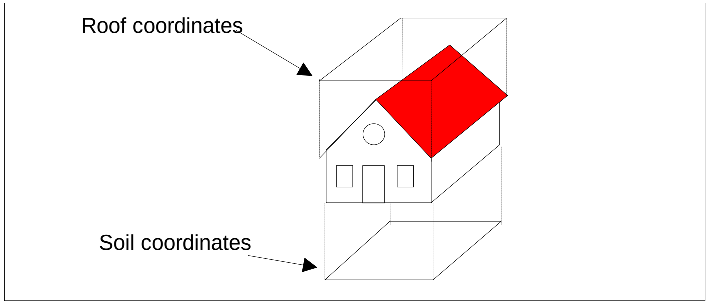


Distortions





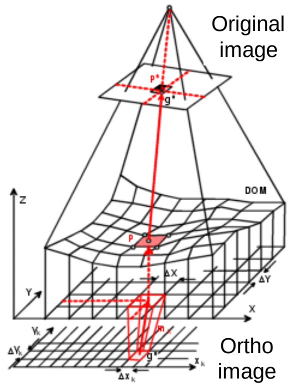
Roof and soil coordinates

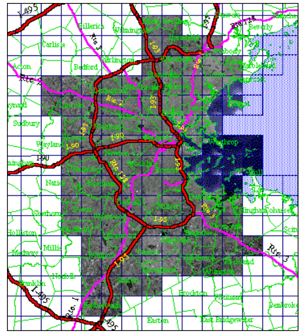


Realization of orthophotos

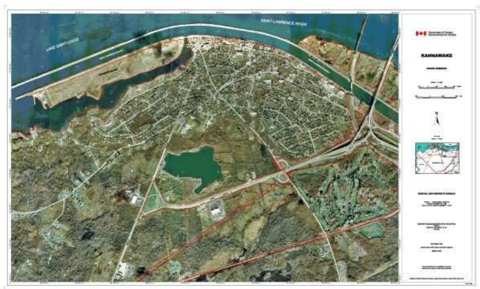
- Overlap : 60 % longitudinal
- 25 % lateral
- Selecting control points
- Elastic transformation (rubber sheeting)
- Corrections of distortions
- Cutting along roads or rivers

Orthophoto principles

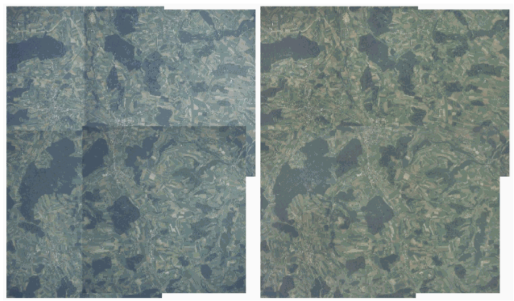




Orthophoto (result)

