OEEPE

European Organization for Experimental Photogrammetric Research

March 2001

Automatic Generalisation Project: Learning Process from Interactive Generalisation

Report by Anne Ruas

Official Publication Nº 39

Published by the Bundesamt für Kartographie und Geodäsie, Frankfurt am Main

Printed by Lausitzer Druck- und Verlagshaus, Bautzen/Budyšin

EUROPEAN ORGANIZATION
for
EXPERIMENTAL PHOTOGRAMMETRIC RESEARCH

STEERING COMMITTEE

(composed of Representatives of the Governments of the Member Countries)

Mr. C. VALORGE Centre National d'Etudes Spatiales 18, Avenue Belin F-31401 Toulouse Cedex	
Prof. Dr. D. FRITSCH Institut für Photogrammetrie Universität Stuttgart Geschwister-Scholl-Straße 24 D-70174 Stuttgart	Germany
Prof. Dr. D. GRÜNREICH Präsident des Bundesamts für Kartographie und Geodäsie Richard-Strauss-Allee 11 D-60598 Frankfurt am Main	
Prof. G. NAGEL Präsident Bayerisches Landesvermessungsamt Alexandrastraße 4 D-80538 München	
Prof. R. GALETTO University of Pavia Via Ferrata 1 I-27100 Pavia	Italy
Prof. Dr. Ir. M. MOLENAAR Rector International Institute für Aerospace Survey and Earth Sciences P.O. Box 6 NL-7500 AA Enschede	Netherlands
Ir. F. A. HAGMAN Survey Department of the Rijkswaterstaat Postbus 5023 NL-2600 GA Delft	
Mr. K. O. SUNDE Technical Director Land Mapping Division Norwegian Mapping Authority N-3511 Hønefoss	Norway
Mrs. E. MALANOWICZ Head Office of Geodesy and Cartography Department of Cartography and Photogrammetry ul. Wspólna 2 PL-00-926 Warszawa	Poland
Mr. F. PAPI MONTANEL Instituto Geographico Nacional General Ibáñez de Ibero 3 E-28003 Madrid	Spain
Mr. J. HERMOSILLA Instituto Geografico Nacional General Ibáñez de Ibero 3 E-28003 Madrid	
Prof. J. TALTS National Land Survey of Sweden S-80182 Gävle	Sweden
Prof. K. TORLEGÅRD The Royal Institute of Technology Dept. of Geodesy and Photogrammetry S-10044 Stockholm 70	

Prof. Dr. O. KÖLBL Institut de Photogrammétrie, EPFL GR-Ecublens CH-1015 Lausanne	Switzerland
Mr. C. EIDENBENZ Bundesamt für Landestopographie Seftigenstrasse 264 CH-3084 Wabern	
Dr. Eng. Col. M. ÖNDER Ministry of National Defence General Command of Mapping MSB Harita Genel Komutanligi Dikimevi TR-06100 Ankara	Turkey
Dipl. Eng Lt. Col. O. AKSU Ministry of National Defence General Command of Mapping MSB Harita Genel Komutanligi Dikimevi TR-06100 Ankara	
Mr. K. J. MURRAY Ordnance Survey Romsey Road Maybush Southampton S016 4GU	United Kingdom
Prof. Dr. I. J. DOWMAN Dept. of Photogrammetry and Surveying University College London Gower Street 6 London WC 1E 6BT	

SCIENCE COMMITTEE

Prof. Dr. Ir. M. MOLENAAR
International Institute für Aerospace
Survey and Earth Sciences
P.O. Box 6
NL-7500 AA Enschede

EXECUTIVE BUREAU

Mr. C. M. PARESI
Secretary General of the OEEPE
International Institute for Aerospace Survey
and Earth Sciences
350 Boulevard 1945, P. O. Box 6
NL-7500 AA Enschede (Netherlands)

Mr. E. HOLLAND
International Institute for
Aerospace Survey and Earth Sciences
P.O. Box 6
NL-7500 AA Enschede

OFFICE OF PUBLICATIONS

Prof. Dr. B.-S. SCHULZ
Bundesamt für Kartographie und Geodäsie
Richard-Strauss-Allee 11
D-60598 Frankfurt am Main

SCIENTIFIC COMMISSIONS

Commission 1:
Sensors, primary data acquisition and geo-referencing

President: Prof. Ø. ANDERSEN
Norges Landbrukshøgskole
Institutt for Kartfag
P.O. Box 5034
N-1432 Ás

Commission 2:
Image analysis and information extraction

President: Prof. Dr.-Ing. C. HEIPKE
Institute of Photogrammetry and
Engineering Surveys
Universität Hannover
Nienburger Str. 1
D-30167 Hannover

Commission 3:
Production systems and processes

President: Dr.Ing. W. MAYR
Vice-President Advanced Technologies
INPHO GmbH
Smaragdweg 1
D-70174 Stuttgart
Commission 4:
Core geospatial data

President: Mr. K. J. MURRAY
Ordnance Survey
Romsay Road
Maybush
Southampton SO164GU
Commission 5:
Integration and delivery of data and services

President: Mr. P. WOODSFORD
Deputy Chairman
Laser-Scan Ltd.
101 Science Park
Milton Road
Cambridge CB4 OFY

Table of Contents

Automatic generalisation project: learning process from interactive generalisation

with 43 figures, 46 tables and 1 appendix

Report by Anne Ruas
Chairwoman of the OEEPE working group on generalisation

Cogit Laboratory – IGN France
2 avenue Pasteur
94165 Saint Mandé Cedex – France

Summary

Since 1993, the OEEPE working group on generalisation is studying in a practical way algorithms, systems and processes in order to know more about current solutions in generalisation.

This report is the synthesis of the OEEPE test on generalisation started in 1996 and finished in late 1999. This test aimed at learning about generalisation sequences. Some data have been generalised on different platforms. Each process was described, step by step, and analysed. Finally processes were compared in order to identify common rules and particularities. The aim was neither to evaluate generalisation packages nor to evaluate generalised data.

The report is presented in seven chapters:

1. An introduction to the adopted evaluation principals: we present briefly current principals that could be used for evaluation: what and how to evaluate.
2. A presentation of the OEEPE test on generalisation: we present the rules of the test as well as the documentation used to perform it.
3. A presentation of the test data used: namely BDCarto® to study road generalisation, BDTopo® to study town generalisation and the ICC topographic database to study building generalisation.
4. The BDCarto® test: this test aimed at generalising road network from 1 : 50.000 to 1 : 250.000. As the processes were rather homogeneous, the results are presented all together.
5. The BDTopo® test: this test aimed at generalising a small town from 1 : 15.000 to 1 : 50.000. This test is the more complex one as this generalisation is very contextual. According to the heterogeneity of the platforms, it was not possible to group the tests results that are presented one by one. A synthesis enlightens the main properties and difficulties identified.
6. The ICC test: this test aimed at generalising buildings from 1 : 5.000 to 1 : 25.000. As the operation mainly consisted in applying very few algorithms, it was not interesting to analyse the sequence of generalisation. This generalisation is rather not contextual. Consequently we just identified main tendency and weak points that are coming from the current limitation of algorithms.
7. Conclusion and Outlook.

To conclude, I wish to remind that the packages used for the test correspond to the versions available at the beginning of the test (1997) and that some of them are proposing today some more functionality.

Acknowledgement

As chairwoman of the OEEPE working group on generalisation, I wish to thank all the partners who have accepted to participate to this working group by either performing some tests or reading the reports *Baella Blanca*, *Annabelle Boffet*, *Lars Harrie*, *Brigitte Husen*, *Monika Jordan*, *Sébastien Mustière*, *Maria Pla* and *Christophe Roux*. I wish I would have been able to test more systems and I apologise for systems that could not be tested (such as ESRI) due to lack of human resources. I wish to add that the intention was never to criticise systems but to provide information to help GIS providers in improving their systems as well as researchers in improving their prototypes and knowledge.

I wish to thank *Michael Lonergan* and *Christine Studer* who helped me to correct the final report and more specifically *Christoph Eidenbenz* who helped me a lot during the final stage.

Collective actions are never easy, they take time and money, but are nevertheless very helpful in making things better, even if 'what is better' is difficult to define. At least it creates a network of people having similar interests.

1 Principles Adopted for Evaluating Generalisation

Generalisation is the process which is used to provide a simplified representation of geographical information. But there is an infinite number of simplified representations of an individual area. As was already the case in manual cartography, this tendency is even more important in digital cartography. However, we cannot say that all representations are equivalent. One can argue that generalisation depends on user needs and, as user needs are non-predictable, we cannot find a generic approach to assess generalisation. Even if this were true, this attitude slows down our efforts in evaluating generalisation.

This part of the OEEPE report tries to propose some principles and methods that can be applied to evaluation. These principles are used to present the core of the OEEPE test on generalisation in the next chapters.

1.1 Why to Evaluate

1. Generalisation is a reduction of information: we should be able to describe what has been preserved, what has been removed, what has been enhanced.
2. Generalisation relies on the aggregation of information (in a semantic meaning). However, there are many ways to aggregate information: which one has been chosen?
3. Without taking parameter values into account, if one algorithm of generalisation is used, a solution is obtained. If an operation with the same name is used on another GIS generalisation package, another solution is obtained. What are the characteristics of the proposed solutions? Is there one which is more relevant than another one? The generalisation community defined a set of operations necessary for generalisation purposes. However, the problem is not solved because an algorithm is related to an operation: there are different ways to simplify, aggregate, collapse, displace, remove, emphasise and caricature objects. What is behind those terms? What is the guarantee of obtaining something which preserves geographical meaning?
4. In the context of web, data will become more and more available at different levels of detail. But how do they reflect a geographical reality? How can we be sure that this aggregated and simplified information is appropriate for a specific purpose?

The challenge in evaluating generalised data is to be able to understand the value of generalised data. Two main questions are:

- Is the data appropriate for the user needs (fitness for use)?
- Does the generalised data still reflect reality? How?

1.2 Which products to Evaluate

Evaluation requires not only criteria (what to measure) but also a real methodology as it is presented in quality management by the National Mapping Agencies (NMAs) [*Dassonville* 1999]. Quality management has two objectives:

– being able to qualify a product
– being able to improve the quality of a product .

The first stage to improve the quality consists in describing accurately how a product is done, its life cycle: how is the production team organised, which process is used, what are the input data, what are the controls during the process, when do they occur, what is the feedback of the user, etc.

The product to be evaluated is either the result or the system (figure 1). These viewpoints are coherent with quality methodology as:

– Evaluating the results is useful in itself but can also be used to improve the system.
– Evaluating the system is useful to improve the system itself and consequently the future results (output).

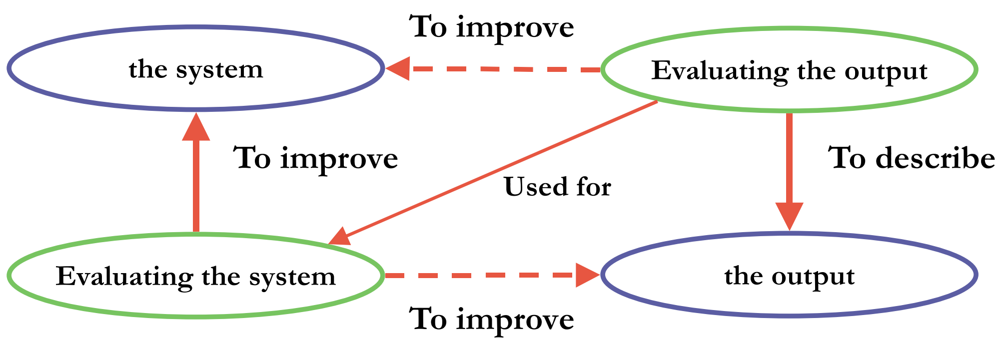


Figure 1 – Loop in evaluation between system and output.

This loop of evaluation between system and result is even more crucial in a research context where the aim is to be able to improve system by proposing new methods. In a production environment, quality control is very expensive, therefore it is restricted to a minimum whereby a maximum of quality with respect to customer satisfaction, price and the strategy of the company (concurrency, image, etc.) can still be ensured.

Consequently, for generalisation we should evaluate both:

– the generalised data,
– the system used to generalise the data.

1.3 What to Evaluate

Research performed in the GIS community concerning quality focuses mainly on identifying geometrical measurements for evaluating the accuracy of digitised data. Even if this aspect is essential for the acquisition stage, it is very incomplete for generalisation. A study [*Joao* 1998] enlarged the focus in measuring the effect of displacement on generalised data, but it also concerns geometric accuracy. A first remark we can make is that generalisation requires more control than geometry. From a research point of view, we can evaluate the generalised data, one generalisation programme, several generalisation programmes, several generalisation algorithms or generalisation processes (table 1). Such evaluations will provide different outputs.

Table 1 – What to evaluate, and what for.

WHAT to evaluate	What for
Generalised data	To know their quality To know how to improve their evaluation
One generalisation programme	To know its quality To know the kind of generalisation it can perform
Its generalisation algorithms	To know how they work To know what is missing
Its dynamics	To know how far we are from automation
Its results (generalised data)	To know what we can obtain with this software
Several generalisation programmes	To know which one is the best To know which ones are relevant for which type of generalisation
Several generalisation algorithms	To compare and classify them
Generalisation process(es)	To learn a generalisation sequence To improve procedural knowledge To identify what is missing or not perfect

To understand accurately what to evaluate and what for, it is useful to identify the components of one generalisation as shown in figure 2. Such a sequential approach corresponds to the current GIS generalisation packages. The 'sequences used' are by default not provided as output in production environment but are very helpful in research.

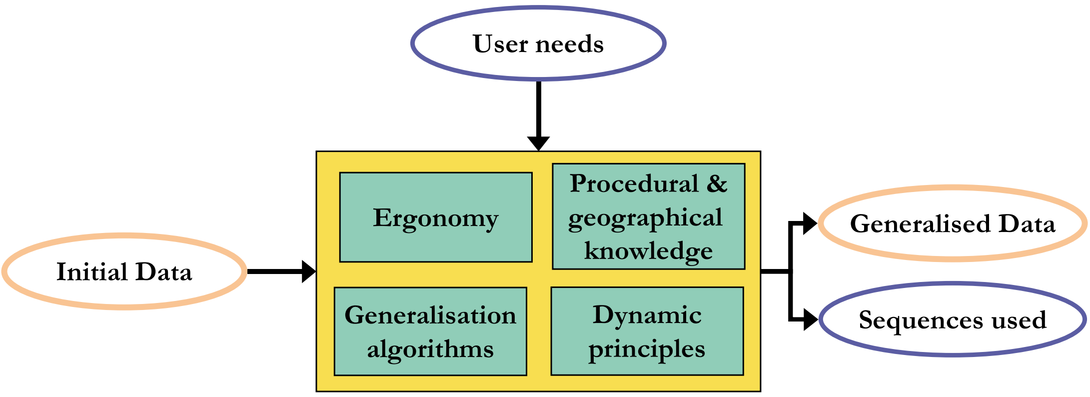


Figure 2 – Component for a generalisation.

We divide the generalisation evaluation into four parts:

- Evaluation of the software without its dynamics (algorithms, ergonomic, etc.)
- Evaluation of the dynamic capacity of the software (which uses knowledge and data to trigger appropriate algorithms)
- Evaluation of generalised data
- Evaluation of one process used for a specific generalisation (how did it work?)

1. **Evaluation of the software without its dynamics**
   1.1. Which software is used?
   1.2. What is required in terms of computer configuration?
   1.3. What is required in terms of input information (data modelling, quantity of data)
   1.4. What are the characteristics of the algorithms:
      1.4.1. For which information are they adapted?
      1.4.2. How do they change information?
      1.4.3. What are their parameters?
      1.4.4. What is the meaning of the parameter? Are they linked to the qualities the system knows?
      1.4.5. What is their robustness?
      1.4.6. Which cases can they not manage?
      1.4.7. What is their computational speed (complexity)?
   1.5. What is the capacity of the software for introducing user specifications? Is it possible? Is it easy?
   1.6. Are there internal and final checks to ensure that user specifications are taken into account during the process?
   1.7. Is the system ergonomic? In particular for the following aspect:
      1.7.1. Loading the data
      1.7.2. Setting the specifications
      1.7.3. Choosing the algorithms
      1.7.4. Choosing the parameter values
   1.8. Is the software able to follow research evolution? Is it possible to include new algorithms?
2. **Evaluation of the dynamics of the software**
   2.1. Is it a batch process?
   2.2. Is it an interactive process? To what extent?
   2.3. In case of an automated process:
      2.3.1. How are algorithms chosen?
      2.3.2. How are parameter values chosen?
      2.3.3. How is the sequence of generalisation chosen?
      2.3.4. Which mechanisms of choice are used?
         2.3.4.1. Does it use heuristics? When? What for?
         2.3.4.2. Does it choose between different solutions? How? When? What for?
         2.3.4.3. Is there a backtracking capacity?
         2.3.4.4. Is the process stochastic? When? What for?
   2.4. Are there cases which the system cannot handle? What can be done in such a case?
3. **Evaluation of generalised data**
   3.1. Do the generalised data satisfy the user needs? In what way?
   3.2. Does the solution respect the specifications included in the system from the user needs?
   3.3. Are there data which could not be generalised? Which ones? Why?
   3.4. Is the geographical meaning retained?
4. **Evaluation of the process used for one generalisation**
   4.1. Was it easy to load the data?
   4.2. Was it possible and easy to set the specifications?
   4.3. Which sequence was used? (objects and operations)
   4.4. What are the algorithms used? In which case? With which parameter values?
   4.5. Are there some algorithms that were not used? Why?
   4.6. Are there cases the system could not handle? Are these cases identified by the system?

4.7. How long does the system need to reach a solution?
    4.7.1. How many solutions were tried before a solution was reached?
    4.7.2. Are there any convergence problems?
4.8. What is the cost of the process (computer and human time)? How much time was devoted to each action? (enriching, choosing, generalising, checking)

1.4 How to Evaluate

We saw previously that it is possible to evaluate different parts of generalisation for different purposes. Whichever part is to be evaluated, some decisions have to be taken prior to the evaluation:

1. Which questions are raised by this evaluation?
2. Which geographical data set should be used as input?
3. Which criteria should be used?
4. Which references can be used for comparison purposes?
5. What are the evaluation methodologies? Who or which tool will make the evaluation?
6. How can we synthesis evaluation results?

These questions will be used to present the OEEPE test in the next chapter.

Which questions are raised by an evaluation?

