An open-source based toolchain for the georeferencing of old cadastral maps

Follin Jean-Michel, Fahrasmane Maïté, Simonetto Elisabeth

jmfollin@cnam.fr
Outline

- Background and issues
- Materials (study area and data)
- Methodology
- Results
- Conclusion and outlooks
2 issues related to 2 fields of research in G&E laboratory:

- **Methodological**:
  - processing old cadastral maps data in a semi-automated and reproducible way:
    - vectorisation → georeferencing → tiling
  - modelling of spatio-temporal database

- **Thematic**:
  - historical studies of demography, human activities, and land cover
  - in relation with changes of the cadastral parcels’ shapes over time
Specific goal of this study:
How to find and apply the "best" spatial transformation to register old cadastral map vector data?
Implementation of the method using open source software and programming language.
Studying old maps, a very hot topic!

Scanian Economic-Demographic Database

Ritratti di città in un interno

ALPAGE
Some preliminary definitions

- **geometric transformation**: mathematical function that modifies the coordinates of a source dataset in order to make them fit into the system of a target dataset.

- In the case of **georeferencing or map registration** the target dataset has ground-based coordinates (such as latitude/longitude or UTM coordinates).

- **Ground Control Point (GCP)**: during the georeferencing process, a point that can be identified both in the source and in the target dataset (thus has known coordinates in both systems). It can be used:
  - either to compute the geometric transformation,
  - or to control it by computing a residual in the validation step.
The geometric transformations used for georeferencing

Two traditional approaches for map registration:

- **Global**: one transformation model for the whole map.
  Appropriate when map distortions are homogeneous

- **Local**: several transformation models → one for each region.
  Appropriate for historical maps where local distortions cannot be neglected.

A locally sensitive global approach proposed in (Herrault et al., 2013) for registration of historical maps:

- Kernel smoothing regression: one transformation model taking into account local distortions

Source: Takeda et al., 2007
The Napoleonic cadastre...

One cadastral section is drawn on many cadastral sheets.

One index of cadastral sections is defined for one commune.

Parcels are mapped on one cadastral sheet.
...a goldmine of information

- Cadastral sheets are available using web-based viewers (not open data)
Overview of different cadastral maps

1813 Napoleonic plan

1850 Napoleonic plan

Republished plan (1972)
New plan (1974)

PCI Vector data (2010)
Synthesis of the study data

- **64 images** of cadastral sheet covering two communes of southern Sarthe (France): Vaas and Aubigné-Racan

### Aubigné-Racan

<table>
<thead>
<tr>
<th>Year of the series</th>
<th>Number of images</th>
<th>Map scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>1813</td>
<td>3</td>
<td>1:2500</td>
</tr>
<tr>
<td>1850</td>
<td>6</td>
<td>1:2000</td>
</tr>
<tr>
<td>1972</td>
<td>6</td>
<td>1:2000</td>
</tr>
</tbody>
</table>

### Vaas

<table>
<thead>
<tr>
<th>Year of the series</th>
<th>Number of images</th>
<th>Map scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>1813</td>
<td>9</td>
<td>1:1250, 1:2500</td>
</tr>
<tr>
<td>1850</td>
<td>22</td>
<td>1:500, 1:1000, 1:2000</td>
</tr>
<tr>
<td>1974</td>
<td>18</td>
<td>1:1000, 1:2000</td>
</tr>
</tbody>
</table>

- Rural landscapes with few built-up areas
- Parcels’ layout slightly modified since 1813
- Cadastre significantly altered between 1813 and 2010 (railway and highway construction, ...)

