e-Perimetron, Vol. 9, No. 2, 2014 [51-65] www.e-perimetron.org | ISSN 1790-3769

Benoît Costes*

Matching Old Hydrographic Vector Data from Cassini's Maps

Keywords: data matching; hierarchical matching; strokes; hydrographic network; Cassini; old maps; imperfection

Summary: This article focuses on the process of matching old hydrographic networks extracted from Cassini maps.The inherent specificities of such data make conventional methods for matching networks inefficient. We propose a novel hierarchical approach to match hydrographic networks, by extracting natural structures called strokes. Each stroke is then classified according to its weight in terms of hydrographic structure within the graph. The hierarchical selection of candidates and the use of a multi-criteria decision making process (AHP) constitute a good strategy to match imperfect hydrographic networks.

Introduction

Old maps are the legacy of the topographic evolution of the French landscape, the densification of its territory, the evolution of its population distribution, that are the consequences of demographic growth, industrialization and technological improvements. Through the study of Cassini maps and Etat-Major maps (XVIIIth and XIXth centuries), combined with a reference vector topographic database of the French national mapping agency (BDCarto), the research project GeoPeuple proposes a geo-historical and spatio-temporal analysis of the relations between the evolution of the topography of the French territory and its population distribution. This project explores different disciplines and unites a number of research partners: the French mapping agency (cartography and geomatics), LaDeHis (history and demography) and LIP6 (informatics and pattern recognition).

The spatio-temporal analysis to be provided requires the integration of vectorized data extracted from cartographic databases, and thus the matching of homologous entities at different times, taking into account the imperfections of the information provided by the data, the various scales and levels of detail, as well as the goals of the cartographic representations (military or public target, etc.).

The multi-temporal data-matching has already been achieved in the case of punctual entities extracted from Cassini maps (Costes et al. 2012): objects with religious purposes (churches, chapels, etc.), urban vocation (places, center of communes, etc.) or even industrial purposes (mines and mills).

This article proposes a novel approach to match hydrographic networks vectorized from Cassini maps, taking into account the various inherent imperfections of data: significant geometric discrepancies between networks, varying levels of detail, inaccurate location of Cassini vector data, especially in steeply sloping areas, and significant semantic and toponymic incompleteness.

At first, we detail the data used and their imperfections. In a second step, we describe the novel network-matching process implemented. Finally, we provide an evaluation of our methodology and a discussion about the results obtained.

* PhD Student, Université Paris-Est, IGN/SR, COGIT, 73 avenue de Paris, 94160 saint-Mandé [benoit.costes@ign.fr]

Matching Cassini's hydrographic network

Vectorized hydrographic network

During the first stage of the Geopeuple project, three Cassini maps have been georeferenced and vectorized, each one covering different landscapes (Geopeuple 2012): Saint-Malo (sea coast), Reims (lowland area) and Grenoble (mountain area). To compare such extracted entities with current topographic information, we use a reference vector database of the French mapping agency (BDCarto) (Fig. 1).

The inherent imperfection of geographical data (Goodchild 1995, Hunter 1998) is especially pronounced for data extracted from Cassini maps. Indeed, the georeferencing study has highlighted strong inaccuracies characterized by significant and random discrepancies between past and present networks (Geopeuple 2012, Costes et al., 2012), more particuraly marked in steeply sloping areas. Moreover, a lot of uncertainties such as rivers passing north of villages in the past whereas they go around it nowadays, but also relative positioning of localities some with regard to the others not consistent with their current location (Fig. 2), decrease the confidence one can have about the reliability of old data. Last but not least, they suffer from a significant semantic and toponymic incompleteness (most of watercourses are not named).

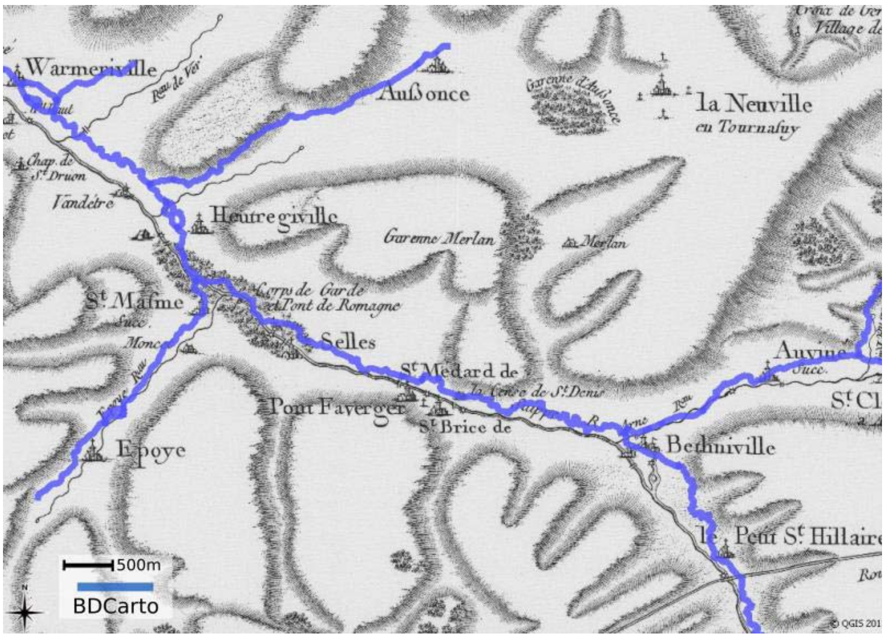

Figure 1: Nowadays hydrographic network superimposed on old Cassini map.

