Development of a new generalisation flowline for topographic maps

François Lecordix, Jean-Marc Le Gallic, Loïc Gondol, Arnaud Braun
Institut Géographique National (France),
francois.lecordix@ign.fr, jean-marc.le-gallic@ign.fr, loic.gondol@ign.fr, arnaud.braun@ign.fr

Abstract

In June 2004, the French National Mapping Agency IGN launched an important project, called New Base Map Project, to develop a new flow line to produce topographic maps at 1:25 000 and 1:50 000 scale from the vector geographic database BDTopo®. Even if IGN has produced topographic maps at 1:25 000 scale from BDTopo® since more than 10 years, this project has to introduce a lot of new solutions to create and manage cartographic databases on whole France which allow producing topographic maps. These new solutions and the new flowline will be presented in detail during the International Cartographic Conference at Moscow by Arnaud Braun. This paper in this generalisation workshop will focus on two specific parts of the problem to produce cartographic database and maps: the generalisation process and, to be able to automate generalisation process, the preliminary process to introduce consistency in the data.

Introduction

Twenty years ago, IGN started to produce a vector geographic database BDTopo® with analytical stereophotogrammetric restitution and field collection. This flowline allowed the storage of planimetric and altimetric informations on 25% of the country with a first specification. This specification contained all informations which were on topographic maps and were extracted from only one main source (aeronautical photo) with data consistency. A numerical cartographic flowline was then developed in 1993 to produce numerically topographic maps at 1:25 000 scale from this BDTopo® and 450 sheets has been produced with this solution.

To speed up the creation of the database BDTopo® on whole France, IGN decided in 2000 to simplify the specifications of this database to be able to finish the creation of this database in 2006. This aim was reached in December 2006, but less informations were stored in database BDTopo® and different sources (aeronautical photos, old maps, old DTM …) were used to obtain this new version of BDTopo®.

With less informations in the vector database BDTopo® and without consistency for some themes, it is necessary to redefine a new flow line to be able to continue to produce as similar topographic maps as before. The New Base Map Project was launched to solve first this problem. But other important aims were added to provide a very new production flowline for topographic maps:

- to produce too maps at 1:50 000 scale – and not only at 1:25 000 scale –, generalisation solutions must be provided;
- to decrease maps updating cost, automatic solutions to propagate updates data from geographic database (or DLM) to cartographic database (or DCM) must be developed;
- to replace a data management sheets by sheets, a new data base management system seamless must be proposed.

The description of the complete process (fig. 1) will be presented during the International Cartographic Conference by [Braun, 2007]. To obtain this new flowline, the New Base Map Project has to work on a lot of different subjects: collection, extraction from different sources, data matching, generalization, symbolization, edition, updating, cartographic data base management…In this paper for the workshop on generalisation and multiple representations, we will focus on the problem to generalise data as automatically as possible, particularly for buildings. But, to be able to automate generalisation, it is necessary first to introduce consistency in the data that we will describe in a first part.

Many works on these two subjects use mainly COGIT laboratory results which were already presented in previous workshops on Generalisation and Multiple Representation. But the New Base Map Project adapted these works to apply on a large dataset and so to be able to produce first a complete prototype and then to provide a new flowline for 100 operators in production department at IGN.

