The open discussion version of this paper is available at: Follin J, Fahrasmane M, Simonetto E. (2016) An open-source based toolchain for the georeferencing of old cadastral maps. PeerJ Preprints 4:e2222v3 https://doi.org/10.7287/peerj.preprints.2222v3

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

Jean-Michel Follin[1], Maïté Fahrasmane[1] and Elisabeth Simonetto[1]

[1]École Supérieure des Géomètres Topographes (ESGT – CNAM), Laboratoire Géomatique et Foncier (GeF), Le Mans, France

Corresponding author:
Jean-Michel Follin[1]

Email address: jmfollin@cnam.fr

ABSTRACT

More and more historical data are available on the web. In France, old cadastral maps are regularly published by the "*départements*". Such material is relevant to various applications (on-the-field search of specific objects such as old boundary stakes, historical studies of demography, human activities, land cover…). The GeF laboratory is working on the development of a complete methodological toolchain to vectorise, correct and analyse cadastral parcels and their evolution, using open source software and programming language only (QGIS, GDAL, Python). This article details the use of a part of this toolchain - georeferencing old cadastral data - on parcels located near the Loir river, in two villages of southern Sarthe: Vaas and Aubigné-Racan. After a presentation of our methodological toolchain, we will discuss our first results.

Keywords: geometric transformation, old cadastral maps georeferencing, quality assessment

INTRODUCTION

An increasing number of old geographical documents (maps, aerial photographs, etc.) are digitised and published on the web. Those documents represent a goldmine of information to humanities and/or social science researchers. Many *départements* of France offer to browse their digitised collections, including old cadastral sheets, using web-based visualisers. Yet, according to their date of creation, the information contained in those documents can be more or less detailed or accurate (Clergeot P., 2007).

The *Géomatique et Foncier* (GeF) laboratory has endeavoured to study territorial evolution by monitoring changes to the shape of the cadastral parcels over time. Part of this task involves the development of an open-source based, semi-automated and reproducible methodological toolchain. This toolchain first processes data (vectorisation, registration and tiling of old cadastral maps) and then stores the result in a spatio-temporal database. This database will be used to carry out analyses like quantification and qualification of changes, study of relations with other data, identification of recurring patterns…

In this paper, we focus on a mainstay of the first step: how to find the best spatial transformation to register old cadastral map vector data. We will survey both conventional global and local

Please cite this paper as: Jean-Michel Follin, Maïté Fahrasmane, Élisabeth Simonetto (2018) An open-source based toolchain for the georeferencing of old cadastral maps. In Marchesini I. & Pierleoni A. (Eds.) Proceedings of the 4th Open Source Geospatial Research and Education Symposium (OGRS2016). Perugia, 12-14 October 2016.
https://doi.org/10.30437/ogrs2016_paper_20

methods, in addition to a nonparametric algorithm called "gaussian kernel regression" whose relevance has been pointed out by (Herrault et al., 2013).

BACKGROUND

Several research projects have or are studying old map data. They reflect the recent interest for geo-historical databases and their use in analytical or prospective applications as shown in the following examples.

The *GéoPeuple* project created an historical spatial database by georeferencing and vectorising old french topographic maps, namely Cassini and État-Major (drawn in the 18th and 19th century respectively), at scales around 1:80,000, in order to study the densification of population (Grosso E. et al., 2012).

The ALPAGE project has developed an historical GIS encompassing pre-industrial Paris's territory. The oldest cadastral map of Paris, called "*cadastre par îlots de Vasserot*" (1810-1836), has been georeferenced and vectorised. The scale of this map, 1:200, has allowed historians to carry out in-depth analysis of Paris's urban demography (Noizet et Grosso, 2012).

The scope of the MODE RESPYR project is, among others, to recreate the history of land cover over the french Pyrenees region. Various geometric transformations applicable to an excerpt of the 19th century "Carte de la France" have been studied and compared (Herrault et al., 2013). The best transformation has been determined by a cross-validation protocol. This protocol demonstrated that the "gaussian kernel regression", defined as a locally sensitive global method, was more efficient than conventional methods.