Those imperfections are directly attributable to data entry, as well as to the maps themselves. Indeed, the location of the different objects during the mapping process by triangulation techniques was approximate. The routes of watercourses could have been even modified by the cartographers, in order not to disturb the drawing of others entities such as religious information that was considered the highest priority. All those imperfections make conventional network matching methods inefficient.

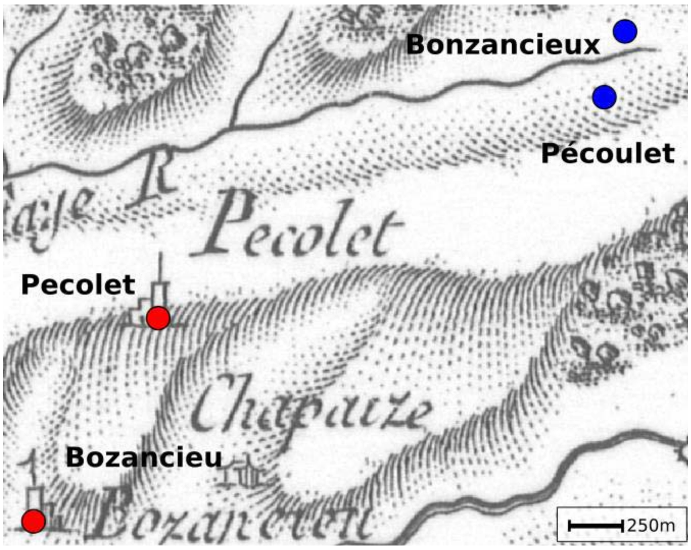

Figure 2: Topological contradictions between relative positions of localities in the past (red) and today (blue).

Existing approaches of matching network

To match networks depicting the same reality at different time, that is to say in order to find out homologous entities in various datasets, one needs to properly define in what extent two polylines are similar. Many criteria can be used, such as topological, geometrical or semantic comparison, and each one of them can be measured by different distances. Developing a matching process involves making a choice of what criteria will be used, and how they will be combined. Devogele (1997) matches networks with different level of details. The matching of edges relies on a pre-matching upstream of nodes, based on criteria of geometric proximity, topological similarity and shortest paths. Mustiere (2006) and Mustiere & Devogele (2008) test and improve this approach on databases with various resolutions. (1:25000 and 1:50000). A similar method is proposed by Lusher et al (2007) to match data with significantly different scales (1:2500 and 1:200000), based on the combination of topological and angular criteria for the pre-matching of nodes.

Walter and Fritsch (1999) use a statistical approach to match networks with relatively similar levels of detail. Candidates for matching are selected by a growing buffer technique, then a set of indicators are calculated for each one: length, baseline orientation, etc. The decision is made by a local optimization algorithm. In another approach, Zhang et al (2005) also use a growing-buffer-based method with an auto-adaptive radius.

All these approaches essentially rely on geometric and topological criteria. Thus, they are inefficient to match networks with as much imperfections as the hydrographic network extracted from Cassini maps. Figure 3 shows an example of strong discrepancy between the two networks that make difficult their matching with such methods.

To our knowledge, only a few network matching processes are adapted to imperfect data. The inaccuracies of data can be managed by the development of a multi-criteria method based on cost functions (Costes et al. 2012) or also evidence theory (Olteanu 2008), but such techniques are limited when all the criteria suffer from imperfections. Furthermore, the algorithmic complexity of the approach used by Olteanu (2008) remains a major concern. Finally, such methods generally need additional knowledge to properly define the parameters of the process, as it is the case in Costes et al. (2012) and in Walter and Fritsch (1999). A comparative evaluation of these methods will be given further.

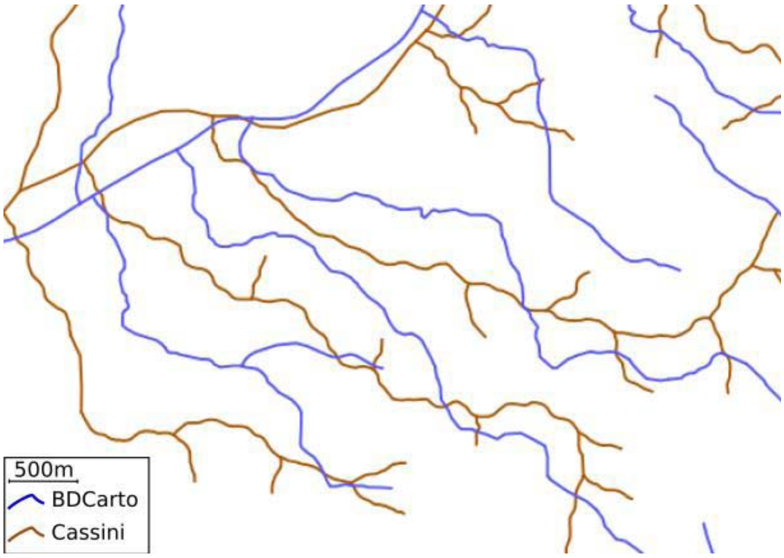

Figure 3: Topological and geometrical discrepancies can make traditional matching approaches inefficient.