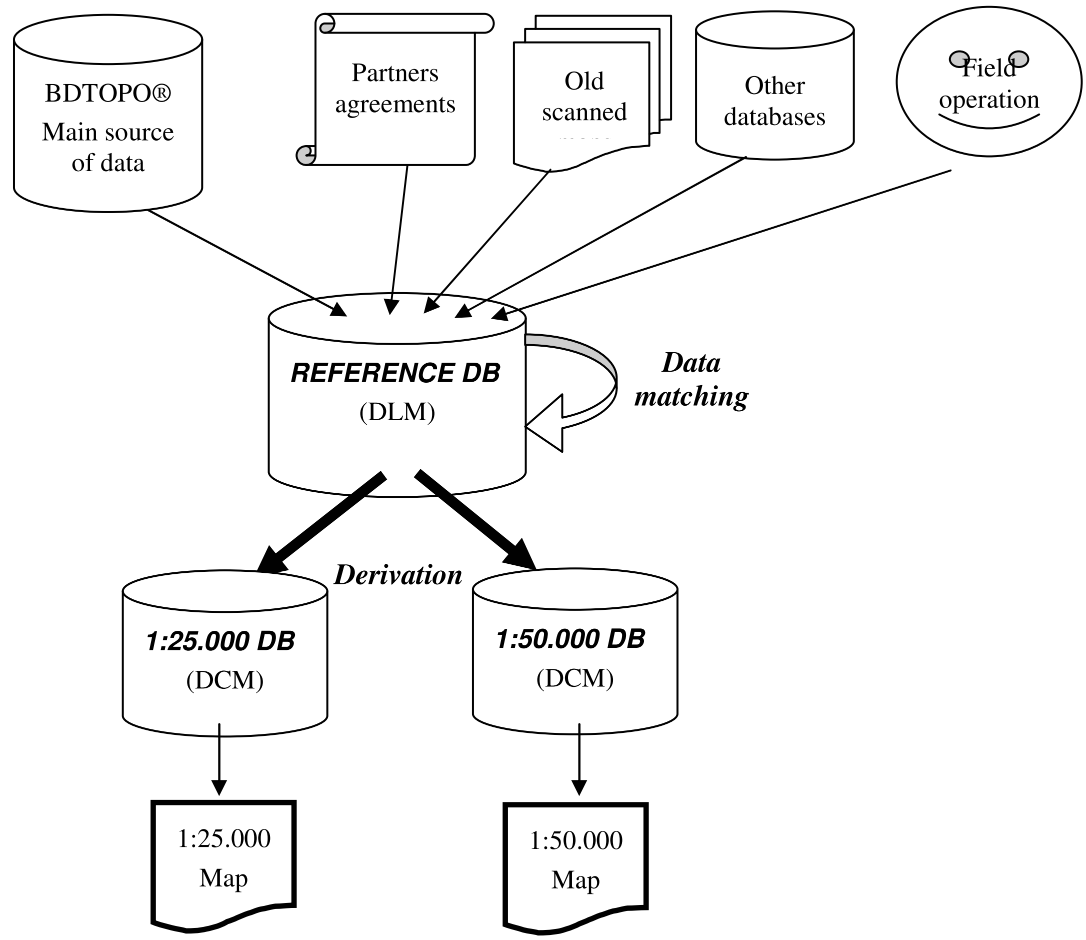


Figure 1: General flowline of the New Base Map

Data Matching

With the new BDTopo®, all informations to produce maps are not present in the database. Even if the data come mainly from BDTopo®, other sources will be used for the maps. These other informations can be provided by partners or by IGN products.

To share the expensive work of data collection, IGN decided to sign agreement with partners who collects information on the whole France or on a large portion of the territory. Thanks to this agreement, it is possible for the New Base Map Project to use data for the new maps production flowline. Two examples can be provided:

- Foot tracks are collected by an association (Fédération Française de Randonnée Pédestre) who digitalised first the information from an old map raster file and now updates the information, but without links with tracks collected by IGN in the vector database BDTopo®.
- Forests are collected by an another French public institute (Inventaire Forestier National) and independently of other informations (roads, rivers,…) in BDTopo®.

When there is no partner and the vector data don't exist, the project can use IGN products to obtain numerical informations. For vineyard, the project developed an automatic process to recognize patterns on a map raster file and to extract areas limits of these patterns (fig. 2). In mountain areas, another automatic process has been developed to extract rocks, screes, glaciers from orthophotos. For contour lines, the project uses a DTM that IGN built twenty years ago with different processes which depended on the area: maps contour lines digitalization for flat areas, analytical photorestitution for mountain areas.

Figure 2: Automatic extraction of areas with patterns (vineyard and forest) from raster file (on the left)to obtain vector areas which are symbolised in green with the other elements coming from BDTopo® (on the right)

Anyway, for all these data which are collected independently of the BDTopo®, there is no chance that the laying out corresponds exactly with BDtopo® one for elements which share geometries, for instance roads and areas limits. During the generalisation process, roads can be displaced and to keep the data consistency, it is necessary first to introduce data matching.