Before starting an evaluation, the first question is to identify which part is supposed to be evaluated. It can be:

– the results (the generalised data): What does this package give in terms of generalisation result?
– the algorithms: What is the quality and exhaustivity of the generalisation algorithms inside the package? (i.e. how do they work, what can they handle and not handle)
– the dynamics of the system: how good and fast are the decisions made?
– the platform: is the package easy to use? (ergonomy, flexibility to specify, possibility to correct mistakes, etc.)

Such an understanding is necessary to specify the evaluation methodology.

Which geographical data should be used?

For defining an appropriate data set, two options are possible. Either the data are a set of existing data from databases or they are created for test purposes.

– Data from databases:
  – Which data source (content and accuracy)? Is it possible to test only a theme (e.g. road network) or all themes?
  – To which extent? Here the challenge is to find the balance between finding the minimum set for obtaining relevant results with different spatial configurations and the quantity of work for performing and analysing the test
– A representative subset: samples are defined in order to test specific parts of the system.

In order to test only algorithms, [*Duchêne* 2000] proposes the creation of appropriate samples of objects (buildings) where only one variable changes at a time. Thus, the system works on one object only and the aim is to see its behaviour. This method allows questions to be answered such as: Does the orientation of a building influence its generalisation? Is the system able to generalise a building (such as a church) which has a repetitive sequence of the same shape? For such a test, samples are created to allow an analysis of each algorithm behaviour in different but known situations.

Comparing what to what?

Evaluating can be performed in two complementary ways:

– describing the result (system or output) as it is,
– comparing the result (system or output) with others taken as references.

An evaluation is rarely performed without making a comparison – at least the generalised data are compared with the initial ones.

If we follow the hypothesis that a generalisation comprises the components shown in figure 2, we can identify what can be compared to what for evaluation purposes (figure 3). We also add manual generalisation as reference since there are still a lot of examples that can be used.

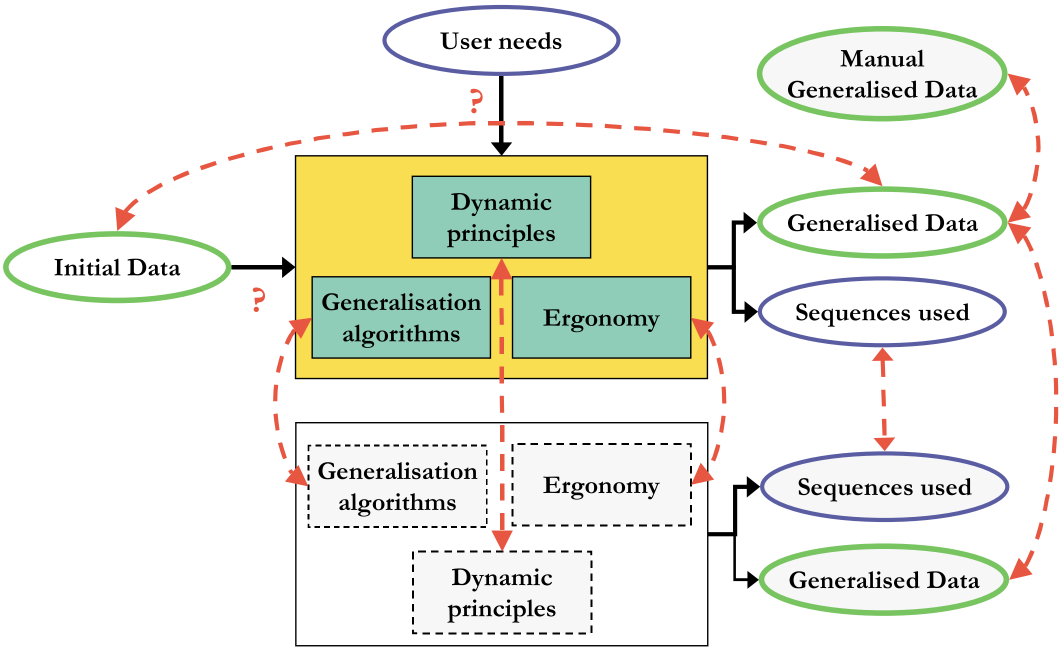


Figure 3 – Comparing what to what when evaluating generalisation.

However, some of these comparisons are easier than others. In order to compare the generalisation packages, normally the easiest way is to compare the outputs because comparing the core of the system is usually not possible. Thus we compare either generalised data with other data or sequences used with other sequences.

In the following we try to identify what can be compared to what, and the advantages and drawbacks of such comparisons:

Compare the generalised data to what?

– to other generalised data,
  – advantages: it allows the identification of differences and tries to understand why dif-

ferences exist. Ideally it allows the identification of the best one according to certain criteria.
  - drawbacks: different generalisations are possible for an identical space. How can the best one be identified? Is the reference used equivalent to a model or is it just another possible solution?
- to the initial data
  - advantages: it is possible to check if the specifications are respected and to evaluate if the generalised data still reflect the initial geographical information.
  - drawbacks: it is not evident if the solution could have been better or not.
- evaluate the generalised data, without references
  - advantages: it is possible to concentrate the checking on the readability of the generalised data.
  - drawbacks: it is impossible to check if the generalised data are still meaningful.

Compare the generalisation process to what?

The process used (sequences of objects, operations, algorithms and parameter values) can be compared to:

- another process performed on another system:
  - advantages: it isolates the influence of the available algorithms on the process and can be of help in finding generic sequences.
  - Drawbacks: since the tools are different (some operations may be missing), a comparison of processes is not easy.
- a process performed interactively on the same system:
  - advantages: because the tools are the same, it helps in analysing why some optimum solutions were not chosen, and in finding out the number of tries necessary to reach a solution.
  - drawbacks: the process performed interactively has to be done by someone who has cartographic skill and who knows the system capacities well enough to really explore all the possible solutions.

Who performs the assessment?

The evaluation of the test has to be done by an entity. This entity can be:

- the system which was used for generalisation
  - advantages: the evaluation is a part of the system which yields directly the generalisation data and some evaluation values.
  - drawbacks: the evaluation criteria might be the same as those which were used for the generalisation process. In such a case, the system might be self-sufficient. Moreover, the comparison between different generalisations performed on different systems is not possible
- another computer module
  - advantages: this module is independent from the system which performed the generalisation. It is possible to compare different generalisations. Quantitative results can be used more easily than qualitative ones for improving the system.
  - drawbacks: it relies on the hypothesis that the research community is able to find a complete set of digital measurements appropriate for evaluating a generalisation.
- an expert cartographer (who did not perform the generalisation)
  - advantages: an expert does not try to defend a system. He has an 'objective view' of the system.

  - drawback: an expert is always subjective and an assessment is his own evaluation of what should be an appropriate generalisation. Comments might be very general or based on fuzzy criteria.
- the user who performed the generalisation or who design the system
  - advantages: the user knows about the system and can interpret the choices made.
  - drawbacks: there is a lack of objectivity of the analysis. The user can have low cartographic skills.

Which criteria?

Finding the appropriate criteria is certainly the most important issue of an evaluation. We can distinguish between the following two different criteria:

- from cartographers: they are not concerned with the capacity of implementing them and concepts do not have accurate definitions. For example, pattern and shape retention will have a different significance for different cartographers.
- from the GIS community: research was carried out in order to try to find appropriate measurements and interpretation of values to qualify some phenomena [*McMaster* 1983] [*Weibel* 1996] [*McMaster* and *Vergerin* 1997] [*Weibel* and *Dutton* 1998] [*Joao* 1999]. But there is still a discrepancy between what is intended to be qualified and the current state of research. For example, the description of patterns remains very inadequate and even shape computation is still difficult.

How to aggregate measurements to provide useful quality indicators?

Another issue is to be able to synthesise a set of evaluation results (quantitative or qualitative description). Three problems occur:

1. How is a range of values synthesised for one criteria? For example, if accuracy is computed for each generalised object (e.g. by means of the Hausdorff distance) there will be a distribution of values: how can they be aggregated? Is the maximum or the average value significant? Is it not more useful to draw the accuracy distribution as a result?
2. Are numbers significant or do we have to find a qualitative interpretation of the value? For example, is an average accuracy of 5 good or not?
3. Since a set of criteria are necessary (shape, accuracy, distance, etc.), do we need to find a common scale between different criteria?

Evaluating the system evaluation:

A given system might provide evaluation criteria as complementary output. Such capacity should also be evaluated for answering the following question: Is the evaluation provided by the system meaningful? One solution may consist in introducing already generalised data (that are considered as good) into the system and matching these data with the non-generalised ones. In such a case, the system can compute its evaluation criteria which can be compared with the evaluation criteria provided by the automated generalisation (figure 4).

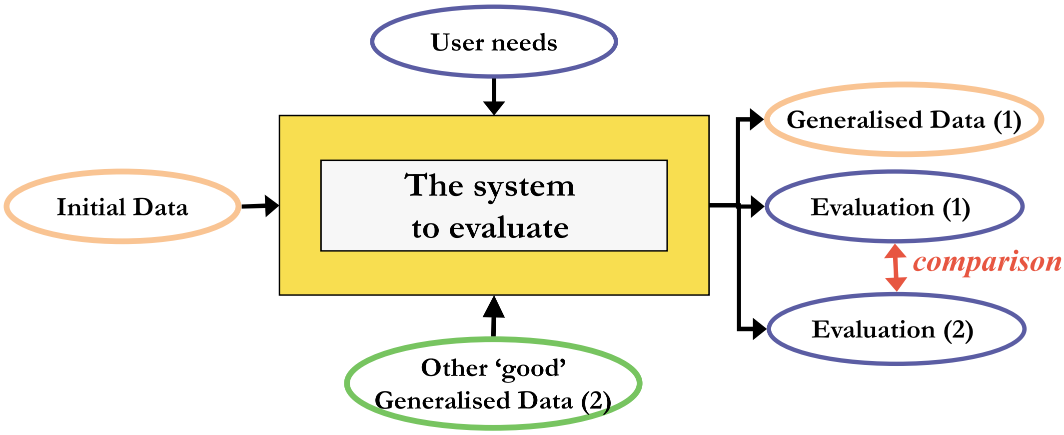


Figure 4 – Evaluation of the system evaluation.

1.5 What has been done by the OEEPE Generalisation Working Group

From the beginning the OEEPE working group in generalisation has undertaken a set of actions for improving our knowledge on generalisation and on generalisation systems. We started with reports to identify generalisation needs.

Some preliminary tests were performed in order to analyse the available generalisation packages:

- **MGE/MG (Intergraph)**
  - IGN 94 *Rousseau, Rousseau, Lecordix*
  - ICC 94 *Baella, Colomer, Pla*
  - GIUZ 95 *Weibel, Erliholzer*
- **CHANGE (University of Hanover, Zeiss at this time)**
  - ICC 94 *Baella, Colomer, Pla*

These preliminary tests were helpful in identifying the quality of the algorithms and their conditions of use. The main interest was to furnish GIS providers with feedback for identifying which parts of their software were appropriate and which parts should be improved according to generalisation needs. All of these actions were carried out by the working group under the supervision of Jean-Philippe Lagrange.

Then we defined two templates by means of questionnaires in order to better describe the content of the tools inside a system. This action was motivated by the fact that behind a word such as "simplification" "aggregation" or "displacement" different transformations are possible and we though that an accurate description would help. Thus we defined:

- A generic template to describe algorithms
- A generic template to describe measurements.

We do not know how far these templates have been used by different partners, but at least the IGN used it to describe a part of their algorithms and measurements. Such documents [OEEPE 1997] have been proposed to the partners. From a practical point of view, such descriptions help a lot not only for communication purposes among the research community but also within a team to make the effort of formalising and homogenising the content of what has been developed. In the future an on-line description of these algorithms and measurements should be available for the user (if required). A development team from the NMAs would at least be better informed as to what is behind the name of a generalisation operation.

The last action of the OEEPE working group was to organise a more global test in order to better understand the process of generalisation using existing GIS packages, either commercial ones (MGE/MG from Intergraph, CHANGE from the university of Hanover and LAMPS2 of Laser-Scan Limited) or prototypes (PlaGe and Stratège from the Cogit laboratory of the IGN). We would have very much liked to analyse other prototypes but under the circumstances we did the best we could. This test is presented in the following chapters as well as its results.

1.6 Conclusions

I let the readers make their own conclusions on the presented test performed by the working group. After having done it, I think that I could have organised it in a better way, I am not fully satisfied by my own methodology. At least it is an attempt to deal with evaluation and it throws light on some difficulties of which I and others were not fully aware. In terms of output, some accurate information is given in the following chapters, but more important is that the test helped to improve the description criteria (see conflict-code templates in Chapter 2.6) and above all it triggered new ideas on the methodologies and points to define in order to make some new generalisation tests. I tried in this chapter to list these aspects that should be considered prior to defining a test and I hope it could be useful for future research in evaluating generalisation.

2 The OEEPE Generalisation Test

2.1 The Context of the Test

The OEEPE working group on generalisation was set up in 1993. Its purpose is to examine the ability of currently available generalisation systems for solving practical problems. One of the objectives is to suggest areas that would benefit from further study. Generalisation is the process used to simplify geographical information in order to satisfy given criteria while maintaining consistency and homogeneity. Up to now this process has been performed by professional cartographer who used their knowledge and experience to eliminate unnecessary details while accentuating information which needs to be maintained. Now that computerised databases are available, we are trying to develop some tools and mechanisms that can carry out the generalisation process with a minimum of human effort. At present GIS offer very limited facilities for such work: either single filtering and smoothing algorithms are provided, or the package has a set of algorithms that must be triggered interactively. However, as generalisation has been studied for years, partial solutions to the problem already exist. These need to be evaluated. The aim of the OEEPE test on generalisation is to study the capacities and deficiencies of these solutions. The assessment focuses on the generalisation process as much as its results. Many algorithms have already been developed to tackle individual generalisation tasks. While this complex task should continue, the most urgent questions address their use: **when, on which objects, with which parameter values, in which order?**

The GIS community can be divided into four sectors: GIS providers, data producers (such as NMAs), GIS users and universities. From the beginning the working group has tried to involve representatives of all four groups. The OEEPE WG includes nearly 40 partners (with a core of 10 active partners) from all over Europe and four sectors. This working group has strong links to the ICA working group in generalisation, though ICA focuses on more theoretical issues.

Initial discussion of this test took place in 1996 during the SDH working group meeting and the first proposal was presented to -and accepted by- the OEEPE commission in October 1996 during the Helsinki meeting. The test involves the generalisation of three small data sets with different contents and resolutions, along with the detailed description of each step: for each object, the cause of generalisation, the chosen method – operation, algorithm and parameter value – and a visual assessment of its result were given. The aim is mainly to create an initial procedural knowledge base, i.e. to identify relations between geographical situations and proposed solutions. As the GIS platforms differ, it is also a way to compare their characteristics, strengths and weaknesses. Of course our common aim is to learn more about generalisation and to improve the platforms in order to develop a more robust and less interactive generalisation process. The test is of the various algorithms and implementations, so no 'point-to-point' corrections were permitted or carried out.

2.2 Partners and Data Sets

As described previously, we wanted to include as many of those involved in GIS as possible. The test was carried out:

– **By two NMAs:** IGN (*Christophe Roux, Sebastien Mustière, Annabelle Boffet, Anne Ruas* – France) and the ICC (*Maria Pla, Baella Blanca* – Catalan) and **five universities:** Hanover (*Brigitte Husen*, IFK, Germany), Munich (*Monika Jordan*, Germany); Glamorgan (*Chris Jones*, UK), Edinburgh (*William Mackaness*, UK) and Lund (*Lars Harrie*, Sweden).
– on five different platforms: **LAMPS 2** from Laser-Scan, conducted by Laser-Scan; Lund and Edinburgh; **Change** from the University of Hanover, conducted by Hanover and the ICC; **MGE/MG** from Intergraph, conducted by the ICC and the University of Munich; **MAGE** from the University of Glamorgan, conducted by the University of Glamorgan, and **PlaGe** and **Stratège** from the IGN Cogit Laboratory, conducted by IGN.

As these platforms have different characteristics and are devoted to different data scales, we used three data sets. These also allow the study of different aspects of generalisation process issues. The three data sets are:

– **BD Carto® network data:** a set of IGN-France roads and railways at a resolution of approx. 10m that has to be generalised at 1/250K scale. The area contains valleys and mountains, thus providing a large and difficult set of line generalisations.
– **BD Topo® data:** a set of IGN-France data set, at a resolution of 1m, which has to be generalised at 1/50K. This area represents a village. Such a data set is too dense to be represented at such a scale without modification and requires contextual operations.
– **ICC large-scale data set:** a set of large-scale urban data used at 1/5K that has to be generalised to the scale 1/25k. Due to the lack of human resources, the result of the test will not be presented in the present report. Only small conclusions are presented in chapter 6. I wish to apologise for this, essentially to the group that performed the test.

To allow the comparison of the treatment of individual objects, each one was given an identifier.

2.3 *The Templates*

A set of documents was prepared to ease the analysis of the test's results. A first draft was written at the IGN and two meetings between partners were organised:

– one at the ICC between IGN, ICC, Hanover and Munich on the 3 and 4 of February 1997;
– a second at Laser-Scan Cambridge between Lund, Edinburgh, Glamorgan, Laser-Scan and the IGN, on the 17 and 18 of February 97.

These two meetings resulted in templates suitable for all the platforms, thus providing a consistent assessments of conflicts and evaluation codes.

1. The process template: This describes the process step by step. Each line of the template describes: the object identifier (or the area); a conflict code; the method for conflict identification; the operation code; the algorithm code; the parameter value(s); the method for algorithm choice; an evaluation of the result (bad-medium-good); the remaining conflicts that cannot be solved; some comments if necessary.
2. The conflict code template contains 62 different codes, organised in different sections of varing specificity. They describe conflicts due to granularity, size, shape, proximity, overlapping, relative position, absolute position and density; within a line, a polygon or between a set of lines, a line and a polygon or a set of polygons.
3. The operation code template contains 20 codes corresponding to the main global transformations performed on objects during the generalisation process.
4. The algorithm code template contains 80 codes which correspond to all the algorithms that exist on the different platforms at this time. They are grouped by their characteristics to facilitate their comparison.

The last three templates were used to fill the following process template which is used for the analysis:

Table 2 – Process template to fill.

Object id.	Main conflicts	Operation	Algorithm	Parameter values	Assessment bad med. good	Remaining conflicts	Comments

Conflict and algorithm were usually attributed interactively.

Each partner was responsible for its test and the submission of a completed, filled process template and a drawing of the generalised data. The content of the conflict code template is given at the end of this chapter.