Hierarchical network matching

We propose a novel approach to match hydrographic networks adapted to old data, based on a hierarchical multi-criteria technique, by highlighting the main watercourses called strokes, classified according to their weight in term of hydrographic structure. The classification proposed takes into account the branches of the networks. This approach consists of three main steps. Firstly, we make the two networks more structurally comparable by changing how they are represent-ed. Then we look for couples of potential candidates for matching using a hierarchical selection that respects the organization of the networks. Finally, we use a multi-criteria decision making to choose the best couples of candidates (Fig. 4).

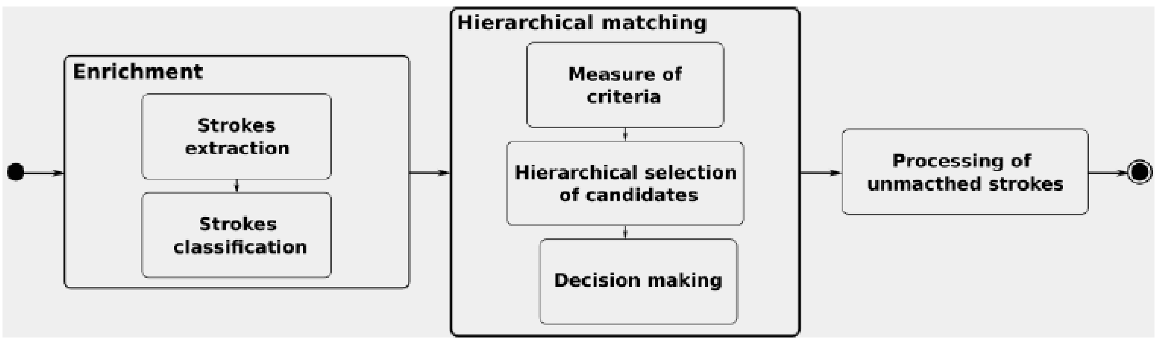

Figure 4: Global scheme of the hierarchical multi-criteria approach.

Making networks more comparable

To make the two networks more structurally comparable, we choose to switch their mode of representation by extracting natural entities called strokes (Thomson and Richardson 1999), or natural roads (Jiang et al. 2008), based on the continuity principle of Gestalt. These strokes consist of a sequence of individual segments that represent continuous hydrographic objects such as rivers and streams.

Jiang et al. (2008) propose an algorithm called « every best fits » to build natural roads that rely on angular criterion at junction point. Although the algorithm fits well for building natural structures with good visual continuity, it does not take into account the topological imperfection of networks (Fig. 5).
To overcome this issue we choose to use an approach similar to that of Jiang et al. (2008), while also taking into account semantic and toponymic information. Strokes thereby represent at best the rivers, streams, brooks, etc.

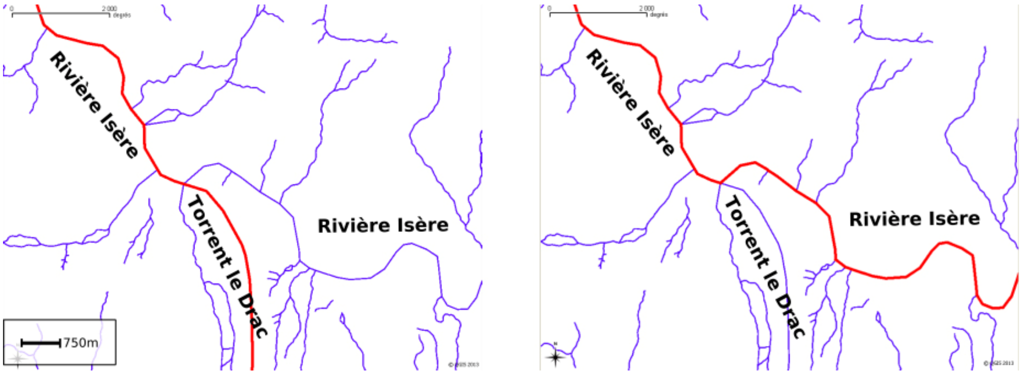

Figure 5: Strokes extraction according to Jiang et al. (2008) on the left, and our method on the right.

A classification of Horton (Horton 1945) on a hydrographic network consist in assigning a number to each stroke, his order, that is a function of the number of its tributaries : «any watercourse without tributary is of order 1, any watercourse with a tributary of order n is of order n+1». Thomson and Brooks (2000) use this classification in order to generalize hydrographic networks. Nevertheless, the classification of Horton strongly depends of the level of detail of the networks. Indeed, as the classification starts on the leafs of the graphs, the main stroke (that is to say the stroke without any tributaries) of respectively a highly detailed network and a network with a lower granularity might have different order. To avoid this situation and make sure that the main strokes of various network will have the same order, we amend the classification of Horton by initializing the algorithm not on the leafs but on the roots of the graphs. Thus, the order 1 is assigned to the main strokes without any tributaries. The order of a stroke reflects its weight (importance) in term of hydrographic structure within the graph. The classification obtained remains however dependant on the selected area for the study. Therefore we need to make sure that the selection of the networks does not induce artifacts in the classification made. For the rest of the article, the parent of a stroke of order n will be the stroke of order n-1 which has the stroke of order n as tributary.