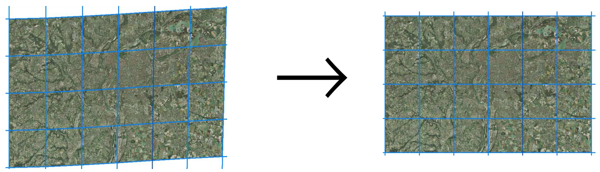


Color balancing



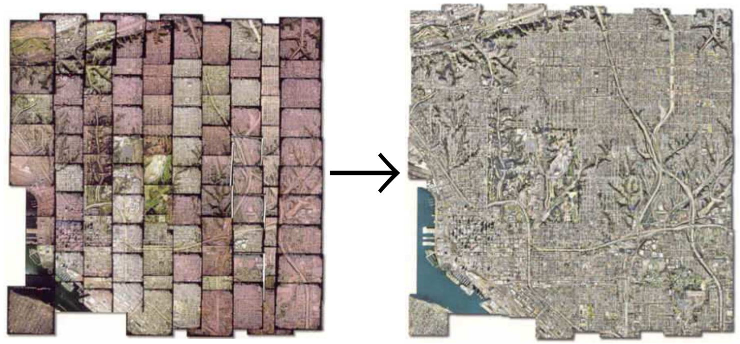
Before After

Image deformation

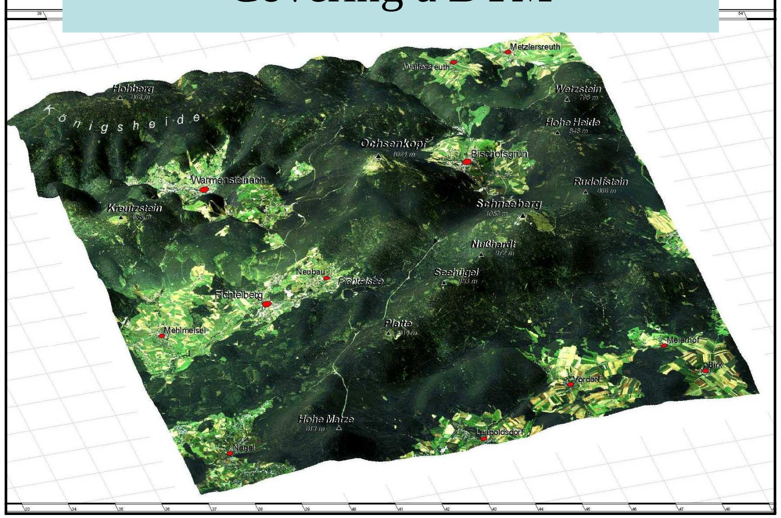


Before After

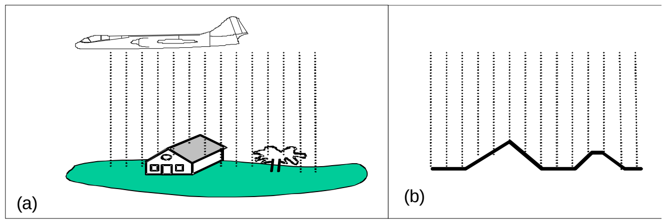
Orthophoto mosaicking



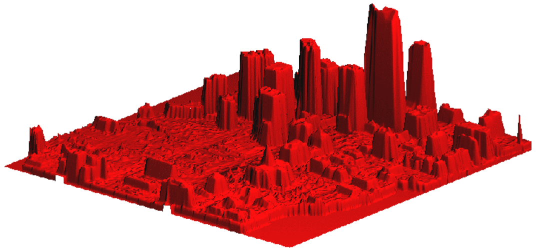
Covering a DTM

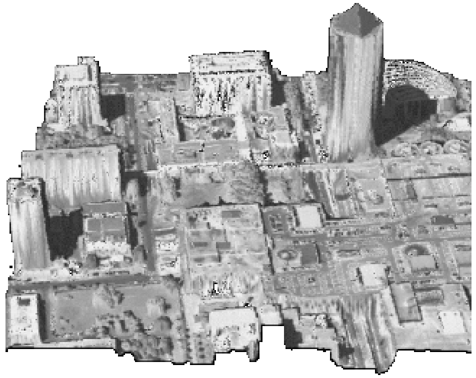


Principle of Laser Scanning



Laser range scanning

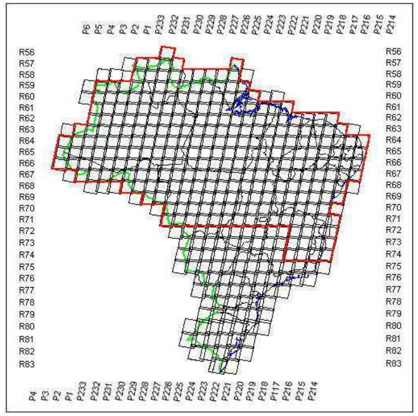




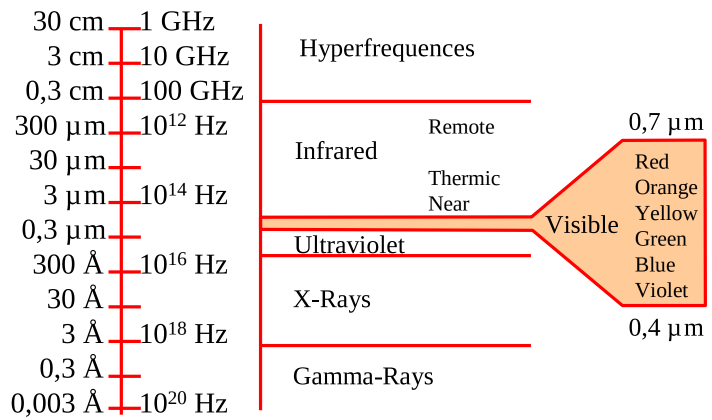
3D modeling
City of Lyon

Satellite Images

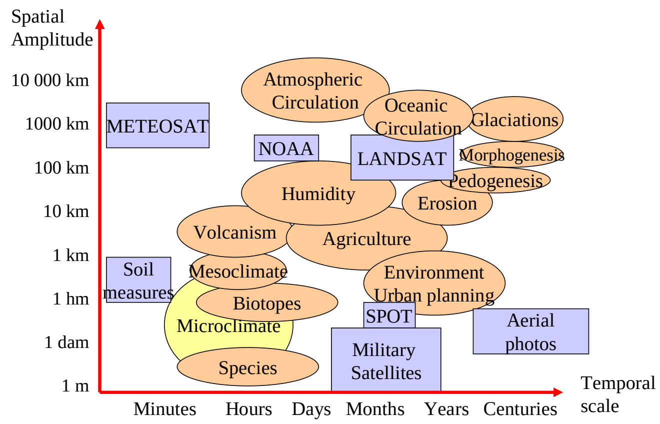
Brazil: Satellite image organization



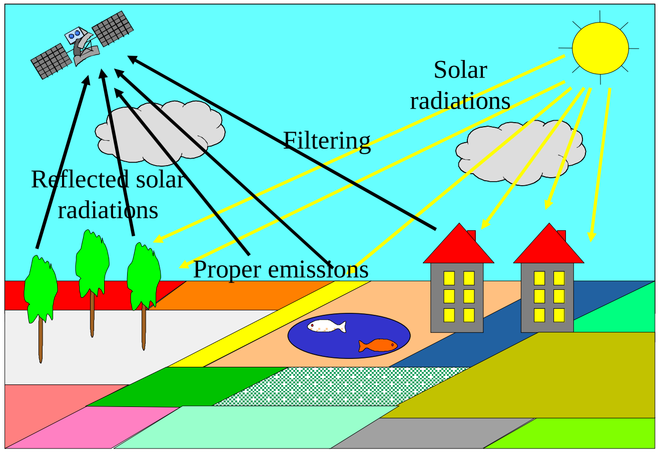
Waves

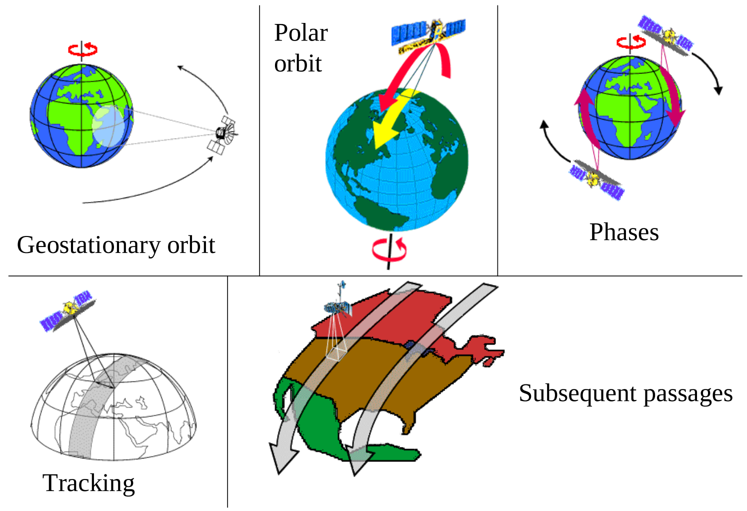


Satellites and usage



Remote sensing principles

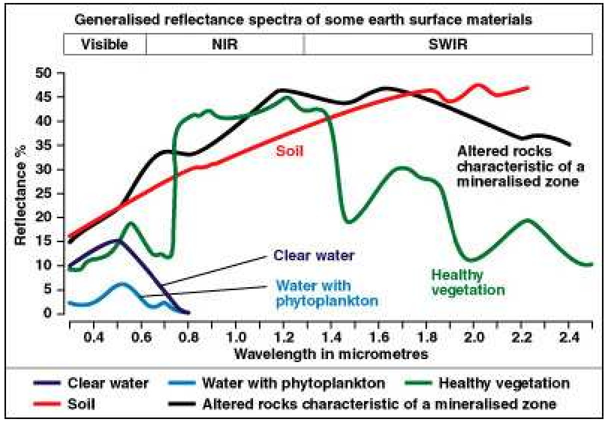




http://www.ccrs.nrcan.gc.ca/ccrs/eduref/tutorial/chap2/c2p2e.html

Spectral signature