2.4 *Analysis of the Results*

The generalisations were performed on different sites and the resulting completed process templates sent to the IGN in Excel forms (see example in table 3).

The differing capacities of the GIS, along with the complexity and subjectivity of the test, generated some differences in the results which explain why the analysis is not entirely homogeneous. Each Excel file was reorganised for each test in order to try to extract as much information as possible. The analyses are both qualitative (understanding the global process) and quantitative (analysis of quantity, frequency and correlation). It may be too quantitative (at least for BDTopo® analysis) but how to prove the conclusions?

Table 3 – Example of template for one urban area (number 24) generalisation.

Work. Areas	On ..	type of obj.	conflict code	conflict code 1	conflict code 2	operation code	algorithm code	parameter value [m]	conf assess	final assess.	f-c 1	f-c 2	f-c 3	f-c 4
24	some small buildings	Bs	121, 480	121	480	4			2					
24	small street	S	280	280		4			2					
24	buildings in the south	Bs	121, 110, 130	121	110	18	143	true centroid area tolerance: 0.0	2					
24		Bps	580, 540	580	540	6	21	25.0 balancing on	2					
24		Bps	361	361		7	131	threshold distance to nearest linear	2					
24	buildings in the north	Bs	121, 110, 130	121	110	18	143	true centroid area tolerance: 0.0	2					
24	two right buildings	Bps	361	361		7	131	threshold distance to nearest linear	2					
24	two left buildings	Bps	361	361		7	131	degrees: 0.0 threshold distance	2					
24									1	131,350, 540,252	131	350	540	252

2.5 Summary of the Decision for the Test

1. **The aim of the test was to evaluate generalisation processes.**
   1.1. The questions raised were
      1.1.1. What is generalised? (Which data? In which proportion?)
      1.1.2. When? **Are there thematic orders to generalise data?**
      1.1.3. How? (operation, algorithm, parameter value)
      1.1.4. Why? **Are there relationships between identified conflicts and chosen solutions.**
   1.2. We did not want to evaluate the generalised data.
   1.3. We did not want to evaluate the self-dynamics of the system because too few systems had self-dynamic capacities.
   1.4. The test also provides some information on the quality of the algorithms used as secondary results. In such a case we wish to describe these results.
   1.5. In order to compare the use of the algorithm, no point by point corrections were allowed. The user was asked not to solve conflicts if no algorithm exists. The only authorised intervention besides using algorithms was an object removal which was necessary for the BDtopo® generalisation test.
2. The data used as input are three data sets, a small amount for studying different kinds of generalisations. Data sets were sent in DXF format with a descriptive file explaining the content of classes and attributes.
3. The criteria used to evaluate the processes were divided in two parts:
   3.1. Several 'generalisation criteria' (called conflict codes) for describing why a decision was taken and what the remaining conflicts – either generated or unsolvable – were..
   3.2. Analysis of homogeneity or heterogeneity of sequence, operations, algorithms and parameter values used. A special statistical analysis of relationships between a conflict and a solution was performed.
4. The references used for this test were divided in two parts:
   4.1. Maps at the final scale for different areas for showing what was requested in terms of generalisation. No digital symbolisation specifications were sent (it was a mistake) .

    4.2. The generalised data could have been used to make a comparison, but this was not the aim of this test.
5. The methodology used for the test was the following:
    5.1. Templates of code references were defined by the group and shared by each one.
    5.2. Generalisation was performed on different sites by a partner, each process was recorded in an Excel file, step by step, according to these code templates.
    5.3. Processes were analysed one by one by *Anne Ruas* using Excel analysis tools. Unclear aspects in each process file generated information exchanges between partners.
    5.4. Some aspects of the processes were compared by *Anne Ruas* using Excel analysis tools.
    5.5. The report was written by *Anne Ruas* who tried to interpret the results.
6. The synthesis of information was done by aggregating values together with respect to the code templates and by processing some statistical analysis when it was possible (enough data).

The first version report was read by each partner who performed the test for validation. The next version was corrected by *Christoph Eidenbenz* and *Christine Studer*. The final version was corrected by *Anne Ruas* and sent to the OEEPE in January 2001.

2.6 The Conflict Code Template

The conflict code template was defined at the beginning by the partners and improved according to the needs.

Table 4 – The conflict code template elaborated during the test.

conflict	code	1 line 000	Code	1 polygon 100	code	set of lines 200
granularity 10..	**10** **11**	Too detailed too short	**110**	too detailed		---
size 20..	**20**	Too thin (double line)	**120** **121**	too thin too small		---
shape 30..	**30** **31**	Too many angularities loss of shape characteristic **	**130** **131**	not squared enough loss of shape characteristic	**230**	too segmented set of continuous lines---------- --
proximity 40..	**40** **41** **42** **43**	Generic internal proximity internal proximity (bend extremity) internal proximity (one bend) internal proximity (set of bends))		-------------------------------------	**240** **241** **242** **243**	generic proximity conflict local proximity continuing proximity (2 lines) continuing proximity (set of // lines)
overlapping 50..	**50** **51** **52** **53**	Generic internal overlapping internal overlapping (bend extremity) internal overlapping (one bend) internal overlapping (set of bends)		-------------------------------------	**250** **251** **252** **253** **254**	local overlapping too sharp connection continuing overlapping (2 lines) continuing overlapping (set of // lines) geometrical overlapping
relative position 60..		--		-------------------------------------	**260** **261** **262** 263	nearly aligned lines (junction) loss of linear spatial structures (gesalt) loss of connectivity **loss of one shortest path**
absolute pos. 70..	**70**	Too far from initial geometry	**170**	too far from initial geometry **		---
density 80..	**80**	Too many points	**180**	too many points	**280** **281**	over density of lines not readable junction (too many close lines)

conflict	Code	1 line and 1 polygon/ 1 symbol 300	Code	set of polygons 400	code	set of symbols 500
shape 30..	**330**	Sausage shapes-(sequence of polygon-line)	421	**Loss of relative building size**------------------- -----------------------		-------------------------------------- --------------
proximity 40..	**340**	Too close	**440**	too close polygons	**540**	too close symbols
overlapping 50..	**350**	Overlapping (with symbolisation)	**450**	polygons overlapping	**550**	symbols overlapping
relative position 60..	**360** **361** **362** **363** **364**	Generic bad relative position Line and polygon (or symbol) badly oriented Loss of relative orientation bad relative position (wrong side of a line) loss of sharing geometry	**460** **461** **462** **463** **464**	generic bad relative position bad relative alignment polygon and hole orientation loss of inclusion loss of polygons spatial structure (gesalt)	**560** **561** **562**	generic bad relative position bad relative alignment loss of spatial distribution (gesalt)
density 80..		--	**480** 481	over density of polygons **homogeneity of density**	**580**	over density of symbols

Semantic : **600** : semantic inconsistency ; **601** : Orogaphic inconsistency **;** **602** : Non acceptable aggregation,
...**610 : non characteristic**

3 Presentation of the Test Data Used

In order to study different aspects of generalisation, we have decided to use three small data sets:

- A road network extracted from medium scale data base: IGN France BDCarto®, to study road generalisation
- A village extracted from IGN France BDTopo® to study contextual town generalisation
- An extract of town at large scale from the ICC Mapa topografica de Catalunya to study building generalisation.

3.1 The BDCarto® Test

3.1.1 The IGN BDCarto® Data

The IGN BDCarto® database is created from 1 : 50.000 scale map digitalisation and actualisation. The data have an accuracy of ten meter, corresponding to a 1 : 50.000 scale map without any generalisation. It is used at the IGN to produce 1 : 100.000 scale map with light generalisation. (see http://www.ign.fr/fr/MP/BDGeo/BDCARTO/)

For the test, only roads and railways were selected in an area near Valence that contains mountains and valley in order to encounter different types of line geometry. Road selection has already been performed before the test to simplify it.

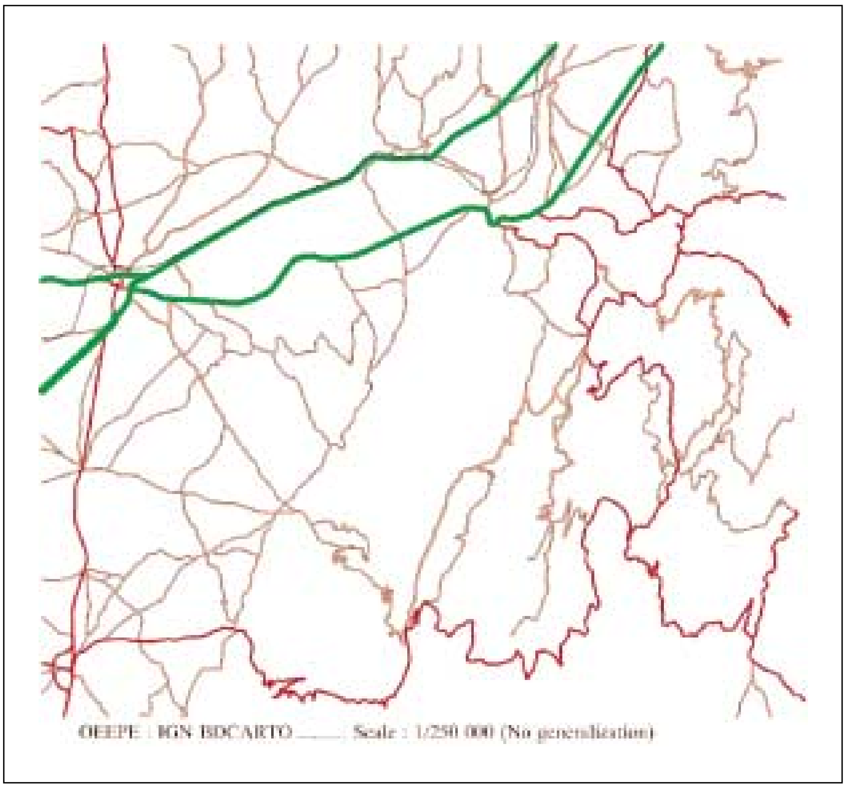


Figure 5 – The BDCarto® road data set before generalisation.

The data set contains: 355 arcs, 8681 points, 747 field kilometres. The remaining objects are divided into 7 classes and related code:

- large route 3
- route 5
- 1st road 10
- 1st narrow road 11
- 2nd road 16
- 2nd narrow road 17
- railway 18

On figure 5 we can see many conflicts due to line symbolisation: overlapping between different lines and within a line. Table 5 summarises the number of objects to generalise.

Table 5 – Number of objects and geometric characteristics.

Kinds of lines →	3	5	10	11	16	17	18	Total
Number of arcs	3	37	1	89	50	161	14	355
Number of points	20	247	2	2246	1350	4666	150	8681
Total length (km)	4,35	53,5	0,03	181,6	105,5	372,7	29,0	747
Average nbr points / arcs	6,7	6,7	2,0	25,2	27,0	29,0	10,7	24,5
Average length / arc	1,45	1,45	0,03	2,04	2,11	2,31	2,07	2,10

In order to prepare the data set of this test, I was helped by *François Lecordix* from the Cogit laboratory.

3.1.2 The Required Generalisation

Partners were asked to generalise the roads without performing any object removal. The final requested scale was the 1 : 250.000 maps. The width of the roads at the final scale was given as a constraint, according to their type. To show the kind of requested generalisation, a 1 : 250.000 scale map was given.

3.2 The BDTopo® Test

3.2.1 The Initial BDTopo® Data

The IGN BDTopo® database is created from stereo plotting and field work. The data has an a accuracy of one meter, corresponding to a 1 : 10.000 scale map without any generalisation (see http://www.ign.fr/fr/MP/BDGeo/BDTOPO/)

To limit the work involved, we chose a small area shown in figure 6 to be generalised.

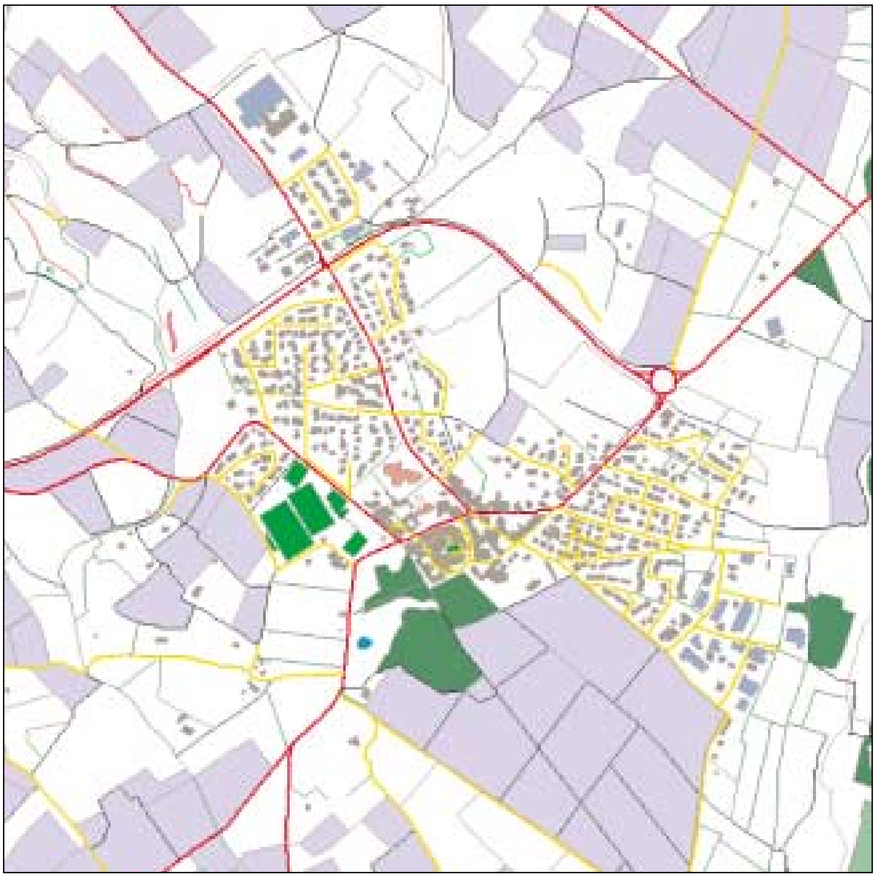


Figure 6 – BDTopo® data to be generalised viewed at 1 : 15.000 (symbolised with Stratège)

This area is a village located near the town of near Montpellier in the south of France. The village contains small houses, a village centre and some industrial buildings and is surrounded by vineyards.

3.2.2 The Required Generalisation

The required generalisation follows the specification of IGN 1 : 50.000 scale maps. These maps are made for military purposes: They must be as descriptive as possible and easily readable even under very difficult conditions (in a tank). For us this kind of generalisation is an interesting study as it tries to preserve the character of the terrain as far as possible without sacrificing legibility. Since there were no clear written specifications, several maps at the scales 1 : 25.000 and 1 : 50.000 were sent as examples of what was requested.

Table 6 – Data content

Data type 1	Data type 2	Nbr of objects	Nbr of points	Length / Area
SYMBOLS	P-crossroad	31		
	Track	6	1313	1 036 m
	River	10	38	586 m
	Line various	16	49	853 m
Road	Road	202	825	14 150 m
	B-road	125	570	8 131 m
	Total Road	*327* ***64%***	*1395*	*22 281 m* ***59%***
	Slope	18	116	1 646 m
	Vegetation boundary	132	527	11 553 m
LINES		**509**	**3438**	**37 955 m**
Buildings	Administrative building	2	50	3 089 m^2
	Industrial building	21	247	57 711 m^2
	Building	473	3244	113 773 m^2
	Total buildings	*496* ***81,6%***	*3541*	*174 573 m^2* ***12,7%***
Vegetation	Agricultural area	88	2143	1 095 260 m^2
	Wood	13	288	86 225 m^2
	Total Vegetation	*101* ***16,6%***	*2431*	*1 181 485 m^2* ***85,6%***
	Hydro area	3	51	919 m^2
	Various areas	8	69	22 491 m^2
AREAS		**608**	**6092**	**1 379 468 m^2**

Figure 7 – Montpellier 1 : 25.000 scale map. Figure 8 – Montpellier 1 : 50.000 scale map.

3.3 The ICC Data Test

3.3.1 The ICC Data

For large scale generalisation, we have chosen the ICC large scale data base

(see http://www.icc.es and http://www.icc.es/cat99/catpub_ingl/topo5.html).

This database is used to produce 1 : 5000 scale map without generalisation.

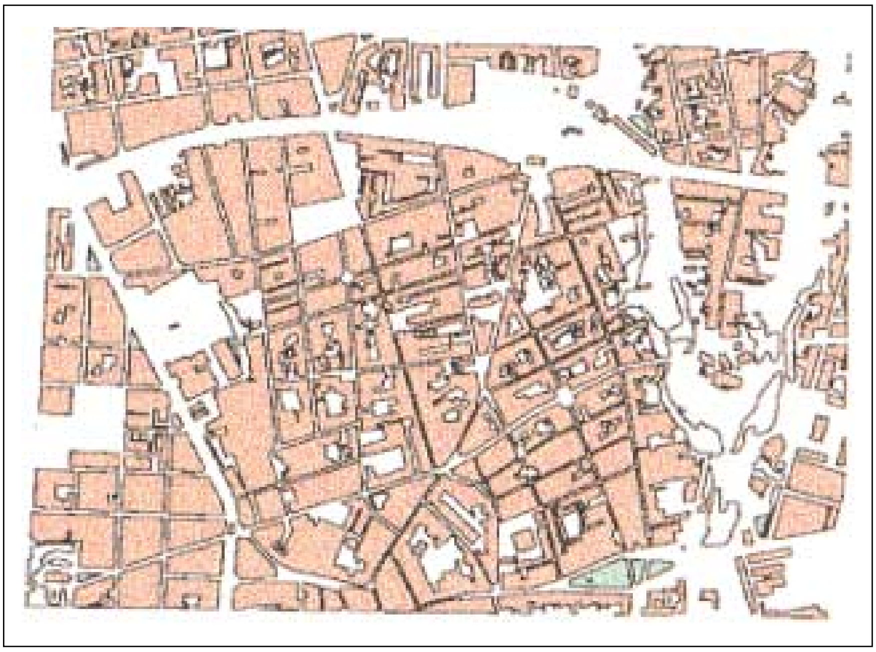


Figure 9 – The ICC initial data to generalise.

The chosen data set contained 1101 objects (around 300 buildings). The database does not explicitly contain streets.

3.3.2 The Required Generalisation

For the test it was asked to generalise only the building layer at 1 : 25.000 scale.