Hierarchical selection of matching candidates

The matching algorithm does not work on the input networks, but uses the strokes as a new modelization of the hydrographic networks. For a given step of the process, let *sref* be the stroke of the reference network to be matched. The *n* candidates $\{c_1,\ldots,c_n\}$ for matching with stroke *sref* are selected as follows: for $i \in \{1,\ldots,n\}$, either *sref* is matched with c_i's parent, or *sref*'s parent is matched with c_i, or *sref*'s parent is matched with c_i's parent (Fig. 6). Thus, the selection al-gorithm depends on the hierarchical structure of the network together with previous matched couples.

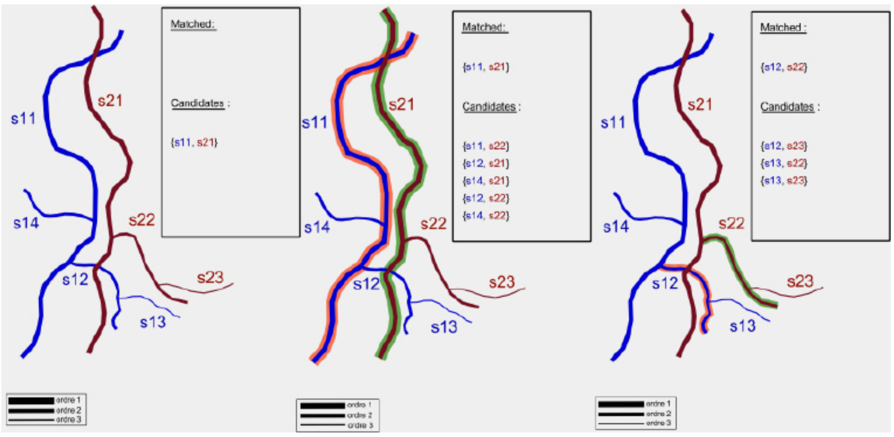

Figure 6: Hierarchical selection and matching of strokes candidates.

Multi-criteria decision making

Each stroke potentially has several candidates for matching. We choose the best one through the use of a multi-criteria analysis technique, the Analytic Hierarchy Process (AHP) implemented by Thomas L. Saaty (Saaty 2008). The AHP is a powerful tool for making complex multi-criteria decisions and is commonly used in fields such as business, industry, etc. It is based on a hierarchical analysis of various criteria and candidates in order to reach a given goal (Fig. 7).

In our context, the objective to achieve is to extract the best candidates from the set of potential candidates for matching if it exists. In that purpose we use 5 criteria. Four of them are quite classical for matching networks and can be used whatever the data to be matched. The fifth is specific to our study. The theoretical principle of the AHP will be illustrated in our matching context.

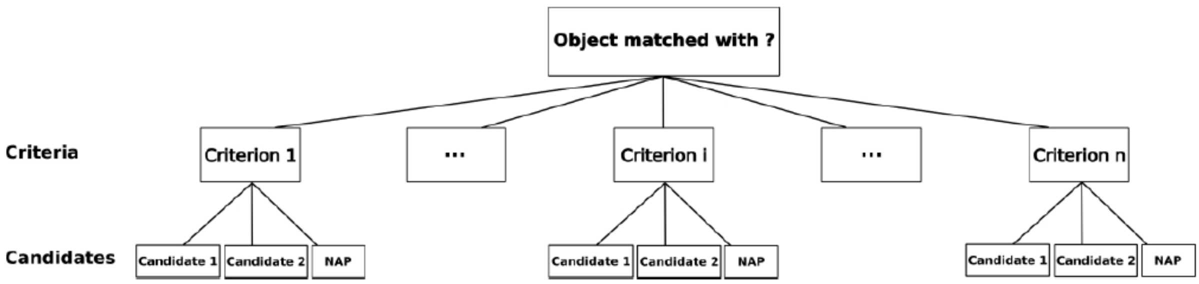

Figure 7: The principle of an AHP in the context of data matching.

Criteria used

We measure the distance between two strokes by the calculus of their discrete Frechet distance, which is an approximation of the Frechet distance calculated in a polynomial time (Eiter and Mannila 1994) (Fig. 8). This distance quantifies not only the space between two polylines, but also their shape difference.

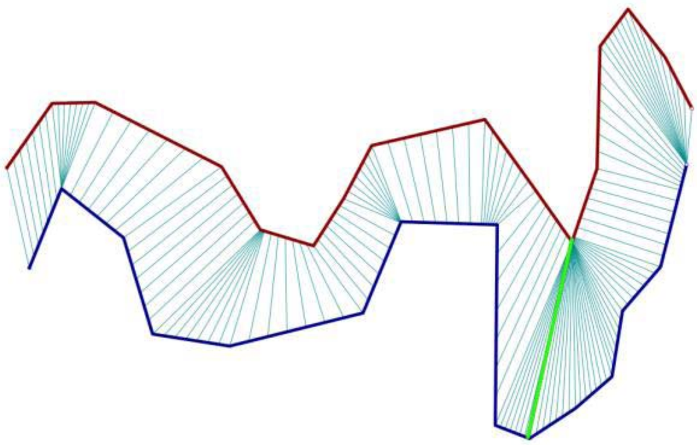

Figure 8: Frechet distance between two polylines.