Similar works include a study of the history of urban developments in Tokyo using a historical GIS to analyze land use and visualise landscapes back from the Edo period up to the present (Shimizu E, Fuse T, 2003). "*Ritratti di città in un interno*" project focus on the historical cadastral cartography of three main Italian cities: Bologna, Milan and Rome (Bitelli G., Gatta G., 2011). In Sweden, the Scanian Economic-Demographic Database (SEDD) project's goal is to create micro-level databases that could be used to enrich longitudinal analysis with geographic context, in the interdisciplinary field of historical demography (Hedefalk, F. et al., 2015).

Our own project is to create a long-term (from 1813 to 2010) geo-historical cadastral database with scales ranging from 1:2,500 to 1:500. Such historical data represent the most detailed description of the entirety of the French territory over a two-century stretch. It allows to accurately reconstruct the land cover throughout the period (Dupouey J-L. et al., 2007). Our project deals with two rural towns of southern Sarthe.

MATERIALS AND METHODS

Study area and data

The two towns under study are located in rural areas of the Loir valley in the south of the French "*département*" of Sarthe. The population is both gathered in the villages and scattered throughout the countryside, made up of middle and large agricultural lots.

The railway from Tours to Le Mans was built in 1855. Its construction induced an upheaval of the cadastre in Vaas and major shifts, or utter disappearance, of paths and roads in all the affected areas. The building of the A28 highway linking Le Mans to Tours east of Vaas similarly caused a reshuffling of most of Vaas's lots. By and large, it should be noted that Vaas's territory has been significantly more altered than Aubigné-Racan's.

The cadastral data under consideration are old maps (1813, 1850 and 1972 - 1974), which have been scanned and are available at the local archives ("*archives départementales*"). They have been vectorised by ESGT's students. We take as reference dataset the 2012 digital cadastral map ("*Plan Cadastral Informatisé*") of the French land taxation bureau.

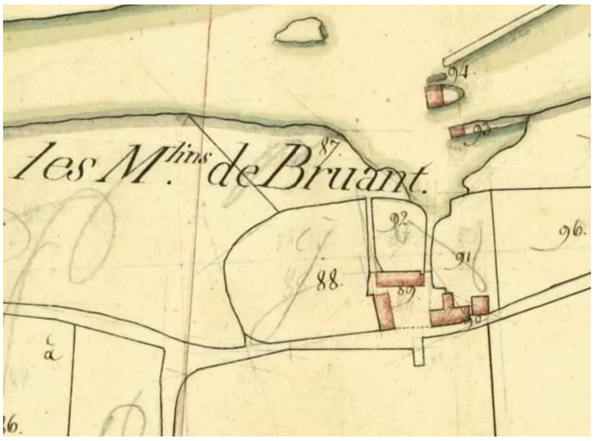

Illustration 1: Excerpt of cadastral map of Vaas village (1813)

The 1813 maps were among the first to be made after Napoleon ordered France's general survey (known as Napoleonic cadastre); survey quality is thus rather poor. The map shows parcels' numbers, paths, roads, buildings, names of places and villages (toponymy).

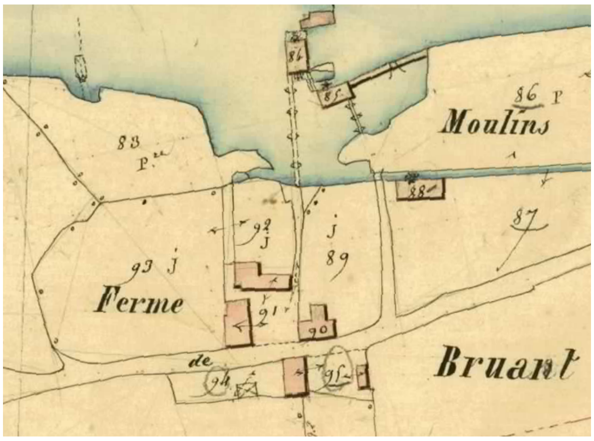

Illustration 2: Excerpt of cadastral map of Vaas village (1850)

The geometric quality of 1850 maps has been greatly improved, thanks to the use of theodolites which allow for faster and more accurate measures. Features are drawn in a completely different way than on the 1813's edition. Furthermore, several types of building ("house" and "sheds") are identified and the use of lots is indicated (vineyard, meadows and gardens). Cellars outlines and party walls or hedges are also included.