4 The BDCarto® Test

The aim of the first test, described below, is to generalise road network from IGN BDCarto® data (10 meter accuracy) to produce a 1 : 250.000 scale map.

4.1 Participating Institutions

The BDCarto® data corresponds to the first data that partners generalised.

5 tests were performed on different sites of three platforms:

– MGE/MG of Intergraph Corporation. 2 tests were performed
– At the University of Munich by *Monika Jordan*, a student in GIS (now teacher);
– At the ICC by *Maria Pla* and *Baella Blanca*, GIS cartographers in production environment
– Lamps 2 of Laser-Scan Limited. One test was performed by:
– University of Lund by *Lars Harrie*, a Ph.D. student in GIS
– PlaGe a prototype developed by the Cogit Laboratory. 2 tests were performed:
– *Christophe Roux*, a student cartographer at the Cogit laboratory
– *Sébastien Mustière*, a Ph.D. student and cartographer of the Cogit laboratory

4.2 *General Analysis*

We first compared the processes to show or emphasise differences due to platform capacities and users choices, and then presented some global results which correspond to a compilation of processes to see whether there are any global rules. This analysis highlights the main tendencies of the results.

Table 7 – Process comparison

		MGE Munich	MGE ICC	Lamps2 Sweden	PlaGe IGN-1	PlaGe IGN-2	Medium Value
1	Interactive actions	128	65	172	424	466	251
2	Sequences on objects	291	596	172	760	480	460
3	Conflict codes used	3	2	5	17	22	10
4	Operation codes used	2	2	3	7	7	4
5	Algorithm codes used	5	3	2	7	14	6
6	Conflict assessment codes	19	4	6	8	12	10
7	Conflict assessments	72	34	21	28	46	40
8	Evaluation % bad-medium	13-74	0-50	1-11	9-20	7-25	6-36

Lines 1 & 2: We can already see that the number of realised operations differ. This is due to two factors:

1. **strategies:** 1/ ICC realised grouped operations according to object codes and then corrected remaining problems; 2/ Munich realised the same operation on a set of neighbouring objects; 3/ IGN-1 generalised objects one by one and realised some final global improvement operations; 4/ IGN-2 and Sweden generalised objects one by one.
2. **segmentation tool:** PlaGe has a line segmentation tool which incites users to segment a line into sub-lines. Such segmentation offers the possibility of finding a more appropriate solution than a global one.

Lines 3 and 6 show that the use of conflict codes (conflict visually identified) is not homogeneous and more accurate in the evaluation stage (line 6) than in the initial stage (line 3) whenever algorithms are missing. The important use of conflict codes at the IGN is probably due to an internal culture of the generalisation process, where conflict is commonly viewed as a way of triggering an operation.

Lines 4 and 5 represent the range of solutions used to solve problems, **and lines 10 and 12** the main algorithms[1]:

- Only filtering and smoothing were performed with MGE. We can also notice that Munich and the ICC used different algorithms: Munich: **Douglas, Lang, Reuman** as filtering and **Brophy and Weighted Average** as smoothing; ICC: **Point relaxation** as filtering and **Brophy and Simple Average** as smoothing.
- Two algorithms were used with Lamps2: **Douglas-Dispike** for filtering and **Akima-Bicubic** for smoothing. **Douglas-Dispike** represents a kind of 'shape-constrained' filtering: it preserves points around the selected vertices. **Akima-Bicubic** allows granularity conflicts to be solved by mixing smoothing and caricature effects, but visual controls show that in order to solve such conflicts, the shape of the line is very much damaged, and the line geometry is very far from its initial position.
- With PlaGe the number of algorithms used is important, which is not surprising as PlaGe is conceived to test existing algorithms and to develop complementary ones. Moreover, between IGN-1 and IGN-2 the number of algorithms used is different. The main reason is that the IGN-1 test was performed by a student who did not know PlaGe before the test, and the IGN-2 test was performed by a researcher. His knowledge in generalisation and in PlaGe allowed him to use a larger set of algorithms which seems to be difficult for less trained people.

Lines 7 & 8 represent users visual evaluation used after each operation. It is thus very subjective but it shows that even interactively, users are not satisfied with the solution.

Table 8 – Process comparison 2.

		MGE-M	**MGE-I**	**Lamps2-S**	**PlaGe-1**	**PlaGe-2**
9	% objects generalised by one operation only	5 % obj. 6% oper.	32% obj. 19% oper.	27% obj. 56 % oper.	37 % obj. 17 % oper.	67 % obj. 50 % oper.
10	Man Algorithm used for a single process ?	Douglas Or Reuman	Pt Relaxation	Douglas- Dispike	Douglas	Douglas
11	% objects generalised by two operations	38 % obj. 93% oper.	68% obj. 81% oper.	11% obj. 44% oper.	38 % obj. 36 % oper.	9 % obj. 14 % oper.
12	Algorithms mainly used for two processes	1st : Douglas or Reuman 2nd : Brophy or W average	Pt Relaxation 2nd : Simple or Weighted average	1st :Douglas-Dispike 2nd : Akima Bicubic	Gauss and Douglas in different order	Smoothing and Douglas /Lang Nickerson and Douglas
13	% objects generalised by more than two operations	0,2 % obj. 1 % oper.	-	-	25 % obj. 47 % oper.	5 % obj. 36 % oper.

Lines 9, 11 and 13 show that users have different perceptions as to the necessity of line generalisation.

Line 13: Whenever more than three operations are performed, it corresponds to lines which have complex and non-homogeneous geometry. The process used by IGN-1 is mainly: 1/ Segmentation; 2/ Smoothing (Gauss or Lowe) and Caricature (mainly Plaster, but some time for

[1] Main algorithms referred: **Filtering:** Douglas, Lang, Point relaxation, Reuman. **Smoothing:** Brophy, Gauss, Simple or Weighted average, Akima Bicubic. **Caricature:** Lowe, Plaster, Accordion. **Line Displacement:** Nickerson.

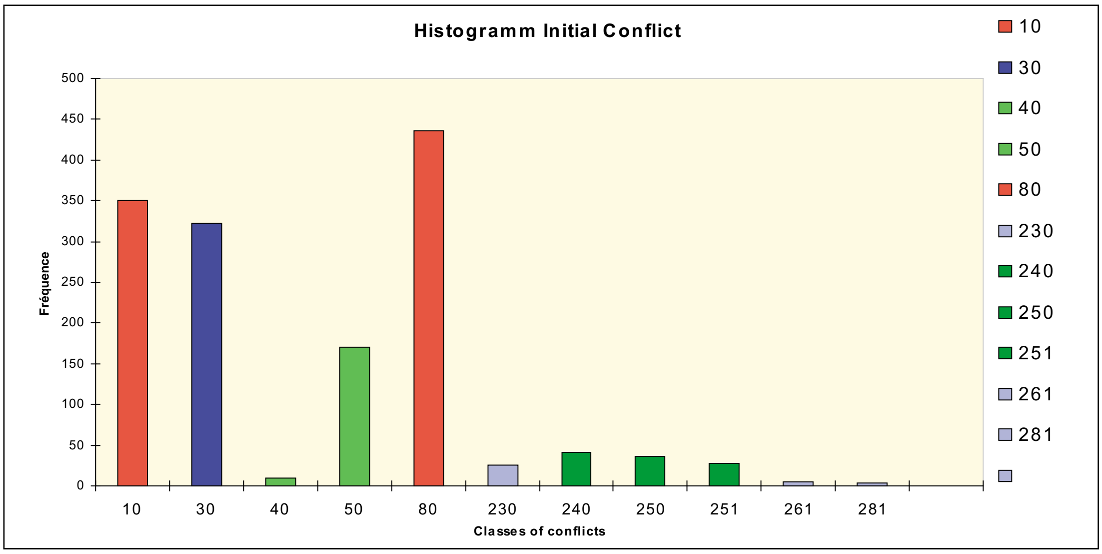
regular bends Accordion); 3/ Filtering with Douglas. IGN-2 used mainly: 1/ Segmentation; 2/ Caricature on complex segments and Smoothing on others; 3/ Fusion; 4/ Smoothing with Gauss; 5/ Filtering with Douglas.

4.3 Accurate Analysis

4.3.1 Conflicts which Motivated an Action

In graph 1 we can see that most of the time the operations were due to:

- the level of detail of the initial data (conflicts 10 and 80)
- the coalescence due to the new road width (40, 50)
- the degradation of the geometry (30: too much angularity) due to the use of filtering algorithms for solving conflicts due to the level of detail.



1 line: 10: too detailed, 30: too angular, 40: to close, 50: self overlapping, 80: too many points
2 lines: 230: too segmented, 240: too close, 250: overlapping, 251: connection to sharp
set of lines: 261: loss of structure, 281: not readable junction

Graph 1 – Initial conflicts.

4.3.2 Relationships between Conflicts and Operations

The set of the following arrays is a synthesis of processes performed by different users. Table 3 shows the correlation between the main conflict visually identified and the operation used to solve it. More accurately, tables 9 and 10 show the correlation between these conflicts and algorithms. Then, table 11 shows a post-visual evaluation of conflicts after each algorithm.

Table 9 – Correlation between initial conflict and best operation to use.

Line conflicts / operation used	Sim	Smo	Caric	Pdis	Ddis	Seg	Fus	Col	Total	%
Line too detailed	**288**	52	6	2	0	1	0	0	349	24,7
Too many points within a line	**433**	1	1	0	0	0	0	0	435	30,8
Shape too angular	0	**323**	0	0	0	0	0	0	323	22,8
Proximity conflict within a line	0	1	7	0	0	1	0	0	9	0,6
Overlapping conflict within a line	18	6	**102**	0	7	**34**	1	0	168	11,9
Line too segmented	0	0	0	0	0	0	**9**	**16**	25	1,8
Proximity conflicts between 2 lines	0	3	0	**11**	**27**	0	0	0	41	2,9
Overlapping between 2 lines	2	0	3	**14**	**16**	0	1	0	36	2,5
Junction too sharp	0	2	1	2	**23**	0	0	0	28	2,0
Total	741	388	120	29	73	36	11	16	**1414**	-
% of this operation / all	52,4	27,4	8,5	2,0	5,2	2,6	0,8	1,1	-	-

Table 9 shows a very strong correlation between conflict and operation. We can just highlight some information:

– Many overlapping conflicts exist within a line (12,5%) and between lines (7,4 %).
– To solve such conflicts caricature (8,5%) and displacement (7,2 %) are often used. Caricature is often preceded by segmentation. These percentages would certainly be higher if corresponding algorithms existed on each platform.
– The conflict "too many points" is hardly distinguished from "too detailed" although it does not represent the same conflict.
– Even if simplification is the main operation used for generalisation, 48% is not simplification but more complex operations, although only road network was represented in this data base.

4.3.3 Relationships between Conflicts and Algorithms

Table 10 and table 11 describe the uses of algorithms for solving conflicts which were identified visually. In the following tables, the algorithm code corresponds to a subset of available algorithms on GIS.

During these tests the following algorithms were used:

– No algorithm: 0
– Shape removal: Schematisation (bend removal) 28
– Simplification: 31: Douglas; 32: Douglas-Dispike; 34: Thapa; 37: Lang; 38: Point relaxation; 39: Reuman-Witkam
– Smoothing: 41: Brophy; 42: Gauss; 44: Simple average; 45: Weighted average; 48: Akima Bicubic
– Caricature: 61: Lowe; 62: Plaster; 71: Balloon; 72: Accordion
– Displacement: 101: Nickerson (line displacement)
– Topological changes: 152: Segmentation; 157: Fusion

Table 10 – Correlation between conflicts within a line and appropriate algorithm.

Conf \ algo	0	28	31	32	34	37	38	39	41	42	44	45	48	61	62	71	72	101	152	157	Tot
Too detailed	1	0	202	82	0	0	6	0	0	45	0	1	0	3	8	0	1	0	0	0	349
Too many pts	0	0	353	0	4	8	0	67	0	2	0	0	0	0	0	0	1	0	0	0	435
Too angular	0	0	0	0	0	0	0	0	69	47	29	101	75	0	2	0	0	0	0	0	323
P proximity	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	4
Bend proxi.	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	2
Bends proxi.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	2
Overlapping	2	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	18
Punctual over	0	0	1	0	0	0	0	0	0	1	0	0	0	2	29	2	2	2	1	0	40
Bend over.	1	0	0	3	0	0	0	0	0	0	0	0	0	0	10	0	4	0	2	0	20
Bends over.	17	1	1	0	0	0	0	0	0	2	0	0	0	3	33	1	20	1	6	0	85
Too segmen.	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	25
First total:	*30*	*1*	*557*	*96*	*4*	*8*	*6*	*67*	*69*	*97*	*29*	*102*	*75*	*9*	*87*	*3*	*30*	*3*	*14*	*16*	*1303*

Table 11 – Correlation between conflicts on a set of lines and appropriate algorithm.

Conf \ algo	**0**	**28**	**31**	**32**	**34**	**37**	**38**	**39**	**41**	**42**	**44**	**45**	**48**	**61**	**62**	**71**	**72**	**101**	**152**	**157**	**Tot**
P Proximity	18	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	5	0	0	26
2 lines prox.	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	12
Lines prox.	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Jct too sharp	5	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	20	0	0	28
Overlap	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	6
2 lines over	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	7	0	0	26
Lines over	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2
Geom. Over	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Second total	48	0	3	0	0	0	0	0	0	4	0	0	0	0	2	2	0	45	0	0	104
Total 4-5	**78**	**1**	**560**	**96**	**4**	**8**	**6**	**67**	**69**	**101**	**29**	**102**	**75**	**9**	**89**	**5**	**30**	**48**	**14**	**16**	**1407**
% algo	**5,5**	**0**	**52,7**						**26,7**					**9,5**				**3,4**	**2,1**		**1407**

– Correlation between conflict identification and algorithms is still strong but not absolute.

– Simplification is used to solve 'too detailed lines' and the main algorithm used is Douglas. It seems that an improved version of Douglas would be appropriate, otherwise (version 31) it creates a 'too angular' line that is solved by smoothing: it explains why around 23% of the results obtained by means of Douglas were qualified as bad or medium (table 12). Moreover, results are often qualified as being 'too far from initial geometry' (graph 2).

– Smoothing algorithms are often used (27% of algorithms) without preference (GISs have different smoothing algorithms). Except at the IGN, smoothing is used after simplification, i.e. to correct the effect due to filtering.

– Caricature algorithms are mainly used to solve internal proximity conflicts. As only the IGN has caricature algorithms, the number of algorithms used is under-evaluated (20 cases of missing algorithm in table 12, column 0). At the IGN, Plaster is the favourite algorithm except for 'set of bends overlapping' where Accordion is also used.

– Algorithms are missing for 46% of the conflicts between different lines (table 11). Such conflicts are solved mainly by line displacement (Nickerson) which seems to give around 54% of appropriate results (table 12).

4.3.4 Algorithm Assessment

After each action the user was asked to add a visual assessment to describe if the proposed solutions were good, average or bad. The following table synthesises these visual related to each algorithm used (algorithm code is given above).

Table 12 – Visual evaluation of solutions after generalisation.

		F	**I**	**L**	**T**	**E**	**R**	**S**	**S**	**M**	**O**	**O**	**T**	**C**	**A**	**R**	**I**	**D**	**S**	**F**	**tot**
Algo.	**0**	**28**	**31**	**32**	**34**	**37**	**38**	**39**	**41**	**42**	**44**	**45**	**48**	**61**	**62**	**71**	**72**	**101**	**152**	**157**	**-**
Bad	15	1	6	2	0	0	0	2	15	3	0	18	0	4	13	0	8	10	4	0	**101**
Medium	26	0	128	10	2	5	4	63	26	15	11	67	9	3	41	2	20	15	7	0	**454**
Good	38	0	427	84	2	4	2	2	28	85	18	17	66	2	35	3	2	29	3	12	**859**
Total	79	1	561	96	4	9	6	67	69	103	29	102	75	9	89	5	30	54	14	12	**1414**

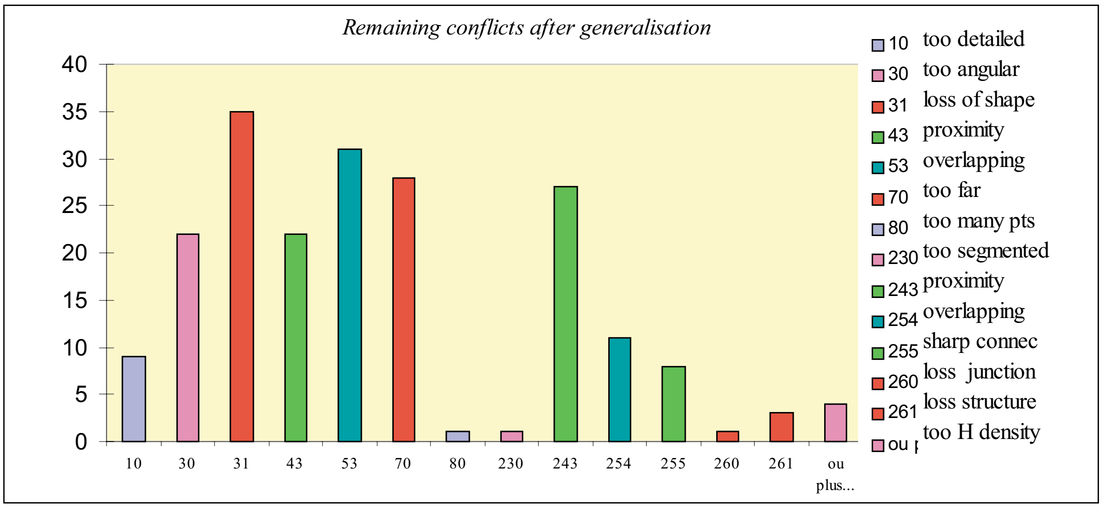
Validation is subjective, therefore these numbers should be regarded with caution. Nevertheless, we can see what kind of algorithms were more or less appreciated.

- 39% of the algorithms did not provide full satisfaction although parameters were tuned interactively.
- **Simplification:**
  - Douglas (31) is the most used but 24% were criticised. Douglas-Dispike (32) is used less but appreciated more.
  - Reuman-Witkam (39) was correct but not perfect.
- **Smoothing:**
  - Platforms used different algorithms, therefore comparison is difficult.
  - In 43% of the results after a smoothing is not exactly appropriate. Among critics, it seems that either angularities remain or lines lose their shapes or some proximity conflicts within a line remain.
- **Caricature:**
  - Caricature was only used by the IGN, but it constitutes 31% of the interactive operation for the IGN. Plaster and Accordion (algorithm developed for generalisation purposes) were used the most.
  - 68 % of the proposed solutions are either 'bad' or 'medium': this shows that such algorithms are useful but research must be continued in order to obtain better results.
- **Displacement:**
  - Only one algorithm was used (Nickerson). Half were criticised but the criticisms were rarely explained.