Using solutions of data matching developed by COGIT laboratory, the New Base Map Project introduced these solutions in the development environment used by the project to automate processes, the software Clarity developed by 1Spatial. These solutions were adapted to be applied on large data set (at least 20 km x 30 km, the extent of one sheet at 1:50 000 scale) and on many specific cases, so to be able to produce a first map prototype. To have more technical details on these solutions used by the project, we recommend the reading of the papers written by the COGIT laboratory researchers. We will focus only on the results, in particular for three cases: foot tracks, vineyards and forests.

For foot tracks (from FFRP) and tracks (from BDTopo®), the problem is a data matching between lines. We used the solution proposed by Sebastien Mustiere in the open source platform Géoxygène [Mustière, 2005]. The data matching source – which will be in open source in the near future -is developed in Java and so it was very easy to introduce it in the Clarity environment of the project. Results were very quickly useable for a large data set and 95% of foot tracks are matched automatically with BDTopo® tracks in few minutes. An example is presented in fig. 3.

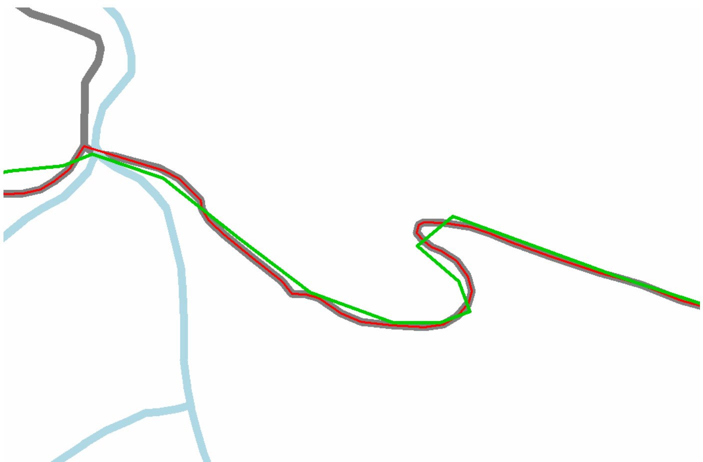


Figure 3: Initial foot tracks (green) are automatically matched with BDTopo® tracks (grey) to provide more accurate foot tracks (red),

For vineyards (from IGN old maps with automatic extraction) to match with road and rivers, the problem is a data matching between areas and lines. After some minor modifications, the same COGIT algorithm is applied on data and provided results presented in figure 4.

Figure 4: Initial vineyards (limit in green) are automatically matched with BDTopo® lines to provide new limits (in red)

For forests (from IFN), the problem is more complex because a lot of specific treatments are necessary. The limits of trees areas are obtained by an automatic extraction from orthophotos. The semantic distinguish between hardwood, conifer wood, brushwood, etc. is collected by IFN operators. But, there are 2 problems:

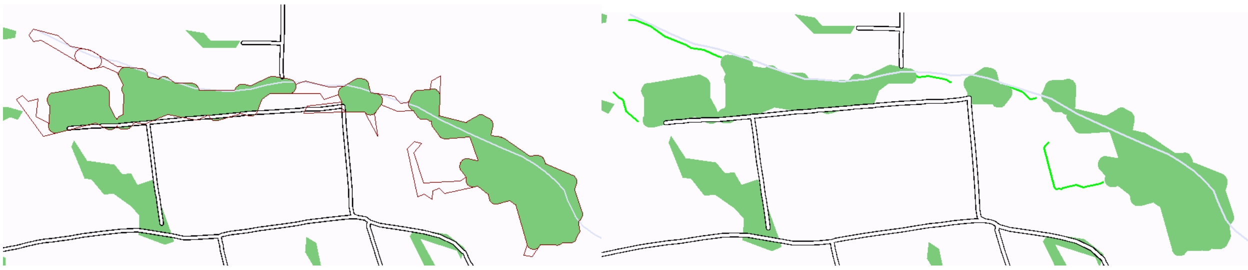
- the partitioning depends on orthophotos one and not on the roads and hydrographic networks;
- hedges are not distinguished and not stored as lines.