The difference of orientation between two strokes is given by the angle between the main directions of the strokes. The main direction of a polyline is defined from the contributions of the orientation of each segment that make the polyline to the global orientation (Hangouët 1998). The longer a segment is, the more its orientation contributes to the main orientation (Fig. 9).

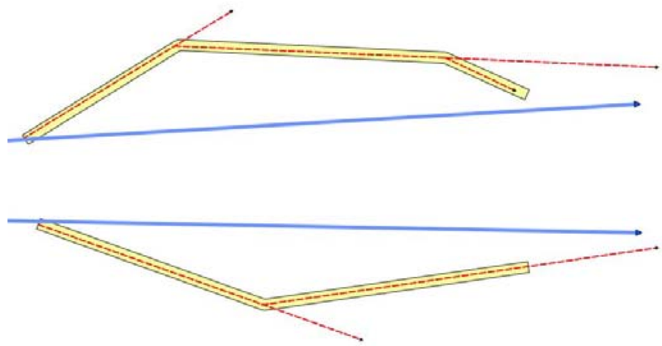

Figure 9: Global orientation of polylines.

The use of only a geometric distance as a geometric criterion is not enough, because two broadly similar geometries can have a significant Frechet distance while the strokes follow rather well overall, if the lines deviate at one end point. This criterion ensures that two geometries do not deviate too far from each other globally. It is based on the measure of the inclusion rate of a polyline to control into the epsilon band (Goodchild 1997) (buffer of ratio epsilon) of the reference polyline (Bel Hadj Ali 2001) (Fig. 10).

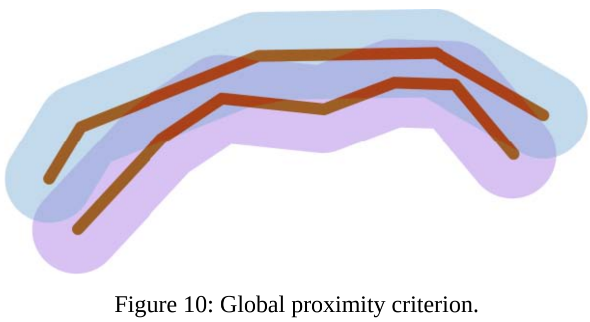

Figure 10: Global proximity criterion.

We use a toponymic distance that is an improvement of the toponymic distance implemented by Samal et al. (2004), based on the Damarau-Levenshtein distance (Damarau 1964), an attractive expansion of the Levenshtein distance (Levensthein 1965) because it take into account more than 80% of spelling mistakes.

We introduce a new criterion that is specific to our study of matching old hydrographic networks. It relies on the assumption that most of the relationships between the hydrographic network and points of interest have been preserved until now. Indeed, in the context of the Geopeuple project, we focus on the major watercourses that serve localities and are potentially involved in the changes affecting them. We suppose that if a river passed close to a village in the past, this remains true today. If we can identify the homologous localities, villages, mills, etc. close to a watercourse in each data base, we might assume that the corresponding watercourses are also matched.

Thus, this criterion requires an upstream matching of punctual objects extracted from Cassini maps, which has already been achieved (Costes et al. 2012): localities, mills, villages, etc. have been matched with another reference topographic vector database of the french mapping agency (BDTopo). For the rest of the article we assume that the matching links established in Costes et al. (2012) have been validated by manual post-processing. An illustration is given in Figure 11. We note that the counterparts of Cassini entities close to a river of the 18th century are close to the existing watercourse, considering that the concept of proximity is defined by a buffer.

The criterion is measured by calculating a ratio of matched punctual objects close to strokes candidates.

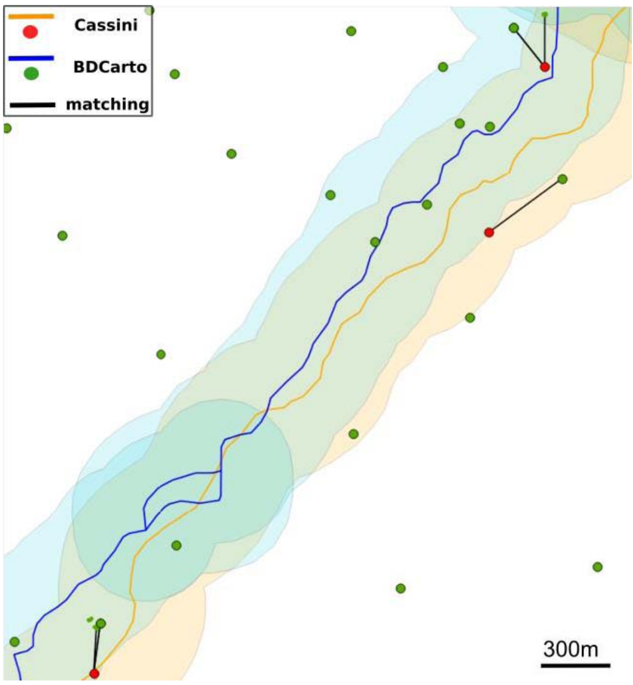

Figure 11: introduction of a new criterion based on the distance to already matched punctual entities.