4.3.5 Remaining Conflicts

After each line generalisation, the user was asked to identify any remaining conflicts. Graph 2 synthesises remaining global conflicts. Some new conflicts appear: 31: loss of shape (17%); 70: too far from initial geometry (14%). However, some conflicts are not solved, especially overlapping within lines (26%) and between lines (23%). If we compare different generalisations, conflict assessment histograms look different: at the IGN the main final assessments are 'loss of shape' (22%) and 'too far from initial geometry' (39%), whereas at Munich, the main final conflicts are overlapping within line (37%) and between lines (37%). Even if such an evaluation is highly subjective, a comparison of proportion is of great interest.

An hypothesis of such diversity of final assessment code repartition is that as long as some unacceptable conflicts such as overlapping remain, we do not focus on other aspects (such as shape and accuracy). Actually, we should distinguish between conflicts which are due to under-generalisation (such as visibility constraints violation: overlapping) and an evaluation of acceptable but not perfect solutions where shape and accuracy will play a major role.



Graph 2 – Remaining conflicts after generalisation.

4.3.6 Parameter Values

Analysing the distribution of parameter values helps mainly to identify the range of possible values used for one algorithm for a specific generalisation. Accurate parameter values for line generalisation can be analysed only if some measures on line characteristics are added (such as shape, coalescence, level of detail). The computation of average values is only meaningful when the set of values is large enough. We will hereafter only analyse Douglas parameter values (frequently used in three tests) and two smoothing parameter values (also frequently used).

Filtering

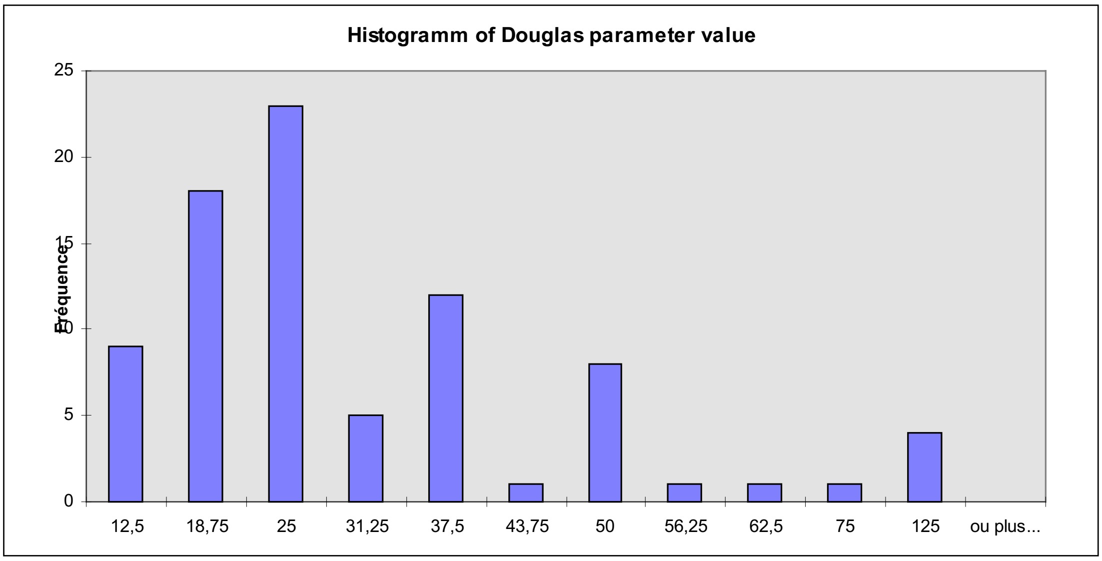
We analysed the use of Douglas parameter values in three tests:

Table 13 – Use of Douglas parameter values.

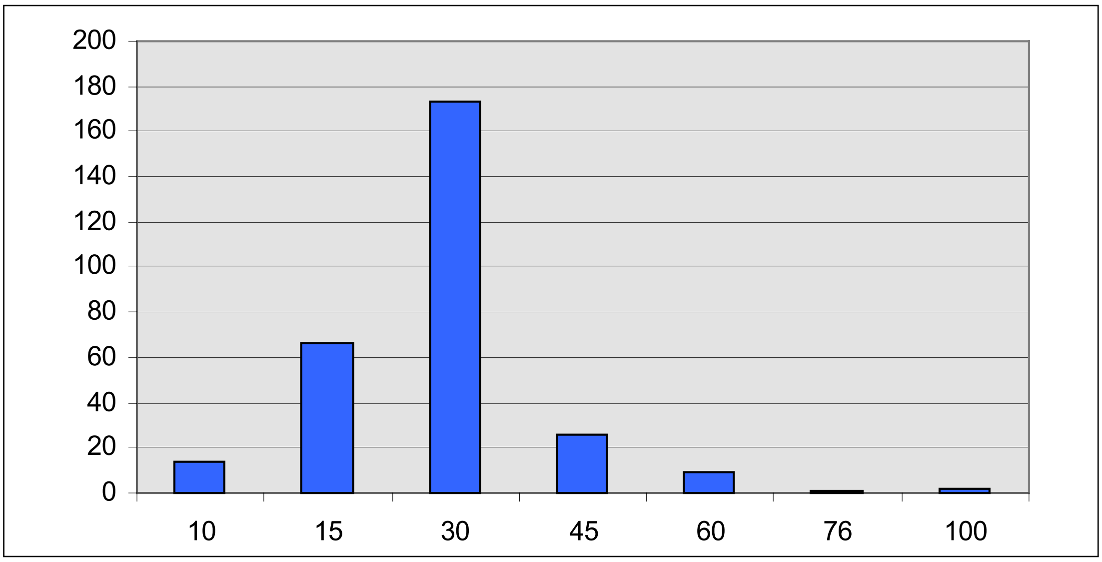
Test	Number of times	Average step between values	Average value	More used value (mode)
MGE/MG	83	6,25 meter	33,3 meter	25 meter
IGN-1	291	15 meter	28,5 meter	30 meter
IGN-2	88	15 meter	22,6 meter	15 meter

We can notice that the values are quiet homogeneous and are corresponding to the minimum level of detail at final scale: 28 meter corresponds to 0.1mm at 1/250 k on the map. The average step between the values often corresponds to the interface capacity.

We also notice on the following histogram that the values are grouped around the average value but that in some cases bigger values are used. The problem – noticed before – is that important values deteriorate line shape property, making them too angular (conflict number 30).



Graph 3 – Distribution of Douglas parameter value (in field meter) on MGE/MG.



Graph 4 – Distribution of Douglas parameter value in field meter (Plage).

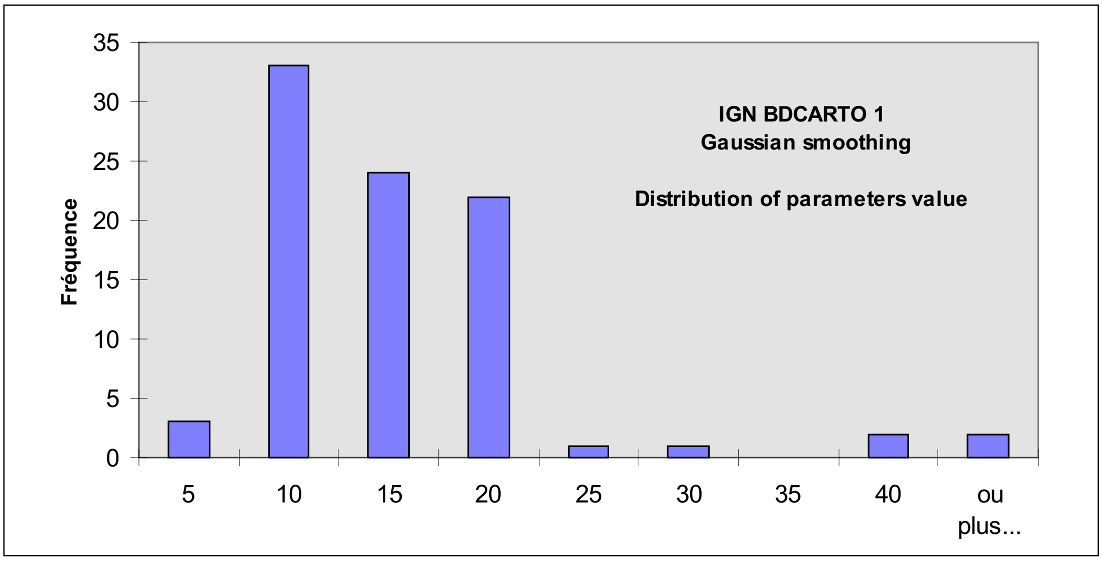
Smoothing

A smoothing generates a line whose curvature is more regular then the initial one. Two methods are usually used. Either the smoothing computes the position of a point according to the position of its neighbouring points (e.g. Gaussian filtering, weighted average) or a new line is computed from the initial points by means of curved line interpolations (e.g. spline, akima bicubic)

Setting the values of the smoothing automatically is difficult because the number of points may describe one shape (a bend) or a set of shapes changing the final shape of a line essentially. Whatever smoothing, it is difficult to choose the right number of points because of the heterogeneity of the line geometry.

We analysed the Gaussian parameter values used at the IGN (graph 5) this algorithm the parameters represents the curvilinear distance considered around each point to smooth. In the algorithm this distance represents the standard deviation of the gaussian curve used to weight points.

We can notice that the values are grouped. he average value is 15,7 meters and the more frequently used is 10 meters.

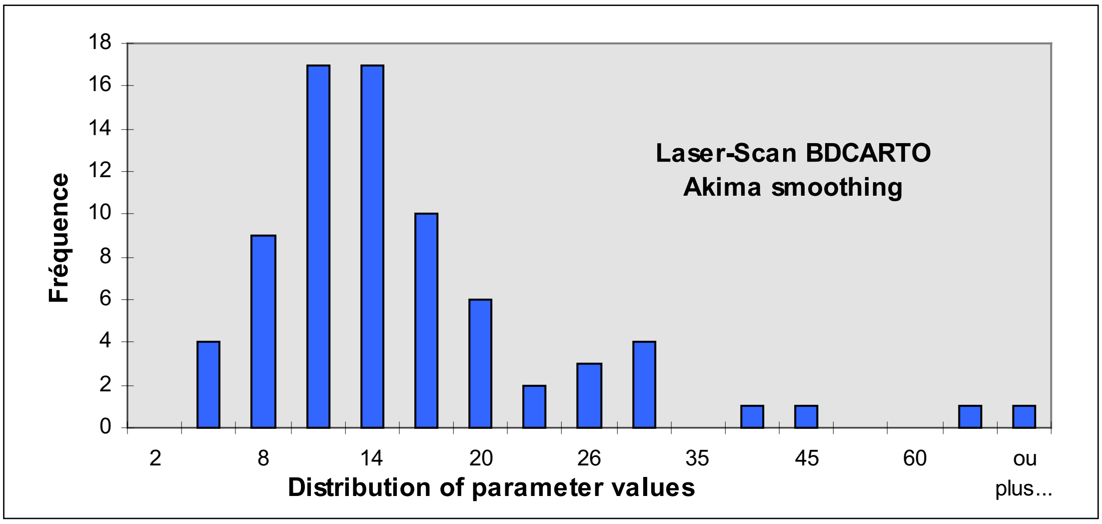


Graph 5 – Distribution of Gaussian smoothing parameter value on Plage.

Laser-Scans smoothing (akima bicubic) generates a set of cubic curves from a set of points. The result is a smoothed line. The advantage of curve interpolation in general is that the line is really curved, the drawback is that the distance between the initial line and the smoothed line can be very important, degrading the line accuracy. Therefore the parameters of Laser-Scans smoothing represents the maximum distance after interpolation between the initial line and the interpolated line. The average value is 15,9 meters, the more frequent one is 10 meters. Here again, the value is under the accuracy distance (i.e. 25 m, 0.1 mm at the final scale). When the chosen parameter values are important they create an accuracy conflict (70) or a proximity conflict with other lines (241).

4.4 Visual Results

Figure 10 shows the five generalisation results. We can see that the generalisation performed on the same platforms look globally the same. The differences are due the the level of generalisation. The ICC and Lund performed stronger generalisation than the others, the IGN tried to preserved as much as possible the initial shape properties, maintaining more roads bends. The lund generalisation performed too strong smoothing that sometimes degraded line shape geometry.



Graph 6 – Distribution of Akima smoothing parameter values.

Figure 10 shows the five generalisation results. We can see that the generalisation performed on the same platforms look globally the same. The differences are due the the level of generalisation. The ICC and Lund performed stronger generalisation than the others, the IGN tried to preserved as much as possible the initial shape properties, maintaining more roads bends. The lund generalisation performed too strong smoothing that sometimes degraded line shape geometry.

4.5 Conclusions

Among the large set of information, we can state some important results:

- The high correlation between conflicts and operators shows, that automation is possible if we work on conflict detection and qualification.
- Correlation between conflicts and algorithms: its seems that the distinction between conflicts should allow a better automation, but the shape of objects, which is not considered, should be taken into account.
- For simplification, Douglas is frequently used but requires a smoothing as post-processing to correct the too angular shapes it gives to the objects. An improved version such as Douglas-Dispike would be very useful.
- A large range of smoothing algorithms are used, but they do not render exactly what users would like.
- Important algorithms such as caricature and line displacement are often missing. For the existing ones, some research must be carried on to improve their behaviour,
- For complex lines, a segmentation is useful to apply an appropriate algorithm on each segment, particularly for caricature,

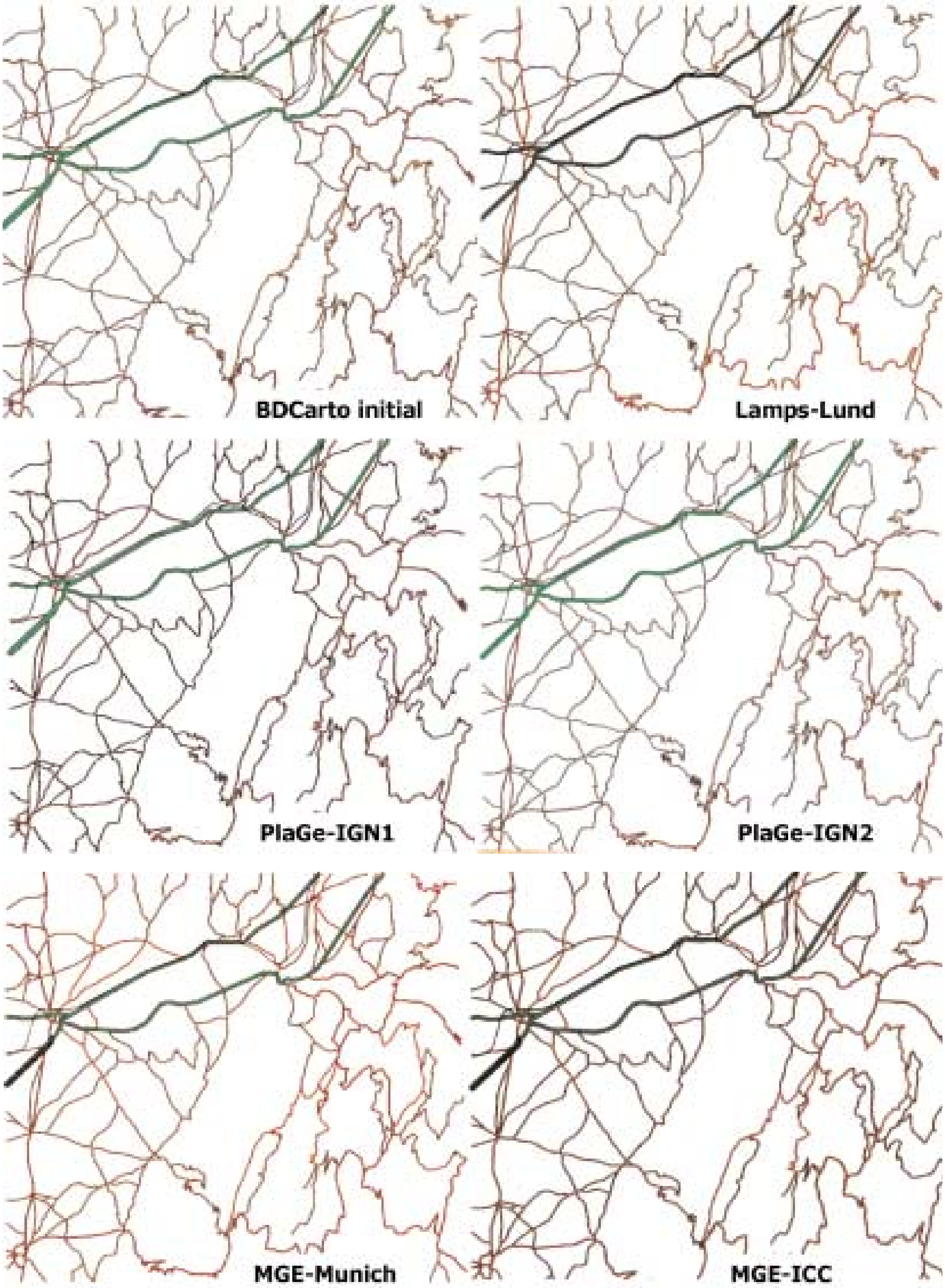


Figure 10 – BDCarto® generalised road.

– To distinguish between different solutions, some characteristic parameters, qualifying accuracy and shape distortion would be useful: Such properties are visible but hard to describe automatically.

5 The BDTopo® Test

The aim of the second test, described below, is to generalise BDTopo® data (1 meter accuracy) to produce a 1 : 50.000 scale map.

5.1 Participating Institutions

This test was performed on different GISs by different users:

– At LUND University, Sweden by *Lars Harrie* on LAMPS2
– At the ICC Barcelona by *Maria Pla* and *Blanca Baella* on their own production system using CHANGE, MGE and some internal algorithms
– At the Technical University of Munich TUM by *Monika Jordan* on MGE
– At the IGN France by *Annabelle Boffet* on LAMPS2 and by *Anne Ruas* on Stratège platform
– Moreover, the University of Hanover (*Brigitte Husen*) sent generalised data on CHANGE and the University of Glamorgan, Wales GB (*Chris Jones* and *Mark Ware*) sent the generalisation of three areas performed on MAGE. Unfortunately these processes could not be compared to others and are not presented hereafter.