A specific study was realized by the COGIT laboratory [Touya, 2007] to solve these problems and a first process was proposed to obtain hedges lines and to cut forests by road and hydrographic networks. Again, the New Base Map Project exploited this solution to be able to apply it on a large dataset. The process is done in 5 steps:

1. Trees areas merging (fig. 5)
2. Cutting up trees areas by networks (fig. 5)
3. Discrimination between forests, hedges and small spot of trees (fig. 6a)
4. Hedges skeleton extraction (fig. 6b) with the straight skeleton [Haunert & Sester, 2004]
5. Forests and hedges matching with networks (fig. 7)

The last step is the same process that we used for vineyards (areas matching) and foot tracks (lines matching). For a large data set (extent 20 km x 30 km) with a lot of forests, the full process takes around 5 hours. The evaluation of full process results will be done in a short time.

Figure 5: Initial data (left) are merging and cutting by networks (right) (Areas limits are in dark green)



(a) Discrimination between forests (green), hedges and small spot of trees (red) *(b)) Hedges skeleton extraction*

Figure 6: Hedges treatment

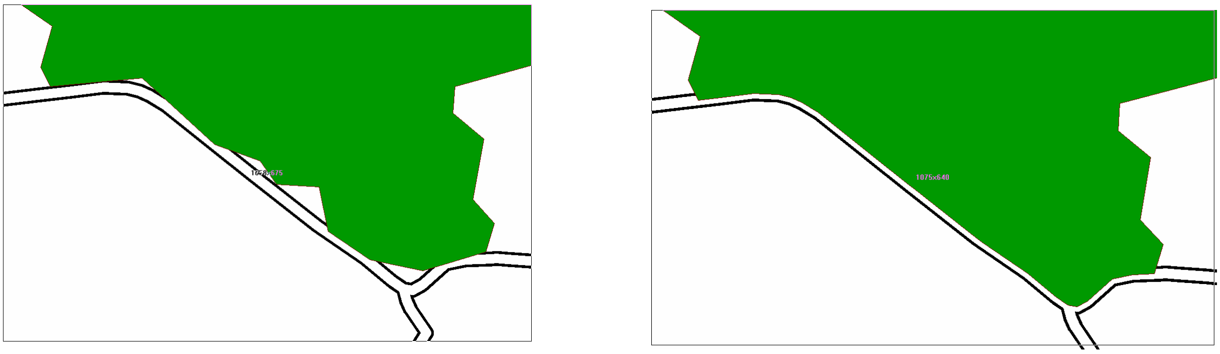


Figure 7: Forests matching with roads network

The last problem of consistency that the New Base Map Project occurs concerns contour lines and hydrographic networks. As these two informations are collected independently with different accuracy, it appears sometimes some inconsistencies: the river intersects many times the same contour line (fig. 8a) or the river is close to the thalweg (fig. 8b). To solve these problems, we experiment solution proposed by [Gaffuri, 2007] with submicro objects. But, at this moment, results are not robust enough for a prototype to present at this workshop.

(a) River intersects many times the same contour line *(b) the river is close to the thalweg*

Figure 8: Example of contour lines and hydrographic network inconsistencies

Generalisation

To produce a map at 1:50 000 scale from the vector geographic database BDTopo®, an important generalisation process is necessary. This generalisation process must be applied more particularly on buildings and networks. A lot of solutions are proposed in research domain but industrial solutions used in NMA are rare at this moment.

The New Base Map Project decided to continue with solutions already used at IGN in the Carto2001 project (Agent technology and Beams algorithm) [Lecordix, 2005] and to extend the use of these solutions to other geographic objects, in particular for buildings. Agent technology [Lamy, 1999] for buildings was already implemented first in LAMPS2 during Agent project [Barrault, 2001] and then in Clarity by COGIT laboratory, but it was not possible to generalise a large data set. So the New Base Map Project provided some efforts to improve this process and the new results are presented in this paper.

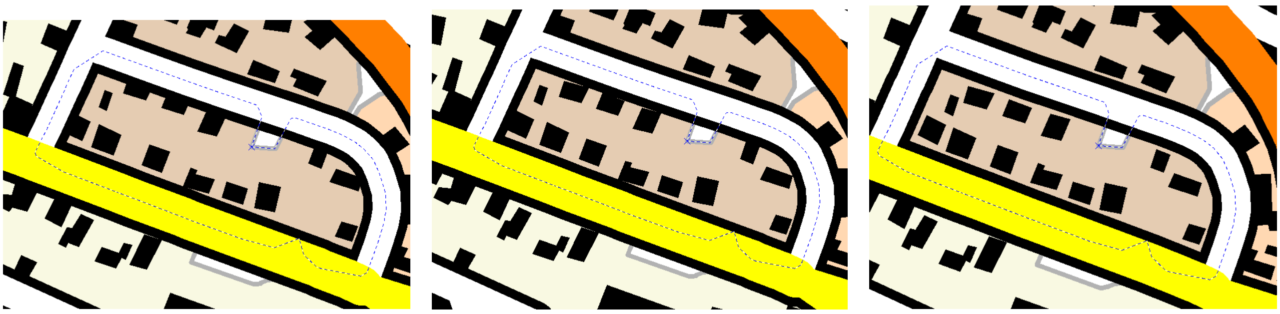
At the last workshop on generalisation and multiple representation, the project presented in details the generalisation process for buildings in the software Clarity [Lecordix, 2006]:

- Urban structures determination [Boffet, 2001]
- Urban structures characterization : unitary blocks, suburban blocks, urban blocks, centre blocks, etc [Boffet, 2001]
- Specific generalisation with Agent technology on each block which is considered as a meso agent (fig. 9)

a) Initial data *b) After elimination* *c) After micro generalisation* *d) After displacement*

Figure 9: Example of an agent generalisation sequence on a suburban block

But, last year, it was not possible to present results on a large dataset because there were memory leaks problems in meso agent strategy and efficiency and quality problems for buildings displacement algorithm. These problems were solved by the student Loïc Gondol during an internship to 1Spatial in 2006. Loïc Gondol started off again from the displacement algorithm proposed by [Anne Ruas, 1999] in her PhD thesis (fig. 10).



a) Initial data without displacement *b) First version of buildings displacement algorithm* *c) Improvement of buildings displacement algorithm*

Figure 10: Improvement of buildings displacement algorithm at 1:25 000 scale

These improvements allowed to the project the launching of complete Agent process in Clarity on a whole map (20 km x 30 km) with 22 000 buildings on Orthez city sheet. 2000 blocks were detected by an automatic process and the automatic generalisation process is executed in around 5 hours. This process is used to obtain the map at 1:25 000 scale (fig. 11) with a weak generalisation and the map at 1:50 000 with more important generalisation (fig. 12). It is interesting to compare final results at 1:50 000 with initial data without any generalization (fig. 13). It is important to indicate that no editing has been done on the results.