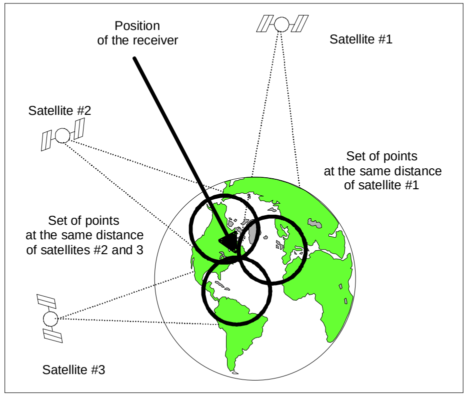
http://www.rsacl.co.uk/remote_sensing/main.htm



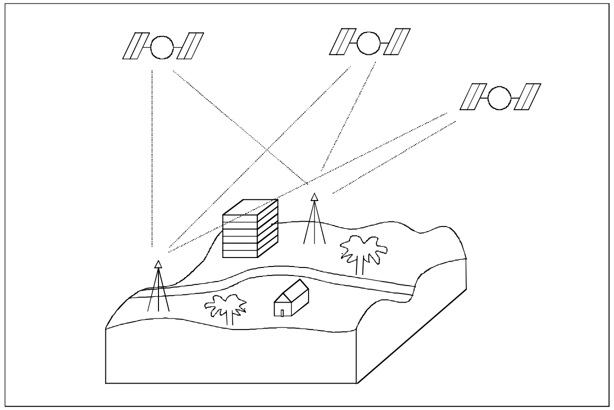
Ikonos

Global Positioning System

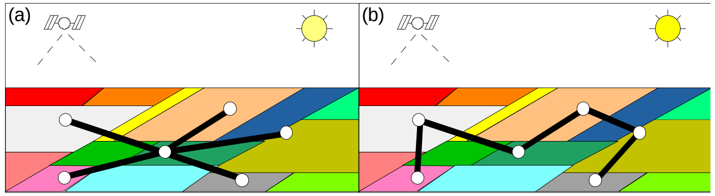
Principles of GPS



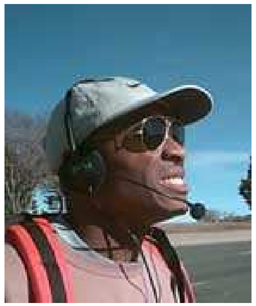
Differential GPS



Measuring with GPS

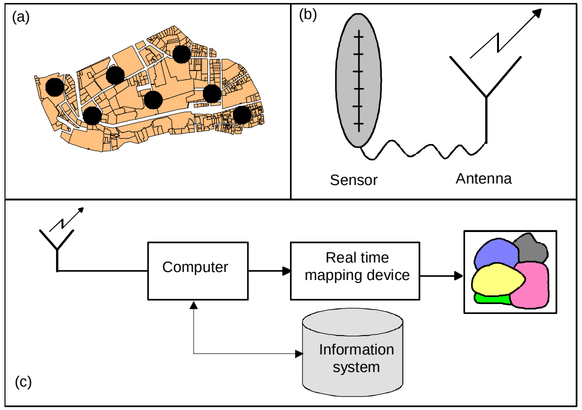


Voice Technology

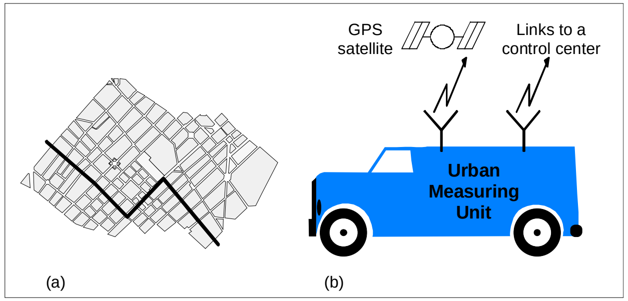


- Provided by Datria / Stantec
- GPS-positioned messages are stored into computers
- Interesting for example to describe certain situations

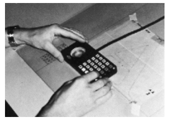
Measures with fixed sensors



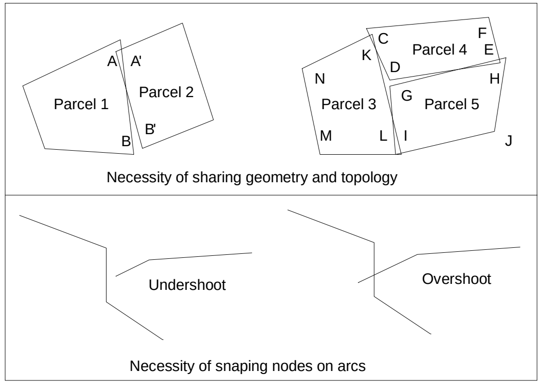
Acquisition with vehicle



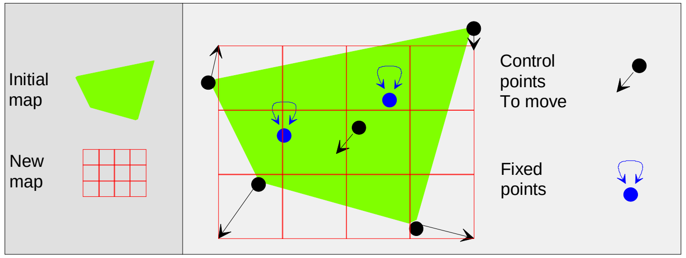
Digitizing



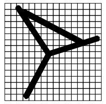
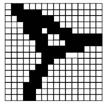
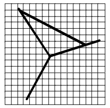
Common errors in digitizing