The results of the tests depend on the operators' choices and abilities as well as the system used. Direct comparison of the results is therefore difficult and must be treated with caution.

5.2 Analysis of the Results

Due to the important difference between the different tools available in the used GIS generalisation packages, it was not possible to synthesise all the generalisation and to present global results as I did for the BDCarto® test. Consequently, the tests are presented one after the other and described in the following scheme:

1. **Global result:** number of operations, quantity of visual and subjective assessments
2. **Operations and algorithms:** the quantity and percentage of operations and algorithms used
3. **The process:** qualitative analysis of the sequence of operations according to working areas and object type. The use of algorithms according to object type.
4. **The conflicts:** quantitative analysis of relations between conflict code and solution. Analysis of remaining conflicts.
5. **The use of algorithms:** Subjective assessment of algorithms and parameter values analysis where possible.
6. **Presentation of cartographic results.**

These presentations are followed by a chapter comparing the different tests.

Before reading the test analysis, the reader should be reminded of three points:

1. The numbers presented in the tables are not always coherent to one another because the users did not fill out all boxes of the process template.

2. Even if some symbol specifications were given, the users did not use the same thresholds (road width and building size). This does not change the conclusions of the test since its aim is to understand the process and the difficulties and not to compare the results.
3. Graphical results are not very satisfactory due to the fact that the principle of the test was to use available algorithms only and to avoid interactive point by point editing. The main goal was to highlight problems, not to correct them. **It should be clear for the reader that each of these platforms can provide excellent generalisation results in a production line with more interactive actions.**

5.3 Test 1: Generalisation on MGE/MG

This test was performed by *Monika Jordan* from the Technical University of Munich. The criteria are applied very strictly. The results of the operations are often considered medium or bad where other, less critical assessors might consider them adequate. For test purposes this is very useful as it allows better identification of problems.

5.3.1 Global Results

- Number of operations 405 operations divided into
  - 268 (66 %) with a generalisation algorithm
  - 135 (34 %) interactive object removal and fusion
  - 2 (0.5 %) dilations on Microstation
- All operations are triggered interactively. One operation is triggered on one or several objects.
- Visual subjective satisfaction on 48 regions (by the user):
  - 5 are well generalised 10,4%
  - 17 are accepted 35,4%
  - 26 are badly generalised 54,2 %

5.3.2 Operations and Algorithms Used

11 MGE algorithms were used:

- Two algorithms to aggregate disjoint polygons. One is assigned to regular shapes.
- **Boundary extent algorithm** to extend the geometry of a line or polygon to connect different geometry
- Two typification algorithms: transforming a set of symbols into one symbol
  - **Point typification** locates the final symbol to an existing position
  - **Centroid typification** locates the symbol at the centre of the cluster
- Two orientations on polygons:
  - **Point/Line orientation** orients a symbol according to the orientation of its neighbouring line
  - **Orientation** rotates a symbol with respect to a given angle
- Two line simplification algorithms:
  - **Douglas** general line simplification
  - **Clarification** to simplify regular shapes such as buildings

– One **squaring algorithm**
– One collapse **area-to-point** algorithm which transforms a polygon into a square symbol.

Neither smoothing nor displacement are used.

Table 14 – Operations and algorithms used (test 1).

Operation	**Number**	**%**	**Algorithm**	**Number**	**%**
Object removal	111	**27,5**	-		
Aggregation	45	11,1	Ortho-aggregation	8	3
			Aggregation	8	3
			Boundary extent	29	10,8%
Area Fusion	24	5,9	-		
Symbol Typification	71	17,5	Point typification	2	0,7
			Centroid typification	69	**25,7**
Chgt Orientation	59	14,6	Point / Line orientation	35	13,1
			Orientation	24	9
Simplification	12	3	Douglas	2	0,7
			Clarification (area)	10	3,7
Squaring	7	1,7	Area squaring	7	2,6
Polygon dilation	2	0,5	-		
Collapse	74	18,3	Area to point	74	**27,6**
Total	**405**	**100**		**268**	**100%**

Table 14 analysis

– Most of the operations concern the reduction of the quantity of information:
  – Object removal, symbol typification, area fusion, aggregation => 62 %
– The typification algorithm works only on symbols. This explains why a lot of collapse operations (from area to symbol) were performed on buildings to be able to use the typification algorithm.
– Very few operations are performed on the object geometry:
  – Simplification, squaring => 4,7 %

5.3.3 Operations on Objects

5.3.3.1 Repetitive Sequences on a Region

This test was performed region by region.

The sequence used for built-up areas in this test was as follows:

1. Interactive removal of objects as small buildings and streets.
2. Thematic and large building generalisation:
   2.1. Simplification (clarification algorithm) (or aggregation).
   2.2. Squaring.
3. LOOP:
   3.1. Transformation of polygon buildings into symbol buildings: Collapse – area-to-point. Medium size buildings were also enlarged during this collapse process .
   3.2. Typification of symbol buildings: most of the time it is performed on two buildings, replacing them with one.
   3.3. Orientation of buildings along a line (mainly streets) (P/L orientation).

4. Interactive object removals: hedges, small buildings. Certain small overlaps between unimportant lines and symbols were solved by eliminating the line.

The sequence used for vegetation areas in this test was mainly interactive:

96 sequences of line or polygon removals were performed interactively.

The only generalisation algorithms used were:

1. 8 aggregations between areas
2. 29 'boundary extend' to remove space between two thematic areas.
3. 2 area simplifications (clarification).

5.3.3.2 Relation Between Object Type and Algorithms:

Table 15 – Algorithms used for each object type (test 1).

		Building polygon		Building symbol		Land area		Land boundary		Street		Mixed lines		Polygons	
Clarification		8				2									
Area squaring		7													
Ortho-aggregation		7												1	
Area-to-point		**73**												1	
Centroid typific.				**69**											
P/ L orientation				**34**										1	
Orientation		1		**23**											
Pt typification				2											
Aggregation						8									
Bound. Extend						1		1				**27**			
Douglas										1		1			
Sum	**%**	95	**35,4**	129	**48,1**	11	4,1	1	0,4	1	0,4	28	10,5	3	1,1
Interactive operations															
Object removal		19		7		6		**23**		7		**46**		3	
Fusion						**24**									
Dilation		2													
All															
Sum	**%**	116	28,6	136	33,6	41	10,1	24	5,9	8	2,0	74	18,3	6	1,5

Table 15 analysis

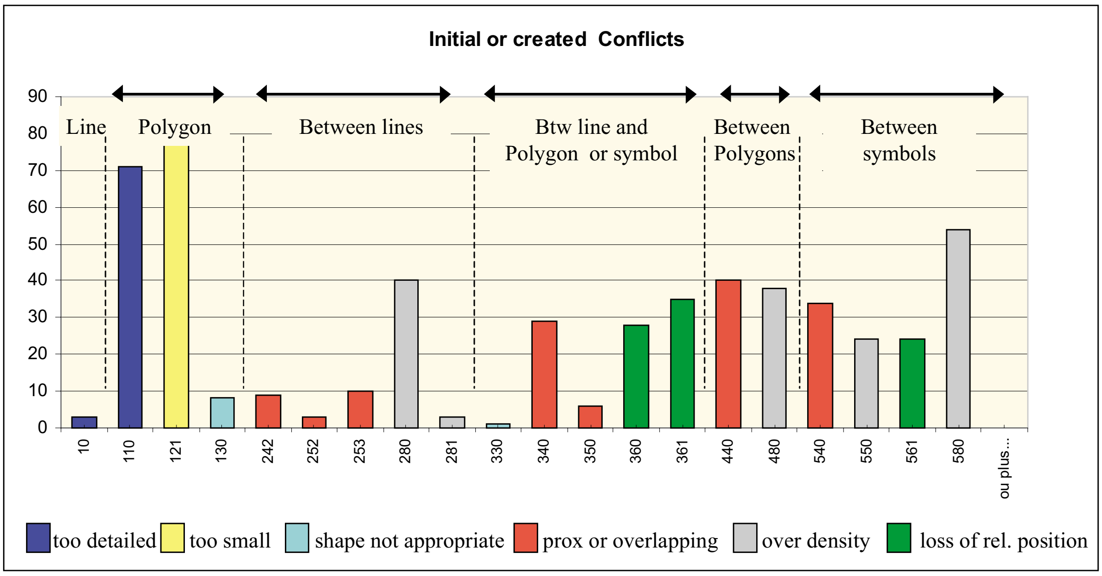
– Most of the generalisation algorithms are performed on buildings: area-to-point; typification (i.e. object removal and displacement of the selected object); orientation.

– Very few operations are performed on streets.

– Vegetation areas are generalised without algorithms. 39% of the interactive deletion operations are performed on vegetation only.

5.3.4 Conflicts and Conflict Resolution

5.3.4.1 Detected Conflicts



Graph 7 – Distribution of conflicts (test 1).

Graph 7 analysis

- Polygons (buildings) are too small and too detailed (96 initial buildings, conflicts 110 and 121),
- Most of the conflicts are related to over proximity, overlapping and over density (242, 252, 253, 280, 281, 340, 350, 440, 480, 540, 550, 580)
- During the process and as a result of the change of polygons into symbols, conflicts between polygons were transformed into conflicts between symbols.
- The area-to-point operation affects the relative position of lines and symbols (360-361) and the loss of the relative symbol position (561).

If we classify the conflict by object character, we obtain the following:

- Over granularity 14,5 %
- Size (too small) 17,2 %
- Bad shape 3 %
- Over proximity 25,7 %
- Over density 7,5 %
- Loss of relative position 32,2 %

Table 16 – Operations used to solve conflicts (test 1).

	operation used to solve conflict									
conflict	simplification	collapse	dilation	squaring	Obj removal	aggregation	orientation	fusion	typification	Total
line too detailed	2									2
Polygon too detailed	10	27								37
Polygon too small		46	2		28					76
Polygon not square		1		7						8
Lines proximity					2					2
2 L. long proximity					7					7
Set of lines proximity					3					3
Lines overlaping					6					6
Lines over density					32					32
not readable junction					2					2
Line &Polygon shape					1					1
Line + (P/S) proximity					5	1			4	10
Line + (P/S) overlapping					3				1	4
L + (P/S) orientation					1	27				28
L + (P/S) loss orientation							35			35
Polygons too close					4	10		24		38
Polygons over density					5	7				12
Symbols too close									32	32
Symbols overlapping					1				21	22
Symbols bad alignement							24			24
Symbol over density					2				13	15
Total	12	74	2	7	102	45	59	24	71	396

Table 17 – Algorithms used to solved conflicts (test 1).

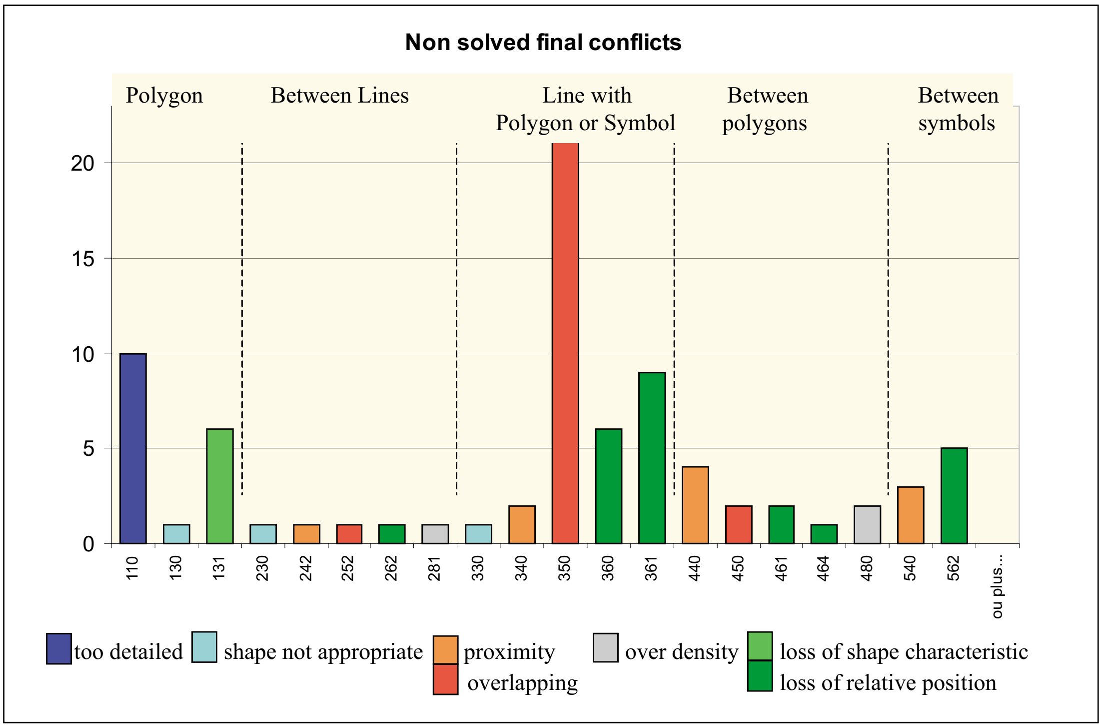
NB conf assess	algorithm code										
conf assess	21	23	31	81	91	96	122	131	132	143	156	Total
Non Satisfaction							1				4	5
Medium Satisfaction		1		1	7	3	5		1	3	12	33
OK	2	68	2	6	1	5	4	35	23	71	12	229
Total	2	69	2	7	8	8	10	35	24	74	28	267

21: Pt typification; 23: C typification; 31: Douglas; 81: area-squaring; 91: ortho-aggregation; 96: aggregation; 122: clarification; 131: S/L orientation; 132: self orientation; 143: area-to-point; 156: boundary extend

Table 16 and table 17 analysis

- Object removal is an answer to several kinds of conflicts.
- Associations between conflict type and solution are very logical
  - e.g. Too detailed => simplification
  - e.g. Too small => removal or dilation or collapse including scaling.
- Most of the algorithms are used for one or two kinds of conflicts only, and vice versa:
  - Polygon too detailed is solved by 1/ collapse or 2/ dilation
  - Collapse is used for 1/ polygon too detailed or 2/ polygon too small
  - Typification is used for 1/ proximity, 2/ overlapping and 3/ over density
- In the absence of a displacement algorithm, proximity and overlapping conflicts between lines or between a line and a polygon or a symbol are often solved by an object removal.
- Proximity, overlapping and density-related problems are solved by an aggregation for polygons and a typification for symbols.

5.3.4.2 Final Assessments



Graph 8 – Remaining conflicts (test 1) (Title of the graph should be identical !).

The user described all remaining conflicts on the 48 regions. Some regions have more than one conflict.

Graph 8 analysis

Remaining conflicts were described for a region and not for each object.

- 80 conflicts were detected (a lot for such a small area)
- Some buildings remained too detailed (110)
- Building shape was not always well maintained (131)
- Overlapping conflicts between lines and symbol or polygon could not be solved with existing algorithms (340, 350)
- The orientation algorithm was not sufficient to maintain the relative position of objects (360 361 461 562)

5.3.5 Use of Algorithms

Table 18 – Algorithm assessment (test 1).

NB conf assess	algorithm code											
conf assess	21	23	31	81	91	96	122	131	132	143	156	Total
Non Satisfaction							1				4	5
Medium Satisfaction		1		1	7	3	5		1	3	12	33
OK	2	68	2	6	1	5	4	35	23	71	12	229
Total	2	69	2	7	8	8	10	35	24	74	28	267

21: Pt typification; 23: C typification; 31: Douglas; 81: area-squaring; 91: ortho-aggregation; 96: Aggregation; 122: clarification; 131: S/L orientation; 132: self orientation; 143: area-to-point; 156: boundary extend

Table 18 analysis

* All the algorithms as well as their parameter values are chosen interactively. However, some algorithms provide results which are not visually assessed as 'good'. We can notice that these algorithms concern **geometric transformation of areas**

- Aggregation: code 91 orthogonal and 96 non orthogonal
- Shape simplification: code 122 clarifications
- Boundary extend: code 156

In the following we are analysing the distribution of parameter values to check if some tendency can be found.

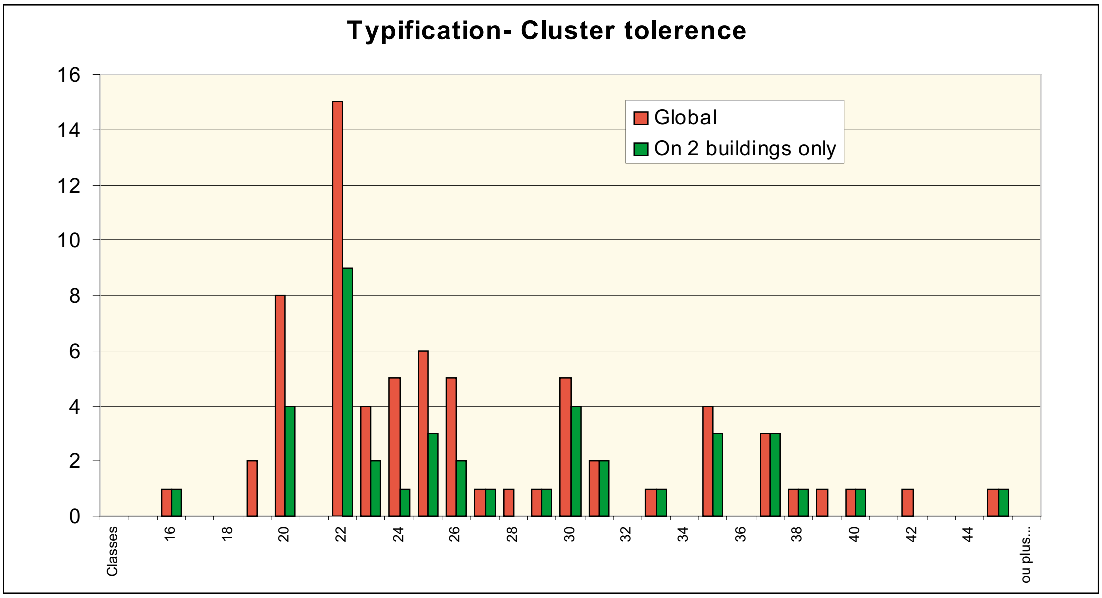
5.3.5.1 Cluster Typification

Cluster typification replaces a set of objects by one object, located in the center of the original group. The parameter value represents the minimum distance between objects to be typified.

Table 19 – Use of typification algorithm on a set of buildings.