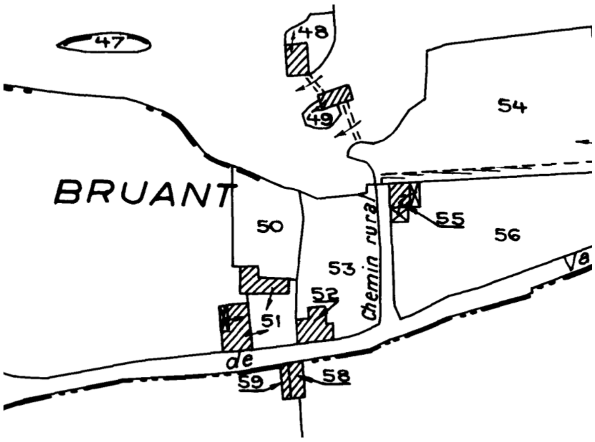

Illustration 3: Excerpt of cadastral map of Vaas village (1974)

1972 - 1974 black and white maps are, by and large, mere updates of the previous versions, except on the centre of Vaas where they were redrawn and have better geometric quality. Different types of buildings are still represented but there is less information than in 1850.

Table 1 summarises the features of old cadastral maps' scanned images.

[redacted]	**1813**		**1850**		**1972 - 1974**	
	200 dpi		200 dpi		400 dpi	
	1:2500	1:1250	1:2000	1:500	1:2000	1:1000
		1:2500		1:1000		1:2000
				1:2000		
	0,3175	0,1588	0,254	0,0635	0,127	0,0635
		0,3175		0,127		0,127
				0,254		
	3	1	6	4	6	10
		7		5		10
				10		

Table 1: Features of old cadastral maps
(medium / light gray cells are related to Aubigné-Racan / Vaas)

Methodology

Building an historical database from scanned maps can be done in two different ways (illustration 4). We can first vectorise, then georeference and finally tile the charts (option 1). In this method, tiling leads to topological error correction of vector data. We can also first georeference and tile images, and only after vectorise them (option 2). Vectorisation, in that second method – the usual one (Baily, B., 2007) — can hardly be automated. Parcels' boundaries can be blurred in overlapped areas and end up distorted in the resulting tiled image. As part of our goal is automated vectorisation, we have chosen the first option. We focus here on the georeferencing step.

Map referencing generally uses global or local transformations (Zitova and Flusser, 2003 ; Bakkouch et al., 2015). Global methods are appropriate when map distortions are homogeneous (only one transformation is computed and applied to the whole chart using all available ground control points). But when local distortions have to be taken into account, a local approach is more efficient: after partitioning the map, a single transformation model is computed for each region. The local and locally sensitive global methods are known to be more efficient to register historical maps (Herrault et al., 2013). We have compared the result of seven transformations on every old cadastral map:

- Global: first, second and third order polynomial (P-ord1, P-ord2 and P-ord3), Helmert transform (H) and Thin Plate Spline (TPS);
- Local: Delaunay triangulation combined with first order polynomial function (TD);
- Locally sensitive global: gaussian kernel regression model (RNG).

GDAL companion tool ogr2ogr perform polynomials and TPS transformations. Since we didn't find any open source tool implementing local or locally sensitive global methods, we developed a Python script using various modules (numpy, shapely, pyshp, scipy and matplotlib). In the case of the RNG implementation, numpy is used among other things to determine the kernel parameters α (amplitude coefficient) and λ (the bias) which minimise the round mean square error (RMSE). This algorithm is detailed in (Herrault et al., 2013).