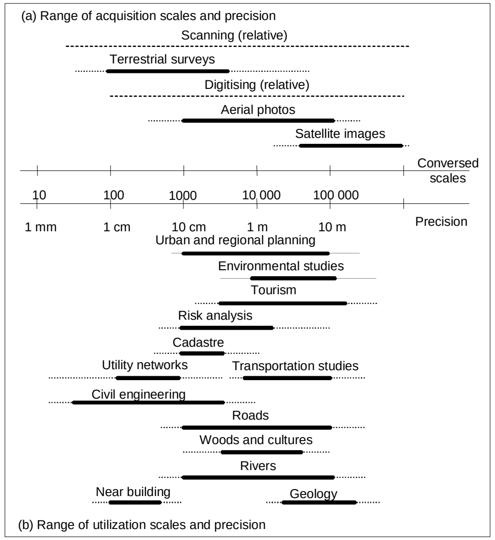
Rubber-sheeting



Map scanning

Original segments	
After scanning	
After skeletonizing	

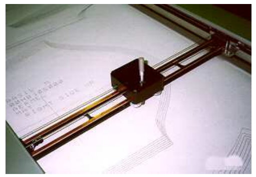
Choosing scales



2.3 – Output devices

- Various devices
- Interactivity level
- Level of Graphic Semiology

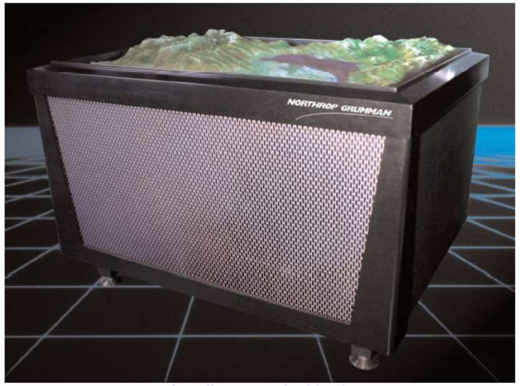
Flatbed plotter



Drum plotter

Control Room at NASA

Generator of terrain models



http://www.touchtable.nl

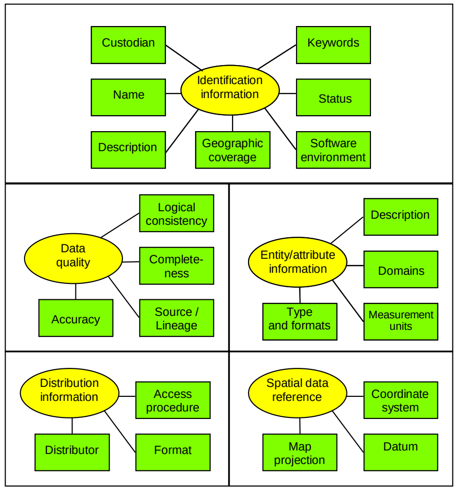
Example of Terrain Model

Tangible Interface of Geodan

http://www.geodan.nl/uk/project/virtual_maquette/HPCfeb05_small.wmv

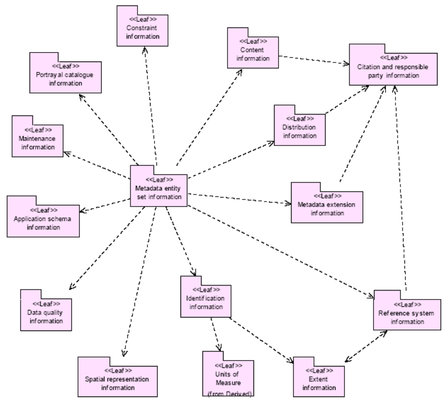
2.4 – Metadata

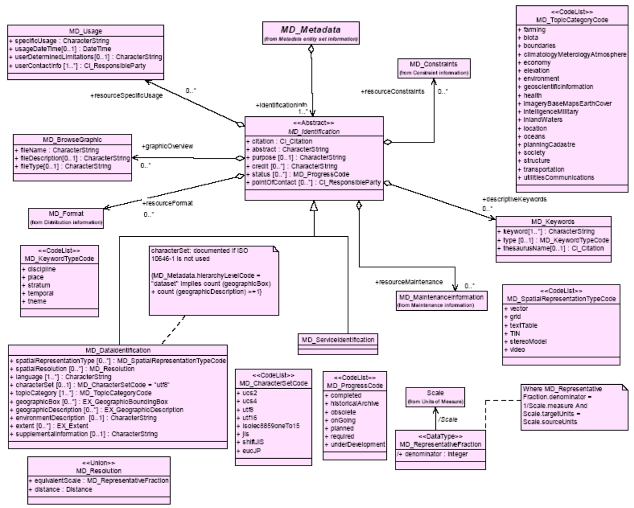
- Data about data
  - Origin
  - Quality
  - Consistency
  - Completeness
  - Updating
- Standards
  - CEN
  - FGDC



FGDC Metadata

Norm ISO19115





2.5 – Consistency of Spatial Data

- Spatial integrity constraints
- Semantics of spatial integrity constraints
- Towards Quality control
- Conclusion

Quality Control

- New databases
  - At creation, powerful controls
  - Then, checking after each updating, insertion and deletion
- Old databases
  - Powerful quality control procedures
  - Correction of erroneous objects
  - Then, checking after each updating, insertion and deletion

Spatial integrity constraints

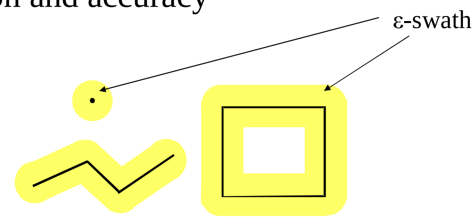
- Elements about integrity constraints in databases
- Consistency and precision
- Semantics and data structures
- Definition of spatial integrity constraints (SIC)
- Examples for terrains
- Constraints and derived spatial data

Elements about integrity constraints in databases

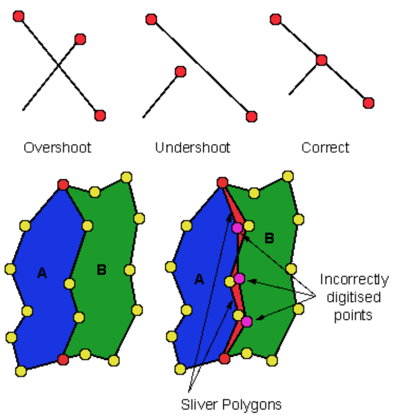
- Likelyhood control of value
- Existential integrity
- Referential integrity
- User-defined constraints

Consistency and precision

- Quality control
- Precision and accuracy



Common errors



Semantics and data structure

- A structure claiming "*I am a square*", is it really a square ?
- Necessity of controls
- In some case, add complimentary data

- Example
- NOQUAD (#quad, #pt1, #pt2, #pt3, #pt4)
- PUNTO(#pt, x, y)