Nbr of buildings / typification.	2	3	4	6	10	More
Number of uses	40	13	3	1	1	9

- Cluster typification is used on 2 and 3 objects (60% resp. 19,4 %). The algorithm is therefore used to replace two objects by one, located between both, and not to reduce a set of objects to another set.



Graph 9 – Distribution of cluster tolerance value (test 1).

Graph 9 analysis

- Large range of values; no regularity; high dispersion
- No link between the quantity of objects and the parameter values

The following table represents the default parameters shown on the Excel spread-sheet 'descriptive statistic' (Kurstosis qualifies the flatness of a distribution). For comparison purposes it is also possible to normalise the data (between 0 and 1). This table is given as an example but in the rest of this report only the most common parameters are evaluated in the same way. We should also remember that most of the parameters are significant only if sufficient data are available.

Table 20 – More accurate analysis of parameter value distribution.

Statistic parameter		Normalised
Average	26,38	0,36
Error	0,77	0,03
Median	24	0,28
Mode (More frequent value)	22	0,21
Standard Deviation	6,40	0,22
Variance	40,94	0,05
Kurstosis	0,30	0,30
Asymmetry	1,03	1,03
Max - Min	29	1
Minimum	16	0
Maximum	45	1
Number of values	69	69

5.3.5.2 Area Clarification

The 'area clarification' algorithm aims to simplify building shapes.
It is only used 10 times, always on thematic or large buildings.

The parameters are: Length threshold
Minimum area threshold

Table 21 – Clarification parameters.

	Length threshold		*Minimum area threshold*	
		Normalised		Normalised
Average	10,8	0,45	209,1	0,22
Median	11	0,46	133,5	0,13
More frequent	11	0,46	122	0,11
Standard deviation	3,29	0,25	244,78	0,29
Asymmetry	0,62	0,62	2,64	2,64
Max-min	13	1	844	1
Minimum	5	0	26	0
Maximum	18	1	870	1
Number of data	10	10	10	10

Table 21 analysis

- This algorithm has two possible behaviours: contour simplification or transformation into a rectangle according to both parameter values. This, however, is not very satisfactory (see table 18).
- The length threshold value is coherent to the final scale (around 10m, i.e. 0.2 mm). Moreover, the dispersion of values is low.
- The area threshold parameter is used to replace the polygon by a rectangle if the required value is larger than the initial area. If this parameter value is low, it is nearly inactive and only the first parameter is significant. We can notice that the average value (209) and the median value (136) are under the minimum building size (between 260 and 300m^2) for this scale.

5.3.5.3 Area Aggregation

Two algorithms were used to aggregate areas:

- Ortho-aggregation for buildings 8 times
- Aggregation for vegetation 8 times

These algorithms have two parameters:

- P1: Threshold tolerance to select buildings close enough to be aggregated
- P2: Zone tolerance the maximum length of connection between the two aggregated areas.

For buildings	P1	Average = 4,9	Minimum = 1,5	Maximum = 12
	P2	Average = 35,7	Minimum = 8	Maximum = 74
For areas	P1	Average = 57,9	Minimum = 12	Maximum = 170
	P2	Average = 110	Minimum = 35	Maximum = 220

We can first notice that tolerances are smaller and less extended for buildings, which is logical. However, the difference between P1 and P2 for one aggregation is very high.

P1	P2	Assessment
12	74	1
2,8	45	1
2,4	10	1
1,5	8	1
6,4	20	1
6,4	51	1
2	70	1
5,5	8	2

It is interesting to notice the difference between P1 and P2 values. Normally, these values should be nearly the same but, to allow for aggregation, the user has been obliged to allow larger P2 values.

Consequently, it is not surprising to see the assessment value always at 1 except when both parameter values are close.

– The problem is certainly due to the fact that aggregation between buildings should not use only initial object coordinates. The relative position and orientation between buildings can generate long distances between initial points even if their boundaries are close. Consequently, parameter P2 is not significant. The user increases the parameter value until the aggregation is accepted.

5.3.6 Visual Results

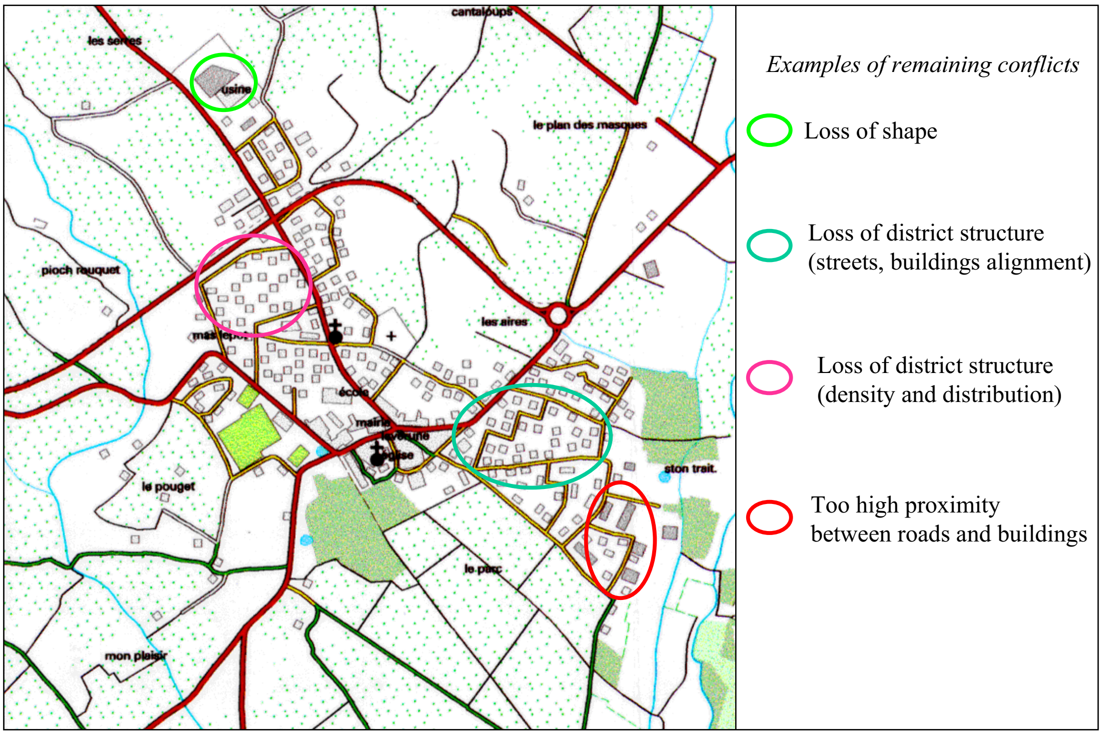


Figure 11 – Visual result (test 1).

Figure 11 analysis

- In order to allow the generalisation of buildings (which represent 43% of objects), the best solution – with the available tools – was to collapse them into symbols, to typify them and to correct their orientation. As a result, the buildings are large enough and far enough apart to be recognised, but their pattern (Gestalt) is often lost and creates a feeling of random distribution. On the other side, the interactive choice of building typification respects successfully the relative distribution of building density between areas.
- For the street network, the connectivity is well maintained though the final network does not preserve the shortest path.
- Another problem is related to the proximity between buildings and streets, which is not treated here.

5.4 Test 2: Generalisation on LAMPS2 – Lund

This test was performed by *Lars Harrie* from the University of Lund (Sweden) on LAMPS2 GIS [*Harrie* 98 47].

5.4.1 Global Results

- Number of operations 122, divided into:
  - 89 (73%) with a generalisation algorithm
  - 33 (27 %) are interactive object removal
- All operations are triggered interactively.
- One operation is triggered generally on several objects. We tried to understand the number of operations on each object according to the templates (to ungroup operations):
  1. Around 500 operations were performed on single objects
  2. Around 130 operations are interactive object removals.
  3. 3 operations are performed on a large set of objects:
     - classification to change the scheme and remove some non-important objects immediately.
     - small building removals: around 70 buildings were removed automatically.
     - area cluster is performed once on all vegetation areas (woods, vineyard, orchard).

Small building removal was performed by an automated selection of all buildings < 70m^2 and their automatic removal without visual checking.

- Visual subjective satisfaction on 122 operations are:

62 are well generalised	50,8%
60 are accepted	49,2 %

5.4.2 Operation and Algorithms Used

7 generalisation algorithms were used but in reality they represent 6 different algorithms.

- **Squaring** is the ‘artificial simplification’ algorithm with a length tolerance value at 0 and an angle tolerance at 30 degrees. In such a way it behaves as a squaring.
- **Exaggerate** results in a dilation for polygons.
- **Corner flipping** is an ‘artificial simplification’ algorithm with 3 parameter values, which changes according to the building.

- **Douglas** is a Douglas-Dispike algorithm with an angle constraint.
- **Aggregation** used in this test is a convex hull. The parameter value is the cluster distance to aggregate objects. It can be performed on an area and generates different aggregations according to the chosen cluster tolerance value. It is often described as a typification by the user.
- **Merge** algorithm aggregates two polygons, even if they overlap. The generated shape is not a convex hull as produced with the aggregation algorithm.
- **Classification** assigns objects to a feature class. At the beginning of the process, it was used to automatically remove some objects: tracks, walls, hedges, fences, ruins, etc. Later it was used to change one road class.

Table 22 – Association between operation and algorithm (test 2).

Operation	squaring	exaggerate	corner-flipping	douglas	aggregation	merge	int. removal	classification	Total	%
squaring	15								15	12,3%
caricature		31							31	25,4%
simplification			12	6			1		19	15,6%
typification					10	1			11	9,0%
aggregation					2	7			9	7,4%
fusion					1				1	0,8%
removal							32	1	33	27,0%
selection								2	2	1,6%
classification								1	1	0,8%
Total	15	31	12	6	13	8	33	4	122	
%	12,3%	25,4%	9,8%	4,9%	10,7%	6,6%	27,0%	3,3%		100%

Table 22 analysis

- A displacement algorithm, available on the platform, was not used as it does not produce appropriate results. This algorithm displaces object coordinates in a decreasing way from a point. Consequently, road and building shapes are locally changed and building size decreases.
- Types of operation:
  - 27,9% of the operations were an improvement of the geometry (simplification, squaring)
  - 25,4% were size growing operations
  - 17,3% were aggregations
- Aggregation is often used as a typification, and often it does not satisfy the user.
- Dilations are described as exaggerations.

5.4.3 Operations on Objects

5.4.3.1 Main Process on Objects

The process began with three grouped operations, then each of the 48 areas were generalised in numerical order.

1. A global classification automatically removed secondary objects.
2. Small buildings (< 70 m^2) were removed
3. Vegetation areas were merged with a very small parameter value: 0.2. This acts as a fusion, combining very close areas which belong to the same class.

4. The generalisation of each area follows this scheme:
   4.1. Sometimes road removals
   4.2. Some building removals (generally smallest ones)
   4.3. Building aggregation (convex hull)
   4.4. Exaggeration of a few buildings
   4.5. Remaining building shape improvement: squaring or sometimes simplification (same algorithm with different parameter value).

5.4.3.2 Link Between ObjectType and Algorithms

Table 23 – Algorithms used on object type (test 2).

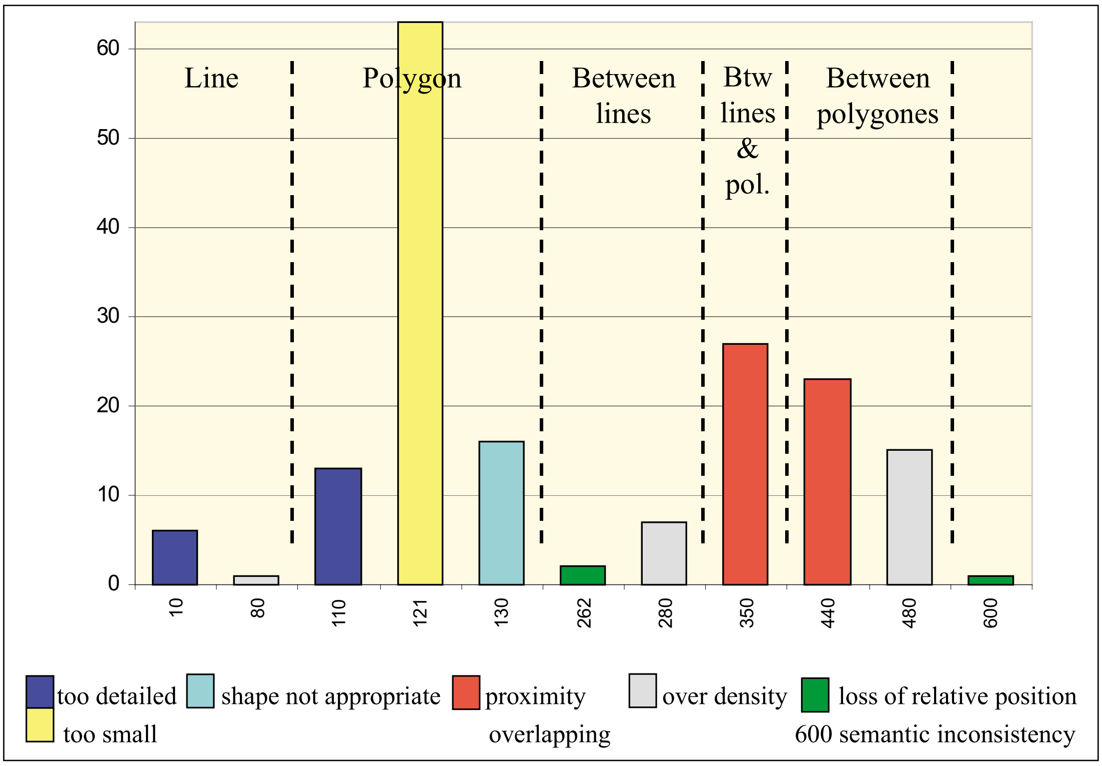
Object type	corner-flipping	douglas	squaring	exaggerate	aggregation	merge	classification	no algorithm	Total	%
Building	12	2	15	29	10	7		23	98	80,3%
Building-all								1	1	0,8%
Sport area						1			1	0,8%
Water tower				1					1	0,8%
Pool				1	1				2	1,6%
Building area	*12*	*2*	*15*	*31*	*11*	*8*		*24*	*103*	***84,4%***
Road							3		3	2,5%
Road department		1							1	0,8%
Road large								8	8	6,6%
Road		*1*					*3*	*8*	*12*	***9,8%***
Fields		3							3	2,5%
Vegetation-all					1				1	0,8%
Wood					1			1	2	1,6%
Vegetation		*3*			*2*			*1*	*6*	***4,9%***
All							1		1	0,8%
Total	12	6	15	31	13	8	4	33	**122**	
%	9,8%	4,9%	12,3%	25,4%	10,7%	6,6%	3,3%	27,0%	100,0%	
% - removal	13,5%	6,7%	16,9%	34,8%	14,6%	9,0%	4,5%		89	

Table 23 analysis

- Most of the operations are performed on buildings (84,4 %). 33 operations of interactive object removals removed around 90 buildings (+ 70 smallest one at the beginning).
- Vegetation is generalised by a global aggregation (cluster) and 6 line simplifications in 3 operations.
- All the 'secondary objects' were removed by a classification process.
- In terms of quantity:
  - 282 buildings were removed by object removal, aggregation and merging. This is 57% of the initial buildings. At the end 214 buildings remained.
  - 28 roads were removed which represent 12% of the initial roads.

5.4.4 Conflicts

5.4.4.1 Detected Conflicts which Trigger Operations



Graph 10 – Conflicts which trigger operations (test 2).

Graph 10 analysis

If we classify the conflict according to object characteristics, we obtain the following:

- Over granularity 14,8%
- Size 39,3%
- Shape inconsistency 13,1%
- Over proximity 12,3%
- Over density 18 %
- Inconsistency 2,4%

Table 24 – Algorithms used to solve conflicts (test 2).

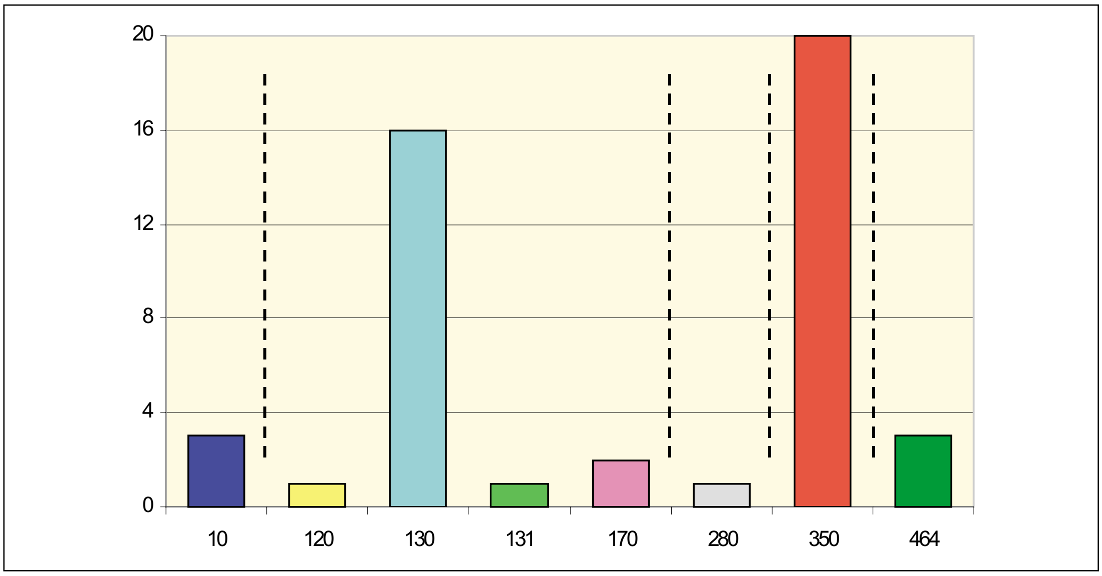
main conflict	douglas	corner-flipping	aggregation	exaggerate	squaring	classification	int. Removal	merge	Total	%
Line too detailed 10	4	1							5	4,1%
Polygoon too detailed 110	2	10					1		13	10,7%
Polygon too small 121			10	31			6	1	48	39,3%
Polygon not square 130		1			15				16	13,1%
Loss of connectivity 262						2			2	1,6%
Lines over density 280							7		7	5,7%
Line and polygons overlap. 350						1	2		3	2,5%
Polygons too close 440			3				2	7	12	9,8%
Polygons over density 480							15		15	12,3%
Smenatic inconsistency 600						1			1	0,8%
Total	6	12	13	31	15	4	33	8	122	100%
Number of objects	10	16	35	126	86	3	197	13	486	
On all objects			1: vegetation			1: all	1: buildings			

Table 24 analysis

- We notice that conflicts related to one object are solved by an appropriate operation:
- Granularity => simplification algorithm (Douglas or corner flipping)
- Size => exaggeration or aggregation
- Shape => squaring (for buildings)
- Proximity or over density are mainly solved by interactive object removals since appropriate algorithms for displacement and selection are missing.
- When polygons are qualified as 'too close' they are aggregated (aggregation or merging), otherwise they are qualified as 'too dense'.
- The relationship between conflicts and solutions is logical although some algorithms are missing.