---

**Background and issues**

**Study area and data**

**Methodology**

**Results**

**Conclusion and outlooks**
A global methodological toolchain

two possible strategies for building cadastral spatio-temporal database

Background and issues

Study area and data

Methodology

Results

Conclusion and outlooks

Old cadastral maps (raster format)

Reference cadastral map (vector format)

Tiling of images

Vectorisation of the parcel polygons and insertion of attributes values

Selection of Ground Control Points

Georeferencing by using best transformation model

Vectorised polygons of parcels

Old cadastral maps (raster format)

Vectorisation of the parcels and insertion of attributes values

Tiling of vectorised parcel

Creation of a temporal cadastral database storing filiation links between parcels

Calculation of indicators pointing out evolution of the cadastre throughout time

Option 1

Option 2
Selection of ground control points

- QGIS' Georeferencer Plugin
- GCPs:
  - Parcels and buildings corners
  - 30 per sheets
Comparison of geometric transformations

- Global approach
  - 1st, 2nd and 3rd order polynomial (P-ord1, P-ord2 and P-ord3)
  - Thin Plate Spline (TPS)
  - Helmert transform (H)

- Locally sensitive
  - Gaussian kernel regression model (RNG).

- Local approach
  - Delaunay triangulation combined with first order polynomial function (TD);

(points file including 30 GCPs both known in the image and in the "PCI-vecteur" systems)

Computing of the transformation model by the least squares method

Georeferencing → Model applied to all the GCPs

Background and issues
Study area and data
Methodology
Results
Conclusion and outlooks
Comparison of geometric transformations

“Leave one-out cross validation” statistical method

Residual: distance between georeferenced GCP and reference point.

Root Mean Square Error (RMSE) computed from the n residuals values of GCPs → indicator of the georeferencing quality.

GCPs:
- Reference point
- Training points
- Control point

Source: DGFiP; Fahrasmane, 2016
Assessment of the georeferencing accuracy (Aubigné-Racan)

Results

Background and issues

Study area and data

Methodology

Global approach
- 1st, 2nd and 3rd order polynomial (P-ord1, P-ord2 and P-ord3)
- Thin Plate Spline (TPS)
- Helmert transform (H)

Locally sensitive
- Gaussian kernel regression model (RNG).

Local approach
- Delaunay triangulation combined with first order polynomial function (TD)
Assessment of the georeferencing accuracy (Vaas)

Results

Background and issues

Study area and data

Methodology

Global approach
- 1st, 2nd and 3rd order polynomial (P-ord1, P-ord2 and P-ord3)
- Thin Plate Spline (TPS)
- Helmert transform (H)

Locally sensitive
- Gaussian kernel regression model (RNG)

Local approach
- Delaunay triangulation combined with first order polynomial function (TD)

Graphs showing EMQ moyen (en m) for different methods and years: Vaas 1813, Vaas 1850, and Vaas 1974.
Assessment of the georeferencing accuracy (synthesis)

Results obtained with the RNG method (always the best)

<table>
<thead>
<tr>
<th>Year of the series</th>
<th>Village</th>
<th>Average of RMSEs in meters</th>
<th>Standard deviation of RMSEs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1813</td>
<td>Aubigné-Racan</td>
<td>3.51</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>Vaas</td>
<td>4.71</td>
<td>1.01</td>
</tr>
<tr>
<td>1850</td>
<td>Aubigné-Racan</td>
<td>1.04</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Vaas</td>
<td>3.32</td>
<td>0.90</td>
</tr>
<tr>
<td>1972 - 1974</td>
<td>Aubigné-Racan</td>
<td>0.50</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Vaas</td>
<td>2.52</td>
<td>1.36</td>
</tr>
</tbody>
</table>
Overlay of reference data with georeferenced 1813 data of Aubigné-Racan
Conclusion

- Consistent results:
  - more recent data → increasing accuracy of georeferencing
  - results on Aubigné-Racan always better

- Encouraging results:
  RNG always performs best especially in the cases of:
  - oldest maps,
  - parcel layout altered significantly over time.

- Concrete results:
  a set of Python programs including:
  - choice of the more efficient geometric transformation
  - application of this method for georeferencing vector data.
Outlooks

Further tests:
- on cadastral maps representing different landscapes

The step before georeferencing:
- semi-automated vectorisation (with a method such as the Hough transform for instance)

The step after georeferencing:
- tiling of the georeferenced data (by correcting topological errors)

The spatio-temporal database:
- implementation
- application, for example, in a long-term study of the evolution of cadastres w/r to socio-demographic, political and economic factors
Outlooks
Tiling

Background and issues
Study area and data
Methodology
Results
Conclusion and outlooks
Thank you for your attention