Standards

- ISO 19113
  - Geographic Information – Quality Principles
- ISO 19114
  - Geographic Information – Quality Evaluation Procedures
- ISO 19115
  - Geographic Information –Metadata

Dimensions of quality

- **Geometric**

  Primitive → Objects → Class of objects → Set of data

- **Semantic**

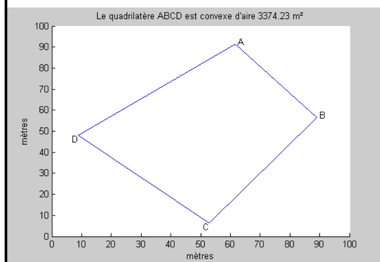
  Value → Domain → Attribute → Class of objectsi → Semantics

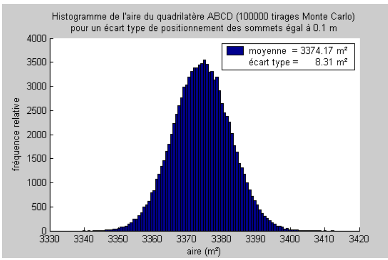
- **Temporal**

  Time primitive → Complete history → All temporal aspects → Set of temporal data

Example of consequences of error

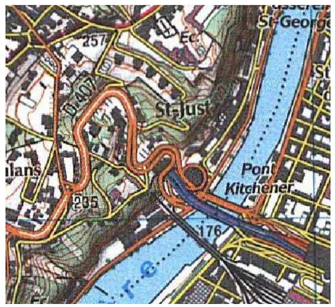
4-sided polygon and its area



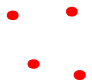
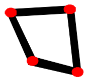
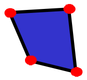


There is no error

The road passes at the third floor of a building!!!



Representation of semantics of different spatial objetcs

Set of points		*R(#object, (#point)*)* No rule No SIC
Closed polyline		*R(#polyline, (#point)*)* Rule: point in a polyline SIC: 2 neighbouring points must differ
Polygon		*R(#polygon, (#point)*)* Rule: point in a polyline SIC: closure and no degenerescence

Definition of spatial integrity constraints

- IC = Predicate against a database
- SIC = Predicate with geometric and/or topological conditions

Examples of a terrain model

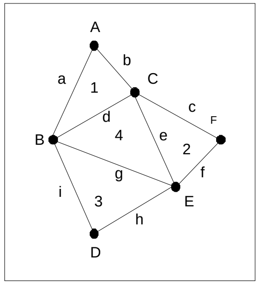
R0 (#terrain, #triangle)

R1 (#triangle, #segment1, #segment2, #segment3)

R2 (#segment, #point1, #point2, #triangle1, #triangle2)

R3 (#point, x, y, z)

Example with tables

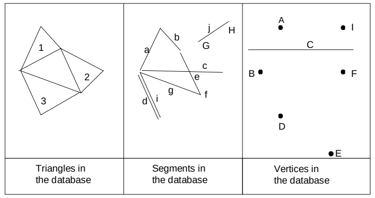


R1	*#triangle*	*#segment1*	*#segment2*	*#segment3*
	1	a	b	d
	2	e	c	f
	3	g	h	i
	4	d	e	g

R2	*#segment*	*#point1*	*#point2*	*#triangle1*	*#triangle2*
	a	A	B	1	null
	b	A	C	1	null
	c	C	F	2	null
	d	B	C	1	4
	e	C	E	2	4
	f	E	F	null	2
	g	B	E	4	3
	h	D	E	null	3
	i	D	B	3	null

R3	*#point*	*x*	*y*	*z*
	A	2	5	10
	B	1	3	3
	C	3	4	12
	D	2	1	3
	E	4	2	8
	F	5	3	15

Examples of inconsistencies



R1	#triangle	#segment1	#segment2	#segment3	
	1	a	null	d	unknown segment2
	2	e	c	f	
	3	g	h	h	modified segment3
					triangle 4 missing

R2	#segment	#point1	#point2	#triang1	#triang2	
	a	A	B	1	null	
	b	A	C	null	null	double null
	c	C	F	2	null	point1 changed
	d	B	C	1	4	point2 changed
	e	C	E	2	4	same triangles
	f	E	F	null	2	same extrémities
	g	B	E	4	3	error in triangle2
	i	D	B	3	null	
	j	G	H	null	7	segment en trop
						segment manquant

R3	#point	x	y	z	
	A	4	3	10	vertex within triangle 2
	B	1	3	3	
	C	null	4	null	unknown coordinates
	D	2	1	3	
	E	0	2	8	coordinate outside
	F	5	3	15	
	I	5	5	18	isolated point

Derived data structures

Example:

Model 1
R1 (#triangle, (#vertex)^3)
R2 (#vertex, x, y).

Model 2
R1bis (#triangle, (#angle)^3)
R2bis (#angle, value-in-degrees).

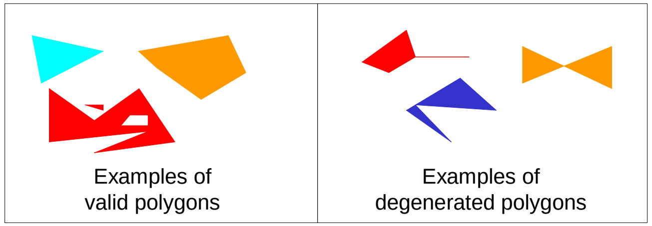
Semantics of SIC

- Geometry and topology
- Example of chaining SIC
- Constraints and representation