Thirty ground control points (GCP) samples have been manually selected on each map (1813, 1850 and 1972 - 1974) by comparison with the reference data (2010) using QGIS's georeferencing tool. We chose a random distribution because it has been proven to lead to more accurate results than a regular distribution for a high number of GCPs (Mather, 1995 and Herrault et al., 2013).

The best geometric transformation has been found after appraising the results with the “leave one-out cross validation” statistical method. Let us consider n ground control points:

– n-1 GCPs are used for computing the transformation model,

– the nth GCP is used for computing a residual.

– Each transformation accuracy is appraised computing the RMSE and associated standard deviation on the GCPs residuals.

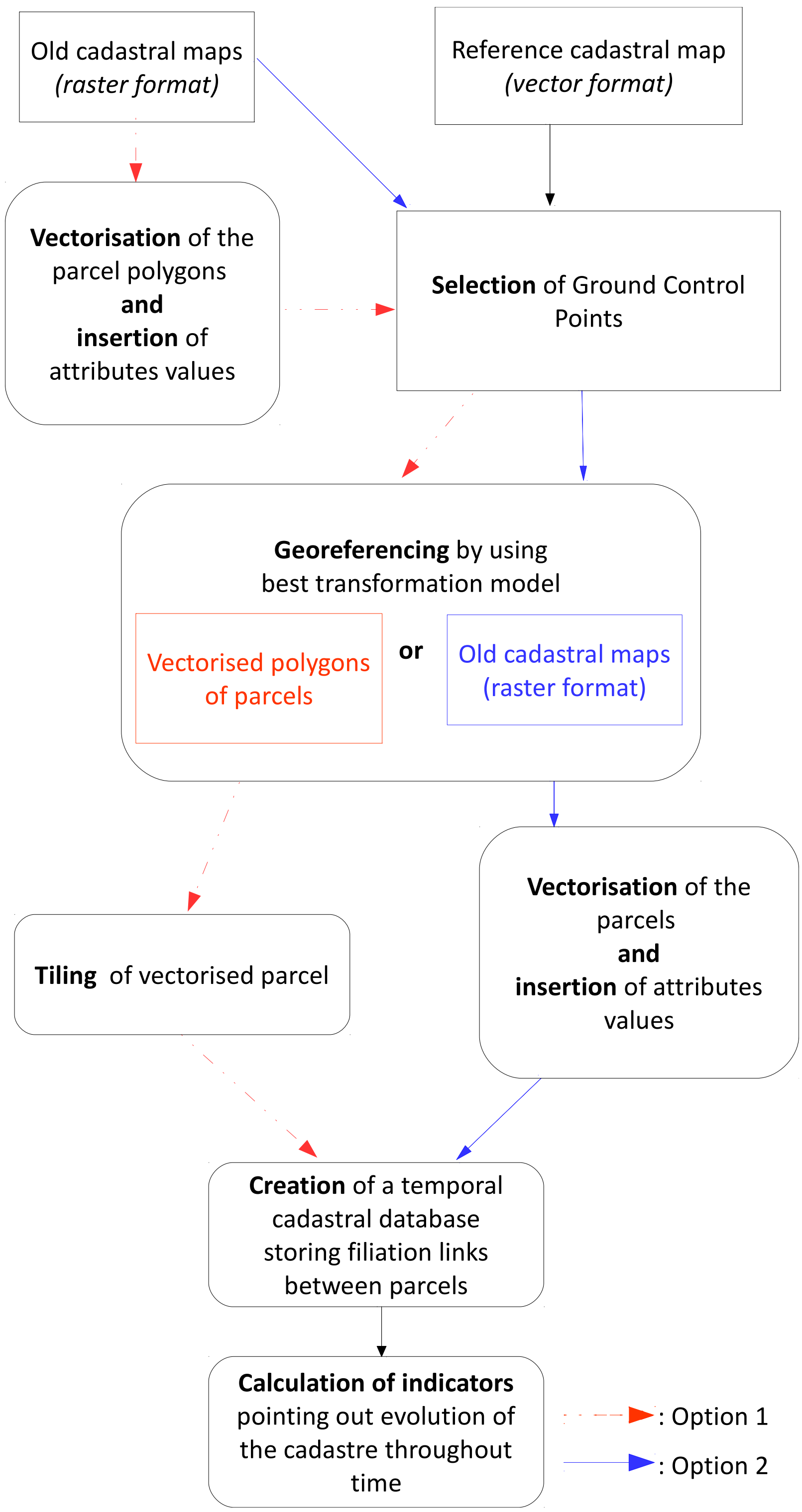

Illustration 4: Possible strategies for building cadastral spatio-temporal database