Candidates' regroupment

Given a stoke to match, all its candidates for matching are selected and compared to each other. For the following, an alternative will be the choice of a candidate among the other. We also create a fictive alternative, denoted NM (not matched): "*the stroke is not matched*", used by the analysis process that will be the best choice if none of the candidates is especially noted.

Confrontation of candidates

For each criterion, the scores of the candidates face pairwise. When a candidate of score m_i faces the alternative NM, we give the latter the score $1-m_i$, in order to illustrate to what extent we do not believe that this candidates is the good match compared to its score.
An artificial decision maker establishes each alternative's preference (the Fundamental Scale for Pairwise Comparisons) (Saaty 2008) (Table 1). It indicates the degree of importance of one alternative against the other according to the score of each one for the measured criterion. Let be the importance of candidate i against candidate j. Then $a_{ji} = \frac{1}{a_{ij}}$.

Intensity of Importance	*Definition*	*Explanation*
1	Equal Importance	Two activities contribute equally to the objective
2	Weak or slight	
3	Moderate importance	Experience and judgement slightly favour one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgement strongly favour one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favoured very strongly over another; its dominance demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favouring one activity over another is of the highest possible order of affirmation
Reciprocals of above	If activity *i* has one of the above non-zero numbers assigned to it when compared with activity *j*, then *j* has the reciprocal value when compared with *i*	A reasonable assumption
1.1–1.9	If the activities are very close	May be difficult to assign the best value but when compared with other contrasting activities the size of the small numbers would not be too noticeable, yet they can still indicate the relative importance of the activities.

Table 1: Example of decision maker based on the principle of *the Fundamental Scale for Pairwise Comparisons (Saaty, 2008).*

The result of the comparison is then recorded in a matrix of size $(n+1)*(n+1)$, where n is the number of confronted candidates. Next, a priorities vector is calculated reflecting the relative importance of each candidate for the measured criterion. This vector actually is the major eigenvector of the matrix. The AHP is particularly interesting because it provides a consistency ratio (*CR*) that allows an assessment of the calculations. In other words, the *CR* indicator indicates in what extent the values established by the decision maker are consistent one with each other. In practice, we try to ensure that $CR<0.10$ (Saaty 2008). For instance in Table 2, the candidate 1 prevails logically for the given criterion.

Scores (between 0 et 1)	Comparisson matrix				Priorities
		Candidate 1	Candidate 2	NM	
Candidate 1 : 0.7	Candidate 1	1	3	4	0.63
Candidate 2 : 0.4	Candidate 2	1/3	1	1/2	0.15
	NM	1/4	2	1	0.22
					CR=0.09

Table 2: Example of comparison matrix for an arbitrary criterion.

Confrontation of criteria

The AHP allows to weight each criterion based on the importance we grant it in respect to each other criterion. That is to say one can define a weights matrix of size $m*m$, where m is the number of criteria (here $m = 5$), which adds much finesse in the definition of relative importance of each criterion, usually done by experts. From this confrontation matrix, a priorities vector and an inconsistency ratio are calculated as done previously.

Decision making

For each candidate, a global priority is calculated, by combining the priorities of the candidate in respect with each criterion, and taking into account the relative priorities of each criterion one with the others. Let $(C_1, ..., C_N)$ be n candidates, $(\alpha_{i,1}, \dots \alpha_{i,N})$ the priorities vector defined by the confrontation of the set of candidates for the i_{th} criterion, and let $(p_1, ..., p_m)$ be the priorities vector calculated by the confrontation of the m criteria. Then the global priority $P(C_j)$ for candidate C_j is given by: $P(C_j) = \sum_{i=1}^{m} p_i * \alpha_{i,j}$. The candidate with the higher global priority is thus chosen. The object is declared not matched if the alternative NM is chosen.

Post-treatment of not matched strokes.

Unmatched strokes are then treated separately, and for each of them we apply the matching algorithm of Mustiere and Devogele (2008). We restrict the matching to the sub-network consist-ed of the stroke studied and strokes connected to it, in order to avoid introducing side effects, giv-en the topological simplification that is made here. The matching links established by this step are then referred to as uncertain and require manual verification post processing.

Assessment of results

Measurements of the quality of the matching

To assess a matching process, we usually calculate the precision and recall of both the matching links and the unmatched objects, respectively noted as P_{app}, R_{app}, and P_{nap}, R_{nap}.
We traditionally call: true positive (*vrais positifs*) the matching links correctly established by the automatic process (*vp*), false positive (*faux positifs*) the links in excess (*fp*), true negative (*vrais negatives*) the objects correctly unmatched (*vn*) and false negative (*faux negatives*) the objects mistakenly unmatched (*fn*). Let N_{app} be the number of matching links expected and N_{nap} the num-

ber of unmatched objects expected. We also intro-duce a measure that equally combines precision and recall, called *FScore*:

$$\begin{cases} P_{app} = \frac{vp}{vp+fp}, R_{app} = \frac{vp}{N_{app}}, P_{nap} = \frac{vn}{vn+fn}, R_{nap} = \frac{vn}{N_{nap}}, Fscore = 2 \cdot \frac{precision \cdot rappel}{precision + rappel} \end{cases}$$

Results

The table 3 compares the assessments of the results of our method and two classical matching algorithms (Mustiere and Devolege 2008, Costes et al. 2012).