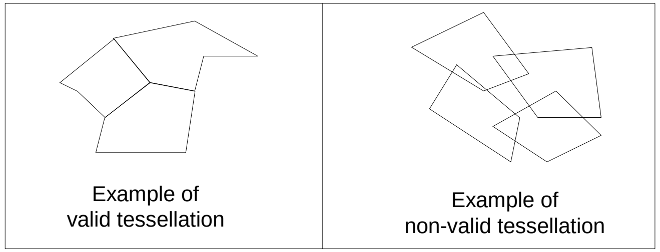
Geometry and topology

- Usage of topology
- Usage of trigonometry
- Usage of some theorems

Examples of valid and degenerated polygons



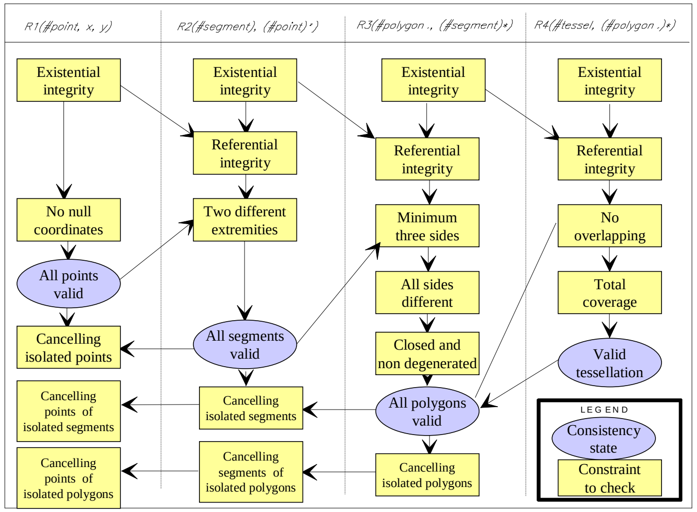
Valid and non-valid tessellations



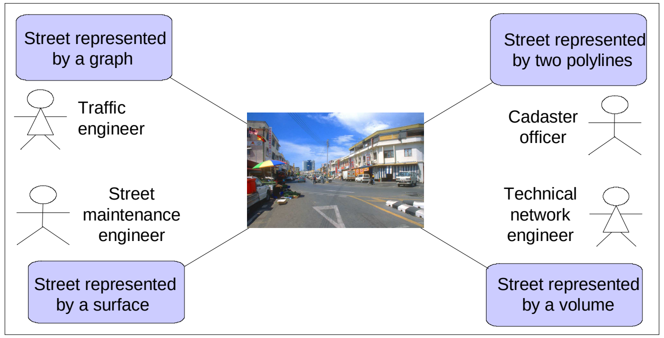
Example of chaining the checking of constraints

Let consider the following database:

R1 (#point, x, y)
R2 (#segment, (#point)²)
R3 (#polygon, (#segment))*
R4 (#tessellation, (#polygon))*



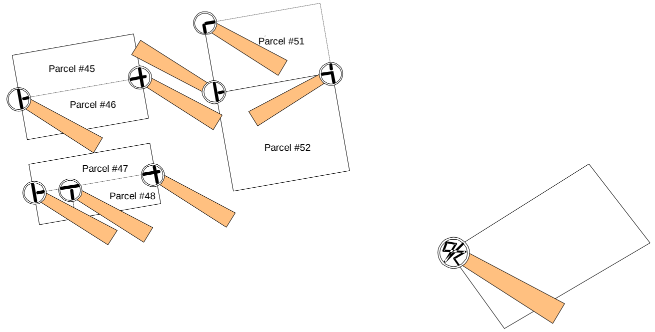
Constraints and multiplicity of representations



Geometry and topology

- Validity of polygonal tessellations
- Validity of networks

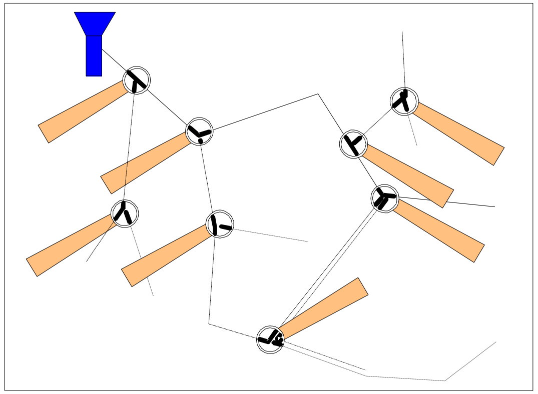
Example of cadaster



Validity of polygonal tessellations

- 1 – validity of points
- 2 – validity of segments
- 3 – validity of polygons
- 4 – validity of tessellation
- 5 – formula of Euler-Poincaré : $P+V = S+1$

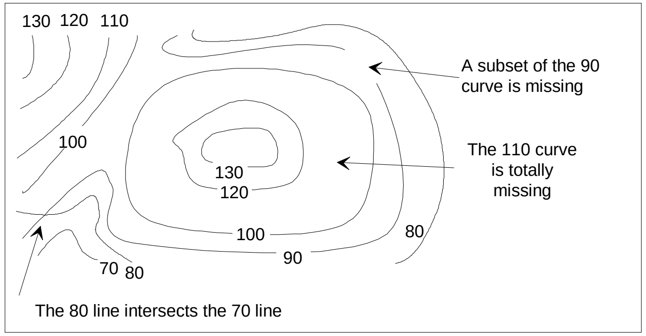
Example of network with errors

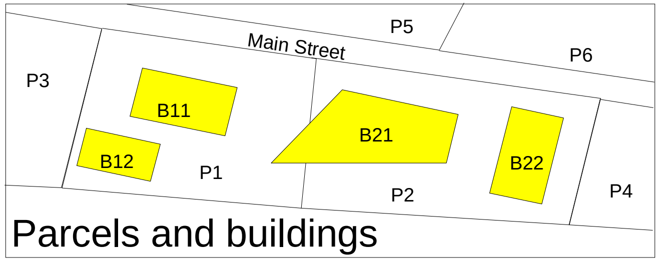


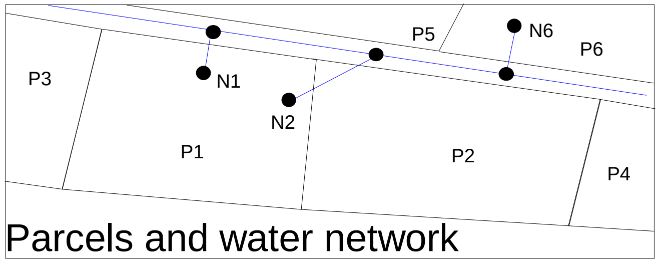
Validity of networks

- 1 – validity of nodes
- 2 – validity of edges/arcs
- 3 – network connexity
- 4 – orientation of the network (iff orientated graph)