RESULTS

Illustrations 5 to 10 show the results obtained on the map set covering both villages: for instance, illustration 9 synthesises the value of the GCPs residuals for the nineteen 1850 maps of Vaas. As expected, results on Aubigné-Racan are always better and the accuracy of georeferencing increases when more recent data is processed.

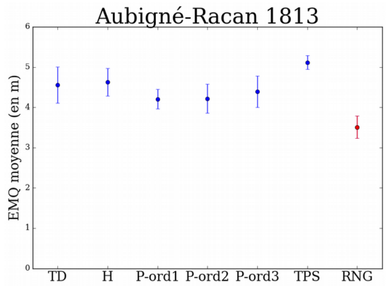

Illustration 5: Results on Aubigné-Racan 1813 data

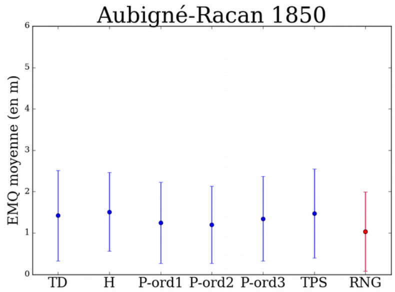

Illustration 6: Results on Aubigné-Racan 1850 data

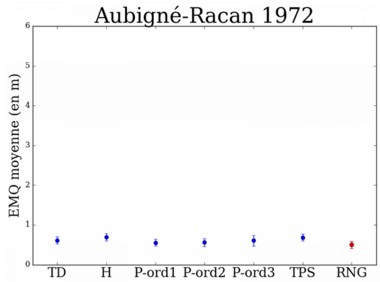

Illustration 7: Results on Aubigné-Racan 1972 data

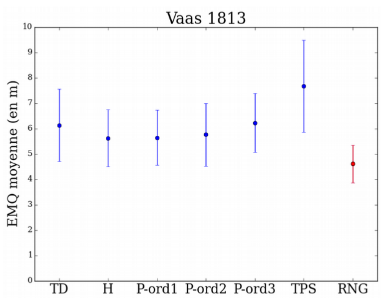

Illustration 8: Results on Vaas 1813 data

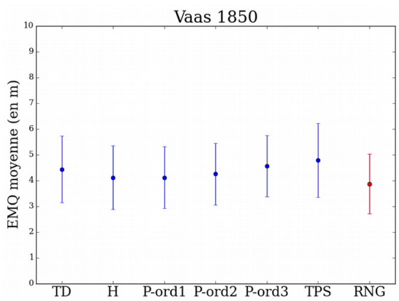

Illustration 9: Results on Vaas 1850 data

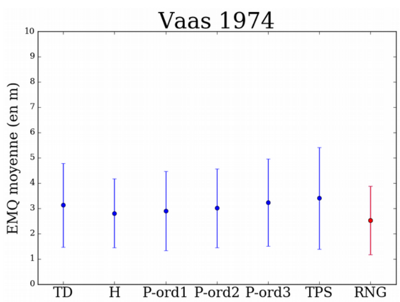

Illustration 10: Results on Vaas 1974 data

Previous illustrations show that the RNG always performs best: it's especially true for oldest maps and when parcels' layout has been altered significantly over time (case of Vaas between 1850 and 2010). In the other cases, the gaps between local transformations, P-ord1, P-ord2 and H, and RNG transformations are less important.

Detailed results of RNG transformation are given in table 2 and illustration 11 overlays reference data with georeferenced 1813 data of Aubigné-Racan obtained with RNG and TD.