	Matching links			Unmatched objects		
Areas	Precision	Recall	Fscore	Precision	Recall	Fscore
	Approach of Mustière and Devogele (2008)					
Reims	73%	85%	78%	44%	80%	57%
Grenoble	58%	76%	66%	89%	85%	87%
St Malo	63%	91%	74%	86%	71%	77%
	Approach of Costes et al. (2012)					
Reims	90%	87%	89%	57%	80%	67%
Grenoble	67%	82%	74%	97%	85%	90%
St Malo	48%	68%	56%	86%	71%	77%
	Hierarchical multi-criteria Approach					
Reims	90%	96%	93%	57%	80%	67%
Grenoble	90%	72%	81%	90%	97%	93%
St Malo	96%	73%	83%	92%	100%	96%

Table 3: Quantitative results of the assessment of thee data matching approaches.

All three methods own many parameters, and several combinations have been tried so that we can compare the best results obtained by each of the approaches. Figure 12 illustrates that most of the main hydrographic structures have been correctly matched. The assessment shown by Table 3 emphasizes that our approach improves significantly the results performed by more classical algorithms: + 18 % (resp. + 16%) for matching links (resp. for unmatched objects) in average compared with the algorithm of Mustiere and Devogele (2008); + 17 % (resp. + 9%) for matching links (resp. for unmatched objects) in average compared with the approach of Costes et al., (2012). The computation time is also improved in average.

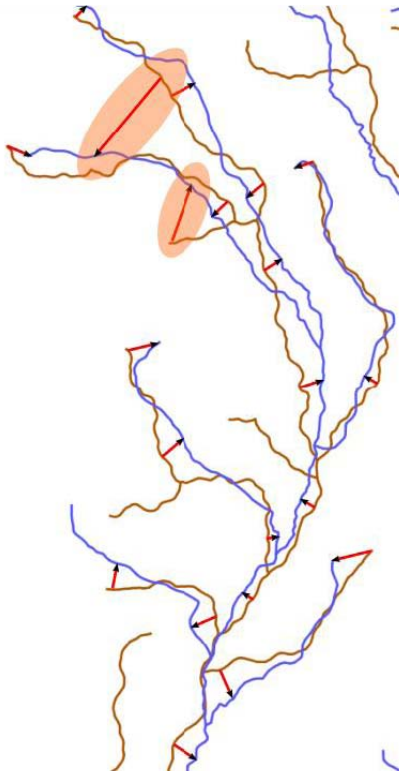

Figure 12: Example of matching for the area of Grenoble (2 mistakes shown in orange).

Discussion

We want to analyze the relative importance of each criterion in our context of matching Cassini hydrographic network. For the study, we consider that all the weights of the criteria matrix are set to one. It is surprising to note (see Table 4) that, except for the Grenoble area, the results are improved when the geometric criterion is not used. We interpret this result as a consequence of the significant geometrical discrepancies between networks. For Reims and Grenoble, orientation, proximity and toponymic criteria seem to be essential. The role of the toponymic criterion depends on the number of nominated Cassini watercourses: as they outnumber on the area of Reims, this criterion is logically more critical for this zone, whereas only less than 5% of St Malo rivers have a name.

The special criterion, measured by proximity to matched punctual entities, appears to be only a little crucial, but still improves the results somewhat. Furthermore, it increases the global score of low order strokes, a fact that can be interpreted as a stronger probability of correct matching.

	F-Score matching links			F-Score unmatched objects		
	Reims	Grenoble	St Malo	Reims	Grenoble	St Malo
All criteria	84%	61%	34%	50%	86%	65%
Without special criterion	83%	59%	31%	40%	85%	61%
Without geometric criterion	86%	60%	44%	63%	83%	66%
Without toponymic criterion	43%	40%	43%	17%	81%	62%
Without orientation criterion	72%	28%	23%	37%	78%	58%
Without proximity criterion	74%	28%	38%	38%	78%	62%

Table 4: Scores obtained by the suppression of some criteria.

Conclusion

In this article we have focused on the matching of an hydrographic network extracted from Cassini maps with a reference topographic vector database of the French mapping agency (BDCarto). The matching of old map is a difficult task because of the inherent imperfections of such data, especially when the data are nearly 250 years apart in age. In our context, these imperfections concern particularly a significant geometric discrepancy between networks, varying levels of detail, inaccurate location of Cassini vector data, especially in steeply sloping areas, significant semantic and toponymic incompleteness and topological inconsistency between homologous entities. These imperfections make conventional network matching methods inefficient. We propose a novel approach of matching hydrographic networks based on the respect of the hierarchical structure of these networks. We firstly switch the representation of the network through an enrichment step, by the detection of natural structures of good continuity called strokes that represent at best the real hydrographic entities such as the rivers, streams, etc. The hierarchical layout so defined thus can be used as a pivot for the matching process, retaining among potential candidates those whose parents or themselves have been previously matched. The hierarchical selection of stroke candidates for matching, combined with the use of an efficient multi-criteria analysis technique (AHP) improve significantly the results obtained by
geometrical-based and topological-based approaches, or non-hierarchical multi-criteria methods. Furthermore, we introduce a new criterion measuring the proximity of each stroke to points of interest. This criterion requires an upstream matching of punctual objects extracted from Cassini maps (places, mills, etc.). The criterion is measured by calculating the ratio of matched punctual objects close to strokes candidates, and relies on the assumption that most of the relationships between the hydrographic network and these points of interest have been preserved until now.
The main limitation of the approach concerns the leaves of a highly branched network due to the hierarchical structure of the process. Nevertheless, the main waterways are properly matched, and their current counterpart clearly identified. The use of the AHP to make a decision raised the issue of its parameterization. A preliminary sensitivity study of the algorithm suggests that the results are improved when the Frechet distance criterion is not used due to the significant discrepancies between networks. Both orientation and proximity criteria seem to be the most important on our tested data, together with the toponymic criterion if enough watercourses are named. The new criterion introduced increases the global score of high order strokes, even if it does not improve that much the results. In the short term, it would be interesting to deepen the sensitivity study of the parameters of the AHP.