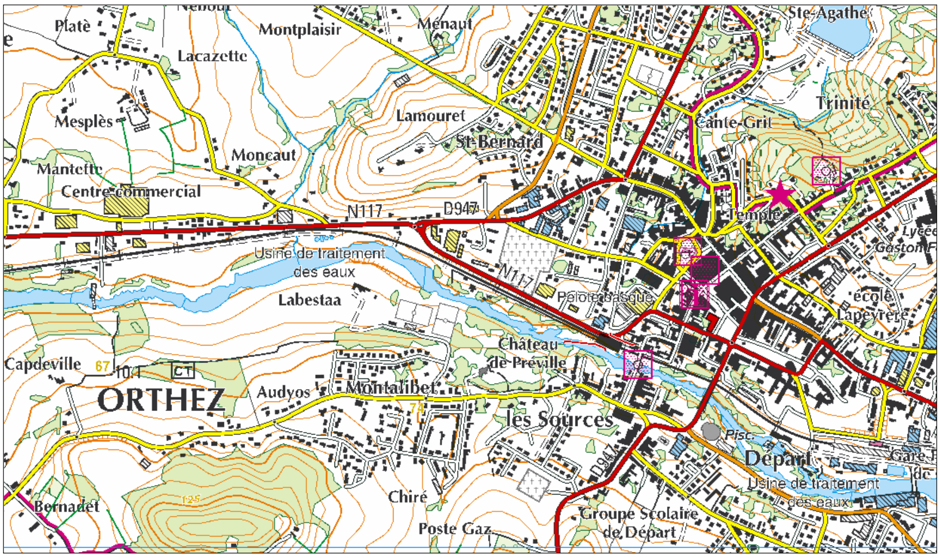


Figure 11: Map extract with buildings generalisation at 1:25 000 scale

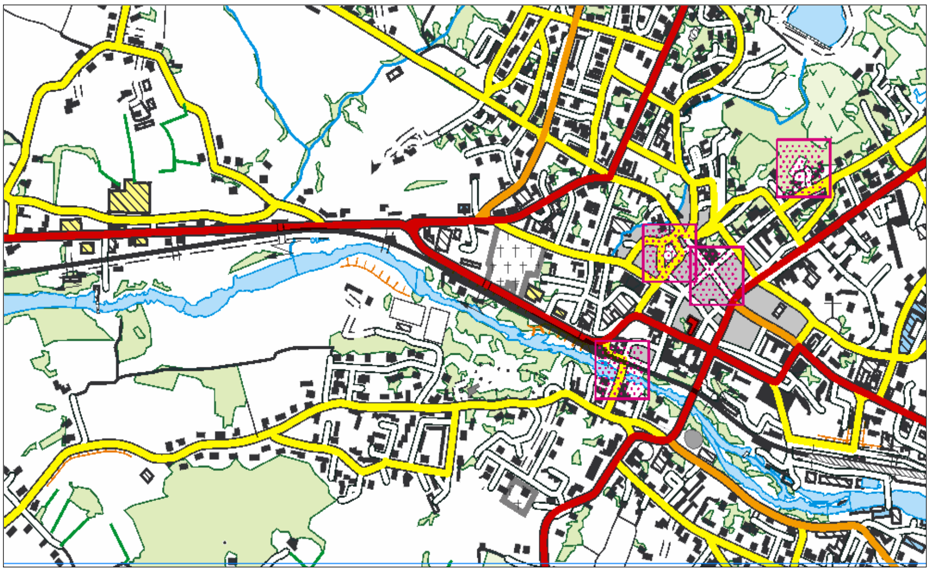


Figure 12: Map extract zoom (factor 2) with buildings generalisation at 1:50 000 scale

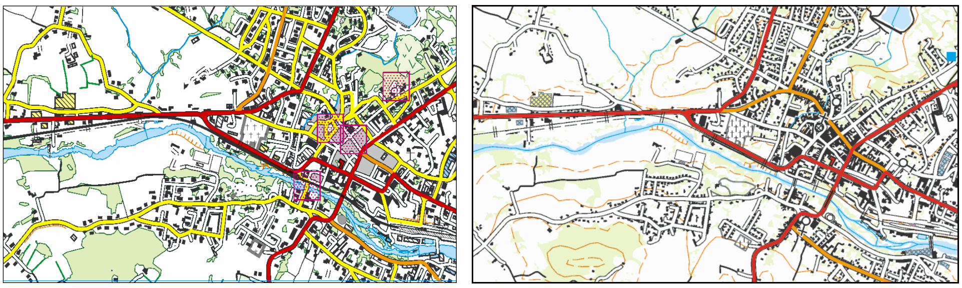


Figure 13: Map extract at 1: 50 000 scale (left) with buildings generalisation to compare with initial data without generalisation (right)

Map was provided to some traditional cartographers at IGN to evaluate buildings generalisation. Their comments are globally positive, with some minor remarks: "*The result describes well the landscape and a manual generalisation on a map extract provides a similar result. But some buildings are a little bit too small after generalisation and more buildings must be deleted*". These comments will be taken into account in the future.

To continue to define the complete process to generalize a map at the 1:50 000 scale, the New base Map Project wishes now to exploit Carto2001 results for roads displacement, with Beams algorithm [Bader, 2001], and flexibility graph [Lecordix & Lemarié, 2007] to automate the process on whole map. First tests had been done at 1:50 000 scale to displace roads and embankments (fig. 14), but the process is still not robust enough in Clarity to apply it on the whole map. We hope that last problems for this part of the process will be solved in few weeks.