Example for contours



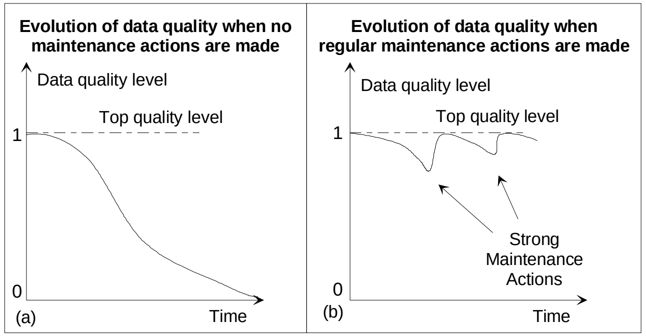




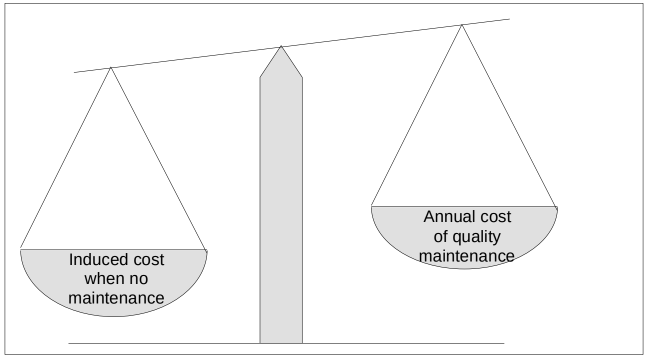
Introduction to Quality control

- New databases
  - at the DB inception
  - checking after each update/delete/insert
- Old databases
  - powerful control procedures and correction of erroneous objects
  - checking after each update/delete/insert

Evolution of quality



Cost Balance



Components of quality

- Lineage
- Accuracy
- Resolution
- Feature completeness
- Timeliness
- Consistency
- Quality of metadata

Conclusion about consistency

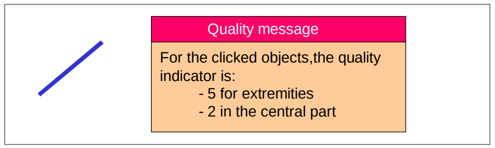
- Importance of quality control
- Cost balance
- *What is the cost of an error*???
- Strong quality control at the creation
- Quality control and lifecycle

Quality Visualization

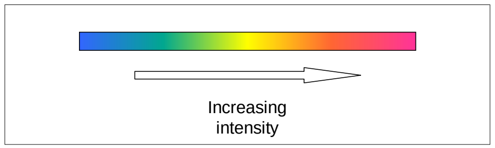
- Various modes
- Example in photogrammetry
- Metadata
- Application

Various modes

With messages

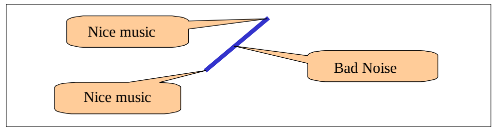


With colors

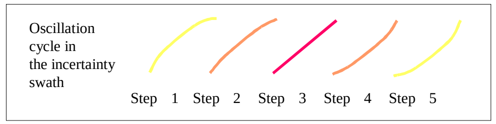


Other modes

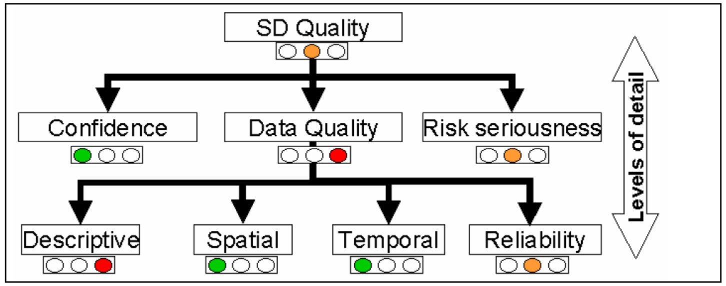
With noise



With animation

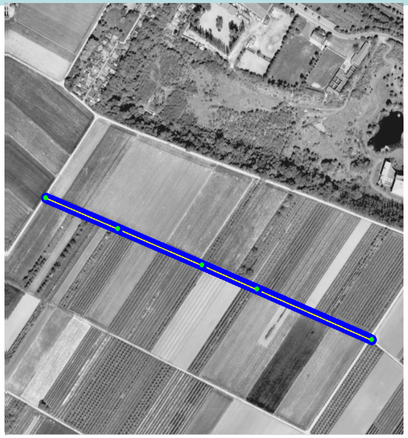


Quality indicator (Traffic light metaphor)



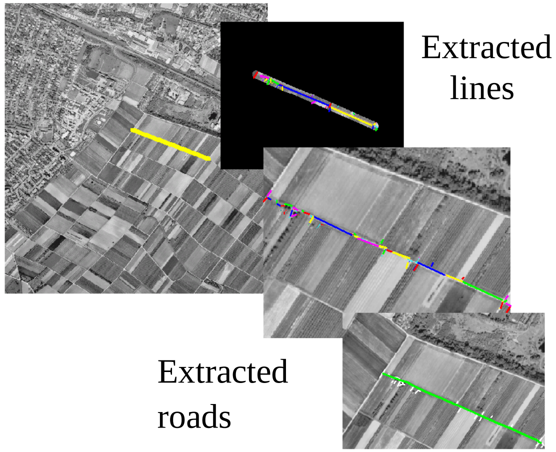
Geometric Precision with colors

Definition of a buffer for road extraction

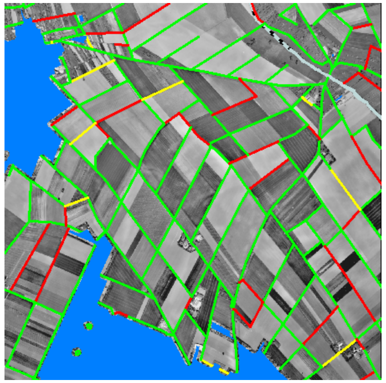


- The buffer depends of
  - positional accuracy
  - of quality of path axes
  - and of attribute quality