Acknowledgment

The author thanks Julien Perret and Eric Grosso for their proofreading and many helpful suggestions.

References

Bel Hadj Ali A. (2001). Qualité géométrique des entités géographiques surfaciques. Application à l'appariement et définition d'une typologie des écarts géométriques. Thèse en Sciences de l'Information Géographique, Université de Marne-La-Vallée

Costes B., Grosso E., Plumejeaud C. (2012). Géoréférencement et appariement des données issues des cartes de Cassini. In: Actes du Colloque International de Géomatique et d'Analyse Spatiale (SAGEO 2012)

Costes B. (2013). Appariement hiérarchique de réseaux hydrographiques imparfait. In : 8th International Workshop on Digital Approaches to Cartographic Heritage, Rome, 19-20 September, 2013.

Damerau F.J.A (1964). Technique for computer detection and correction of spelling errors Commun. ACM, vol. 7, p. 171-176

Devogele T. (1997). Processus d'intégration et d'appariement de bases de données Géographiques, Applications á une base de données routières multi-échelles. Thèse en Sciences de l'Information Géographiques, Université de Versailles

Devogele T., Badard T., Libourel t. (2002). La problématique de la représentation multiple. In: Ruas, A. (éd), Généralisation et représentation multiple. Paris, Lavoisier, 390 p.

Eiter T., Mannila H. (1994). Computing discrete Fréchet distance. Christian Doppler Laboratory for Expert Systems , rapport CD-TR 9464

Geopeuple. ANR Project (2011-2013), www.geopeuple.ign.fr

Geopeuple (2012). L2.1-1 Géoréférencement des feuilles de Cassini. Technical report (16 pages).

Goodchild M.F. (1995). Sharing imperfect data. In H.J. Onsrud and G. Rushton, editors, Sharing Geographic Information, New Brunswick, NJ: Rutgers University Prress, 1995, p. 413-425

Goodchild M.F., Hunter G.A, (1997). A simple positional accuracy measure for linear features. International Journal of Geographical Information Systems, vol. 11, p. 299-306

Hangouët J.F (1998). Approche et méthodes pour l'automatisation de la généralisation cartographique;application en bord de ville. Université de Marne-La-Vallée

Horton, R.A. (1945). Erosional development of Streams and their drainage basins: hydrophisical approach to quantitative morphology. Geological Society of America Bulletin, vol. 56, p. 275-370

Hunter G.J. (1998). Managing uncertainty in GIS. NCGIA Core Curriculum in GISScience, fév. 1998

Jiang B., Zhao S., Yin J. (2008). Self-organized Natural Roads for Predicting Traffic Flow: A Sensitivity Study. Journal of Statistical Mechanics: Theory and Experiment

Levenshtein V. (1965). Binary codes capable of correcting deletions, insertions and reversals. Doklady Akademii Nauk SSSR, vol. 4, p. 845-848

Lüscher P., Burghardt D. Weibel R. (2007). Matching road data of scales with an order of magnitude difference. In : 23th International Cartographic Conference

Mustière S. (2006). Results of experiments on automated matching of network at different scales. ISPRS, vol. 26.

Mustière S., Devogele T. (2008). Matching networks with different levels of detail. GeoInformatica, p. 435-453

Olteanu A.M. (2008). Fusion de connaissances imparfaites pour l'appariement de données géographiques. Proposition d'une approche s'appuyant sur la théorie des fonctions de croyance.Thèse en Sciences de l'Information Géographique, Université Paris-Est

Saaty T. (2008). Decision making with the analytic hierarchy process. International Journal of Services Sciences

Samal A., Seth S., Cueto K.A. (2004). A Feature-based approach to conflation of geospatial sources. International Journal of Geographical Information Sciences, vol. 18, p. 459-489

Thomson R., Richardson D. (1999). The 'Good Continuation' principle of perceptual organization applied to the generalization of road networks. In : 19th International Cartographic Conference

Thomson R., Brooks R. (2000). Efficient generalization and abstraction of network data using perceptual grouping. In : 5th GeoComputation

Walter F., Fritsch D. (1999). Matching spatial data sets : a statistical approach. International Journal of Geographical Information Science, vol. 13:5, p. 445:473

Zhang M., Shi W., Meng, L.A (2005). A generic matching algorithm for line networks of different resolutions. In: ICA Workshop on Generalisation and Multiple Representations