	Village	RMSE in meters	Standard deviation
1813	Aubigné-Racan	3,51	0,28
	Vaas	4,71	1,01
1850	Aubigné-Racan	1,04	0,96
	Vaas	3,32	0,90
1972 - 1974	Aubigné-Racan	0,50	0,08
	Vaas	2,52	1,36

Table 2: Results on RNG transformation models

Illustration 11: Comparison of maps georeferenced with different transformations

CONCLUSION

This paper presented an open source based automated tool providing guidance for the choice of the best transformation to use on old cadastral maps. It will be included in a global methodological toolchain whose key elements have been described. Further tests should be conducted on cadastral maps representing different landscapes (hilly areas, for example). Given the wealth of information contained in old cadastral maps, the resulting database should be a powerful tool for historical analyses of territories and landscapes.

ACKNOWLEDGMENT

Thanks to Pierre-Alexis Herrault (for having provided us with the Matlab code of his application), Emmanuel Labrousse, Pierre Lambert, Vincent Thouzeau and Sébastien Coicaud, former ESGT students who started this project and created the vector data, Vincent Habchi (for English help), and Marie Fournier and Mathieu Bonnefond (for all the discussions on the old cadastre).

REFERENCES

Bakkouch L., Bihr H., Ruggeri P., Follin J-M., Simonetto E., Touzard P., 2015. Comparaison de méthodes de recalage d'images pour le géoréférencement de clichés anciens. Revue XYZ n° 142, 15-23.

Baily B., 2007. The extraction of digital vector data from historic land use maps of Great Britain using image processing techniques, E-perimetron, vol. 2, n°4, 209-223.

Bitelli G., Gatta G., 2011. Experiences on georeferencing of maps from the XIX century Gregorian Cadastre of Bologna (Italy), E-perimetron, vol. 6, n°4, 270-275.

Clergeot P., 2007. Cent millions de parcelles en France : 1807, un cadastre pour l'Empire. Publi-Topex, Paris, 124 p.

Boutoura C., Livieratos E., 2006. Some fundamentals for the study of early maps by compative methods. E-Perimetron, vol.1, n° 1, 60-70.

Grosso E., Plumejeaud C., Parent B., 2012. GeoPeuple project: using RESTful Web API to disseminate geohistorical database as open data, 2nd Open Source Geospatial Research and Education Symposium (OGRS)

Dupouey J-L., Bachacou J., Cosserat R., Aberdam S., Vallauri D., Chappart G., Corvisier de Villièle M-A., 2007. Vers la réalisation d'une carte géoréférencée des forêts anciennes de France. Revue du Comité français de cartographie, n°191, 85-98.

Hedefalk, F., Harrie, L., Svensson, P., 2015. Methods to Create a Longitudinal Integrated Demographic and Geographic Database on the Micro-Level. Historical Methods, vol. 48, n°3, 153-174.

Hennet A.-J.-U., 1811. Recueil méthodique des lois, décrets, réglemens, instructions et décisions sur le cadastre de la France ; approuvé par le Ministre des finances. Paris, Imprimerie Impériale.

Herrault P-A., Sheeren D., Fauvel M., Monteil C., Paegelow M., 2013. A comparative study of geometric transformation models for the historical 'Map of France' registration. Geographia Technica, n° 1, 34-46.

Mather P.M., 1995. Map-image registration accuracy using least-squares polynomials. International Journal of Geographical Information Science, n° 9, 543-554.

Noizet H., Grosso E., 2012. Mesurer la ville : Paris de l'actuel au Moyen Âge. Les apports du système d'information géographique d'ALPAGE. Revue du Comité français de cartographie, n°211, 85-100.

Shimizu E, Fuse T, 2003, Rubber-sheeting of historical maps in GIS and its application to landscape visualization of old-time cities: focusing on Tokyo of the past, in Proceedings of the 8th international conference on computers in urban planning and urban management, 11(3)

Zitova B., Flusser J., 2003. Image registration methods: a survey, Image and Vision Computing, 21, 977-1000.