Road extraction

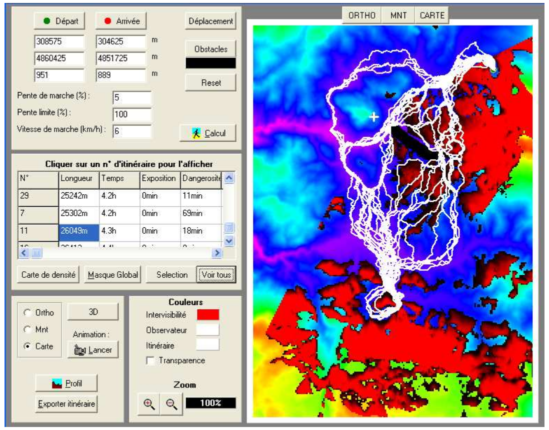


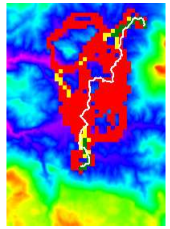
Results in a rural zone



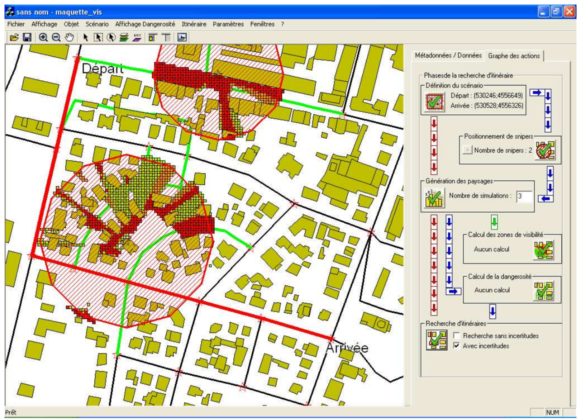
Results:
83 % accepted
13 % non accepted
4 % indecision

Optimal path in a terrain taking quality into account

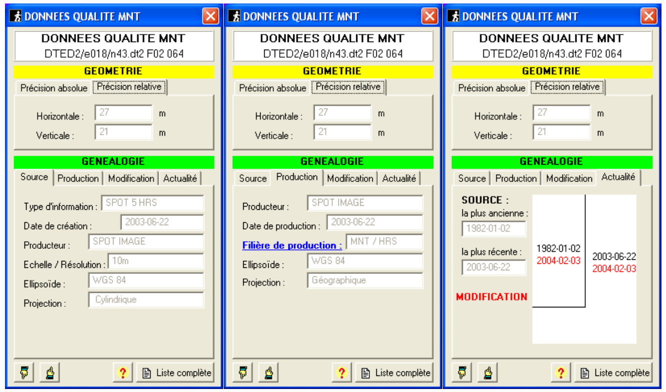




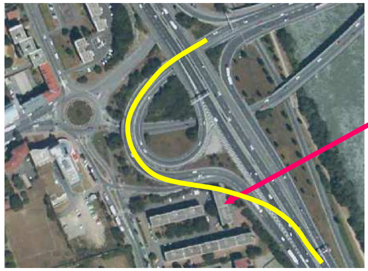
Intervisibility with errors



Metadata



Consequences of an error



At motorway construction,
This building was not
Stored into the DB

Conclusion about quality control

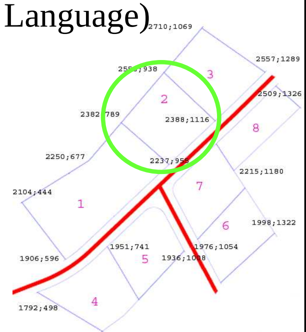
- Importance of quality control
- Cost of checking
- Cost of correction
- Cost in case of errors ????

2.6 – Extensions of XML

- Objective: processing geographic vector data on Internet
- Interest:
  - To reduce server loads
  - To alleviate interchanges between clients and servers
  - To allow queries at client side
  - To install local processing at client level

Extensions

- SVG (Scalable Vector Graphic)
- GML (Geography Markup Language)
- LandXML



SVG

- To increase graphic capabilities of XML
- Not originally planned for cartography
- The maps is seen as a drawing
- Possibility of interactivity
- Possibility to modify some drawing attributes

SVG

```
<desc>Parcel Lot #2</desc>
  <g>
      <polyline points="938.15,-2556,24
                             789.84,-2382,09"/>
      <polyline points="789.84,-2382,09
                             952.92,-2237,08"/>
      <polyline points="955.92,-2237,08
                             1116.15,-2388,54"/>
      <polyline points="1116.15,-2388,54
                             938.15,-2556,24"/>
  </g>
```

GML

- Really encoding geographic vector data
- Targeted applications: mapping and spatial analysis
- Creation of a small GIS at client level
- Capacity of using spatial and non-spatial attributes
- Opening towards interoperability

GML

```
<exMember>
   <Parcel>
      <gml:name>Lot #2</gml:name>
      <area>52129.7703</area>
      <gml:centerOf>
         <gml:Point>
            <gml:coordinates>2392.91 950.79</gml:coordinates>
         </gml:Point>
      </gml:centerOf>
      <gml:extentOf>
         <gml:Polygon srsName="http://www.opengis.net/gml/srs/epsg.xml#4326">
            <gml:outerBoundaryIs>
               <gml:LinearRing>
                  <gml:coordinates>
            2556.24 938.15 2382.09 789.84 2382.09 789.84 2237.08 955.92
            2237.08 955.92 2388.54 1116.15 2388.54 1116.15 2556.24 938.15
                  </gml:coordinates>
               </gml:LinearRing>
            </gml:outerBoundaryIs>
         </gml:Polygon>
      </gml:extentOf>
   </Parcel>
</exMember>
```

LandXML

- Format specification in civil engineering, surveying and architecture
- Transferring data between actors
- 2D and 3D

LandXML

```
<Parcel name="Lot #2" area="52129.77" >
    <Center>2392.91 950.79</Center>
    <CoordGeom>
        <Line length="228.74" dir="229.58" >
                <Start>2556.24 938.15</Start>
                <End>2382.09 789.84</End>
        </Line>
        <Line length="220.48" dir="318.87" >
                <Start>2382.09 789.84</Start>
            <End>2237.08 955.92</End>
        </Line>
        <Line length="220.49" dir="43.38" >
                <Start>2237.08 955.92</Start>
                <End>2388.54 1116.15</End>
        </Line>
        <Line length="244.56" dir="136.70" >
                <Start>2388.54 1116.15</Start>
                <End>2556.24 938.15</End>
        </Line>
    </CoordGeom>
</Parcel>
```

Comparison - usage

Domains	GML	SVG	LandXML
Urban planning	X	X	XX
Environmental planning	X	X	
Cadaster	XX	X	XX
Statistical mapping		XX	
3D	X		X

2.7 – Conclusions

- 80 % of information throughout the world have some spatial component
- Geographic Databases among the biggest in the world
- Usage for other domains
  - Geomarketing
  - Real estate management
  - Location-Based Services
  - Pervasive Information Systems