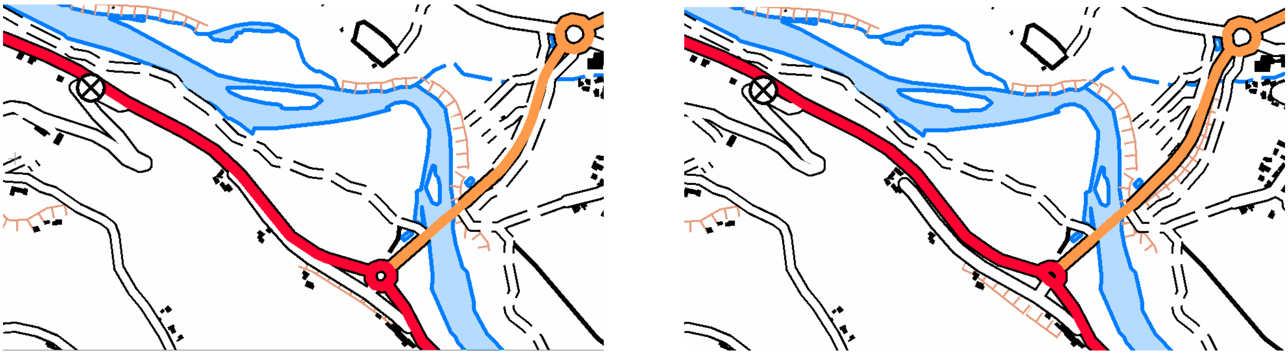


Figure 14: Roads and embankments displacement at 1: 50 000 scale with Beams algorithm and flexibility graph strategy

Conclusion

To develop a new flow line for topographic maps from BDTopo®, the New Base Map Project must solve a lot of different problems. This paper focused on data matching and generalisation that the project must apply on large dataset to obtain a first prototype. At this moment, the complete process is not finished and there is still some works, but important progresses have been done, first to introduce consistencies in data coming from different sources and then in generalisation, particularly for buildings. Solving step by step the principal phases of generalisation for large data set (data matching, buildings generalistion, rods generalisation,…), the New Base Map Project is hopeful to obtain in a very short time a complete prototype of map at 1:50 000 scale.

Reference

Bader, M. (2001). Energy Minimization Methods for Feature Displacement in Map Generalization. Dissertation zur Erlangung der Doktorwürde, *Mathematisch-naturwissenschaftliche Fakultät, Universität* Zürich, 2001.

Barrault, M., Regnault, N., Duchêne, C. Haire, K., Baeijs, C. Demazeau, Y., Hardy, P., Mackaness, W., Ruas, A., Weibel, R. (2001). Integrating Multi-Agent, Object-Oriented and Algorithmic Techniques For Improved Automated Map Generalization, *Proc. Of the 20th International Cartographic Conference*, vol. 3, Beijing, China, 2001, pp. 2110-2116.

Boffet A. (2001). Méthode de création d'informations multi-niveaux pour la généralisation de l'urbain. *Thèse de doctorat, Université de Marne-la-Vallée, 2001.*

Gaffuri, J. (2007) Outflow preservation of the hydrographic network on the relief in map generalisation, *Proc. Of the International Cartographic Conference*, International Cartographic Association, Moscow, Russia.

Haunert, J.-H. & Sester, M. (2004). Using the straight skeleton for generalisation in a multiple representation environment. *6th ICA Workshop on progress in automated map generalisation, Leicester 2004*

Lamy, S., Ruas, A., Demazeau, Y., Jackson, M., Mackaness, W. A. and Weibel, R. (1999). The Application of Agents in Automated Map Generalisation, *Proceedings of the 19th International Cartographic Conference (ICA'99), ICA/ACI (Eds.)*, Ottawa, Canada, pp. 1225-1234.

Lecordix, F. Jahard, Y., Lemarié, C., Hauboin, E. (2005). The end of Carto2001 Project. Top100 based on BD Carto® database. *8th Workshop on Generalisation and Multiple Representation,* Portland

Lecordix, F., Trévisan J., Le Gallic J-M, Gondol L. (2006) Clarity experimentations for cartographic generalisation in production. *9th Workshop on Generalisation and Multiple Representation, Portland*

Lecordix, F., Lemarié, C., (2007). Managing Generalisation Updates in IGN Map Production *Generalisation of Geographic Information: Cartographic Modelling and Applications. Chapter 15. Elsevier*

Mustière, S., (2005). Appariement et recalage sur la BDTopo des sentiers de randonnée saisis par la FFRP sur les SCAN 25®, *technical report*, IGN France (SR/2005.0386).

Ruas A. (1999). Modèle de généralisation de données géographiques à base de contraintes et d'autonomie. *Thèse de doctorat, Université de Marne-la-Vallée, France, 1999.*
Touya, G. (2007) . Etude sur la généralisation et l'intégration pour la construction d'un fond vert commun, *bulletin d'information scientifique et technique de l'IGN*, bilan de la recherche 2007.