5.4.4.2 Final Assessments

Final assessments represent unsolved conflicts. Each described conflict often represents a set of similar conflicts within an area. Of the 48 regions, 22 regions still contained a total of 48 conflicts at the end of the process.



10 Lines too detailed
120 Polygon too thin
170 Polygon too far from initial position
130 Poygon not square
131 Loss of polygon shape
280 Lines over density
350 Line and polygon overlapping
464 Loss of polygon relative position

Graph 11 – Unsolved conflicts (test 2).

Graph 11 analysis

Most of the remaining conflicts are related to:

- Polygons not squared enough: it seems that the squaring algorithm is not sufficient.
- Proximity and overlapping, due to the lack of a displacement function.

5.4.5 Use of Algorithms

5.4.5.1 Algorithm Assessment

After each process, the user specified a level of satisfaction (table 25). Three values were possible:

0= bad; 1= medium; 2= good results. Bad was never mentioned.

Table 25 – Visual assessment of algorithm (test 2).

Conflict as.	aggregation	classification	corner-flipping	douglas	exaggerate	merge	squaring	int. Removal	Total	
1	**10**		4	2	1	2	**14**	**18**	51	45,1%
2	3	4	7	4	29	6	1	8	62	54,9%
Total	13	4	11	6	30	8	15	26	113	100%

Table 25 analysis

The results of 45% of the operations were not considered excellent and we notice that squaring and aggregation were specifically criticised.

- Displacement was never used because it does not give appropriate results.
- The aggregation method used produces a convex hull of the buildings involved and certainly the final shapes are not appropriate.
- The squaring is considered as 'not square enough' by the user.
- The corner-flipping is criticised in 36% of the cases.
- Merge, exaggeration and Douglas were appreciated.

5.4.5.2 Parameter Value

As it has been stated previously, many algorithms are performed on sets of objects.

Aggregation:

- Used 13 times

Cluster tolerance:

- For buildings always 7.6 m
- For vegetation 0.2 m (it acts as a fusion)
- For woods 30 m

Merge:

- Used 8 times

Length tolerance:

- For buildings: min = 6 m; max = 13.3 m; average = 9.5 m; standard deviation = 2.9 m
- For sports area 100 m

Aggregation and merge parameter values are not dispersed and are similar. At the final scale they correspond to a value between 0.15 mm and 0.2 mm, which is a high proximity.

Artificial simplification:

1. **as squaring**

- used 15 times

Corner flipping algorithm (artificial simplification)

- Length tolerance always 0 (no simplification)
- Angle tolerance always 30 degrees

2. **as corner flipping** used 12 times

Three parameters (angle, long line and short line) allow the removal of shapes such as hats or stairs.

If one shape's longest dimension is smaller than 'long line' and the shortest dimension of the shape is smaller than 'short line', the shape is removed, otherwise the shape is preserved. This logical can hardly ensure a stability of generalisation output for intermediate shape size.

On this population, the parameter values are not correlated.

Table 26 – Statistical analysis of simplification parameter values (test2).

	Angle	long line	short line
Average	24,3	61	8,5
Median	30	47,6	7,5
More frequent	30	101	
standard deviation	8,5	31,3	3,8
max-min	19,6	77,7	13,8
min	10,4	23,9	5,6
max	30	101,6	19,4

The average long line length of 61m corresponds to 1.2mm at the final scale, and the average short line length corresponds to 0.17 mm.

Douglas

– Used 6 times

Length tolerance around 35 for vegetation, 28 for roads, 15 for buildings
Angle tolerance always 30 degrees

Exaggeration

– Used 36 times with different parameters (31 times referred but for several applications, sometimes with 2 parameter values). We also tried to evaluate the percentage related to the number objects dilated.

Table 27 – The percentage of size increasing value used (test 2).

% of increase	Frequency	% of time	% of objects
50	21	58.3	63.5
100	11	30.6	22.2
150	3	8.3	13.5
200	1 (on water tower)		0.8

These size dilations are stronger than for the other test on Lamps.

5.4.6 Visual Result

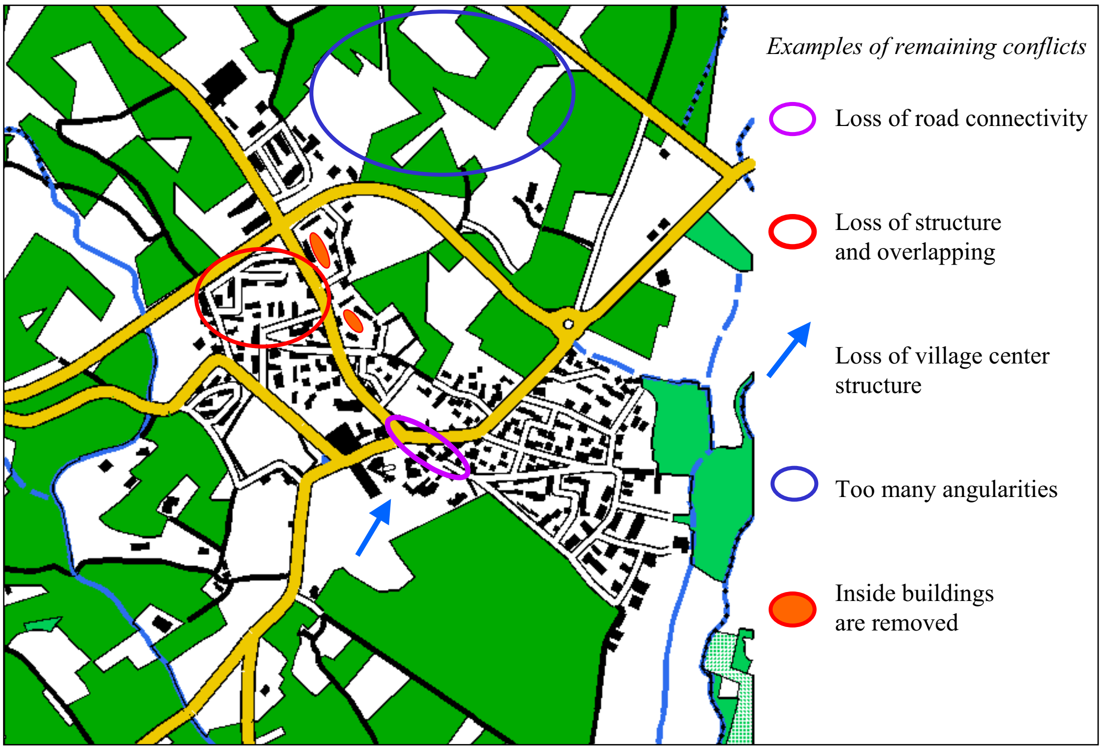


Figure 12 – Visual result (test 2).

Figure 12 analysis

- Generalised data are clearly more readable than non-generalised ones because the data were greatly simplified.
- Vegetation boundaries are over generalised: angles are too sharp.
- A lot of buildings are overlapped by streets, even where free areas exist.
- The village centre disappeared somehow: the high initial density which helps us to identify the village structure was not preserved very well. Moreover, the street network is not continuous in this area.
- Globally the evolution of building density is not homogeneous from the initial to the final scale.
- House patterns are not preserved (distribution of buildings).
- Some buildings remain too detailed and not squared enough.

5.5 Test 3: Generalisation on LAMPS2 – IGN France

This test was performed with LAMPS2 at IGN France in January 1999 using the same set of algorithms as in test 2. It was carried out by *Annabelle Boffet*, a Ph.D. student studying urban area classification for generalisation purposes.

5.5.1 Global Results

- Number of operations — 300 divided in:
  - 205 (68 %) with a generalisation algorithm;
  - 95 (32 %) interactive solutions

    89 % of these interactive operations were object removals
- All operations are triggered interactively. One operation is triggered per object except for object removal.
- Subjective visual satisfaction on 48 regions:
  - 36 are well generalised 75 %
  - 5 are acceptable 10,4 %
  - 7 are badly generalised 14,6 %

5.5.2 Operations and Algorithms Used

Table 28 – Operations and algorithms used (test 3).

Algorithm	simplification	exagerate	squaring	displacement	rotation	split	delete	Total	% Algorithm
32	**37**		1					38	18,5%
74		**67**						67	32,7%
83	1		**93**					94	45,9%
102				**6**				6	2,9%
No Algorithm			3	1	4	2	**85**	**95**	
Total	38	67	97	7	4	2	85	**300**	
% Operation	12,7%	22,3%	32,3%	2,3%	1,3%	0,7%	28,3%		

Table 28 Analysis

- Only four algorithms were used
  - Simplification of boundary. The platform proposed two algorithms:
    - **32: Douglas-Dispike:** This algorithm was used for building simplification although it is made for non-orthogonal lines.
    - **Artificial object simplification** (for buildings). Although this algorithm is assigned to building generalisation, it was not used because its parameter values were too difficult to tune.
  - **74: Exaggeration:** This algorithm is used either to dilate an area or to caricature a line. In this test it was used for building dilation.
  - **83: Squaring**
  - **102: Displacement:** This algorithm displaces the coordinates of objects to allow a minimum separating distance. It is based on cluster displacement. The amount of displacement can take into account object types. For lines, the displacement occurs only in conflict areas, which creates a local distortion of the line geometry.
- Main transformations used are:
  - Squaring
  - Object removal
  - Dilation
  - Simplification
- The correlation between operations and algorithms is nearly perfect. The two cases of non coincidence certainly correspond to mistakes during grid filling.

5.5.3 Operations on Objects

5.5.3.1 Repetitive Sequences on Objects

- The generalisation was performed on each individual region in numerical order.
- The first operation is the removal of unimportant objects such as line land, streets, buildings.
- Then the buildings were generalised, one after the other. All buildings which are generalised are squared.
- The last operation in a region was sometimes an object removal certainly due to the side effect generated by object dilation (exaggeration).

On 496 buildings:

- 216 were removed 43,5 % (grouped in 25 operations)
- 98 were generalised 19,8 %
- 182 remained unchanged 36,7 %

The operations used for buildings is described hereafter:

Table 29 – Sequence of algorithms on buildings (test 3).

Building generalisation	1	1	2	2	3	3	3	3	4
	Delete	Squaring	Exaggerate Squaring	Simplification Squaring	Exaggerate Simplification Squaring	Exaggerate Squaring Rotation	Displacement Simplification Squaring	others	Exaggerate displacement Simplification Squaring
123 operations:	25	13	41	17	16	4	2	3	2
%:	20,3%	10,6%	33,3%	13,8%	13,0%	3,3%	1,6%	2,4%	1,6%
% of gene algo		13,3%		59,2%				25,5%	2,0%

Table 29 analysis

- Most of the time two or three operations are required to generalise buildings.
- Exaggeration and squaring are the most frequent sequences.
- Simplification and squaring are used for large buildings. They are often industrial ones.
- Sometimes a rotation is performed as a final step.
- Displacement is always followed by simplification and squaring because it degrades the building shape.
- At the beginning of the test, the sequences varied (learning phase). Later they became more standardised.

5.5.3.2 Object Type, Operators and Algorithms

Table 30 – Operators used on object type (test 3).

type	simplification	exaggerate	squaring	displacement	rotation	split	delete	Total	%
building	32	65	87	4	3		23	214	71,3%
Special Building	6	1	10	2	1		2	22	7,3%
Land Boundary							27	27	9,0%
Land Area		1						1	0,3%
Street						2	14	16	5,3%
Mixed lines				1			19	20	6,7%
Total	38	67	97	7	4	2	85	300	

Table 30 analysis

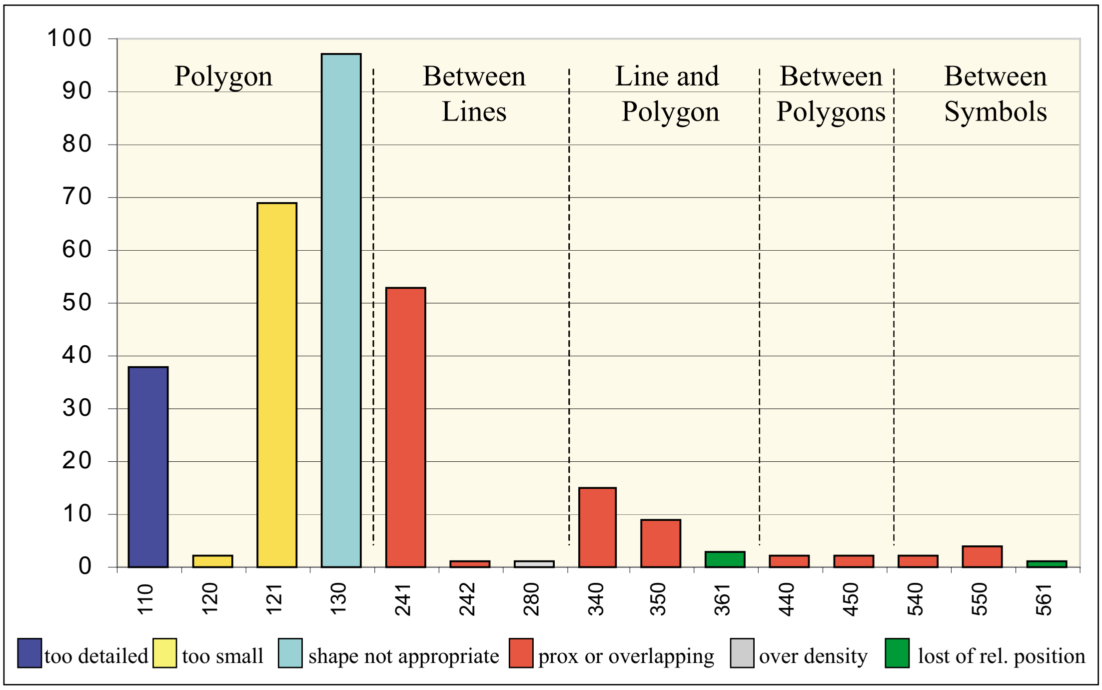
- Most of the operations are realised on buildings (77,6%)
- Other types of objects are mainly removed or not generalised

Table 31 – Algorithms used according to object type (test 3).

type	32: Dougl	74: exag.	83: squaring	102: displace	Algorithms	No algorithm	Total
Building	32	65	84	4	185	29	214
Special Building	6	1	10	2	19	3	22
Land Boundary						27	27
Land Area		1			1		1
Street						16	16
Mixed lines						20	20
Total	38	67	94	6	205	95	300

5.5.4 Conflicts and Conflict Resolution

5.5.4.1 Detected Conflicts which Trigger Operations



Graph 12 – Conflicts which trigger operations (test 3).

Graph 12 analysis

- Polygons are considered as too small (121), too detailed (110), and not squared enough (130)
- Conflicts on lines are related to proximity. They are solved by interactive line removal.
- Polygons too close are almost not described because no solution can be provided.
- The over density is seldom identified. The identification is often local, object by object.

Table 32 – Algorithms used to solve a conflict (test 3).

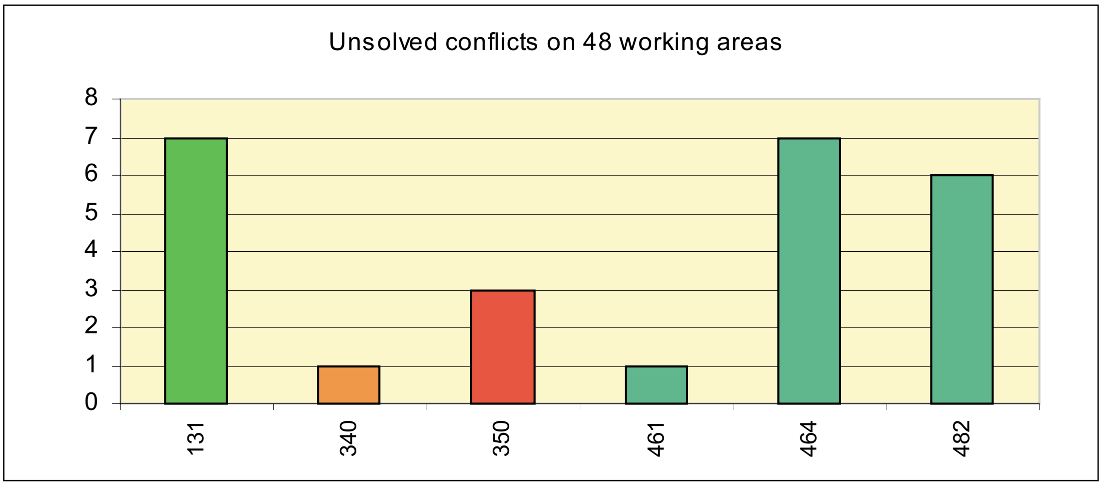
Conflit Code	simplification	exagerate	squaring	delete	displacement	rotation	split	Total
110 Poly too detailed	38							38
120 polygon too thin				2				2
121 polygon too small		67		2				69
130 Shape not appropriate			97					97
241 lines too close				51			2	53
242 2 // lines too close					1			1
280 overdensity of lines				1				1
340 line and pol too close				12	3			15
350 line and pol overlap				6	3			9
361 line and poly orientation								
440 polygons too close				2				2
450 polygons overlapping				2				2
540 symbols too close				2				2
550 symbols overlapping				4				4
561 symbols bad alignment								
Total	38	67	97	84	7	4	2	299

Table 32 analysis

– Whenever an algorithm exists, the association between conflicts and operation is logical.
– Otherwise, the lack of algorithms induces the use of interactive object removal (delete) as a magic operation for solving conflicts which are not really related to density. For example:
  – Polygon too thin, too small
  – Objects too close: lines, line and polygons, polygons, symbols.
– When a line and one or several polygons are too close (340, 350), polygons are removed except for important buildings where a displacement is performed.

5.5.4.2 Final Assessment

The assessment was done globally on the 48 working areas:



131 loss of building shape
340 proximity btw line and polygon
350 overlapping btw line and polygon
461 loss of polygons alignment
464 loss of relative position
482 loss of relative density

Graph 13 – Unsolved conflicts (test 3).

Graph 13 analysis

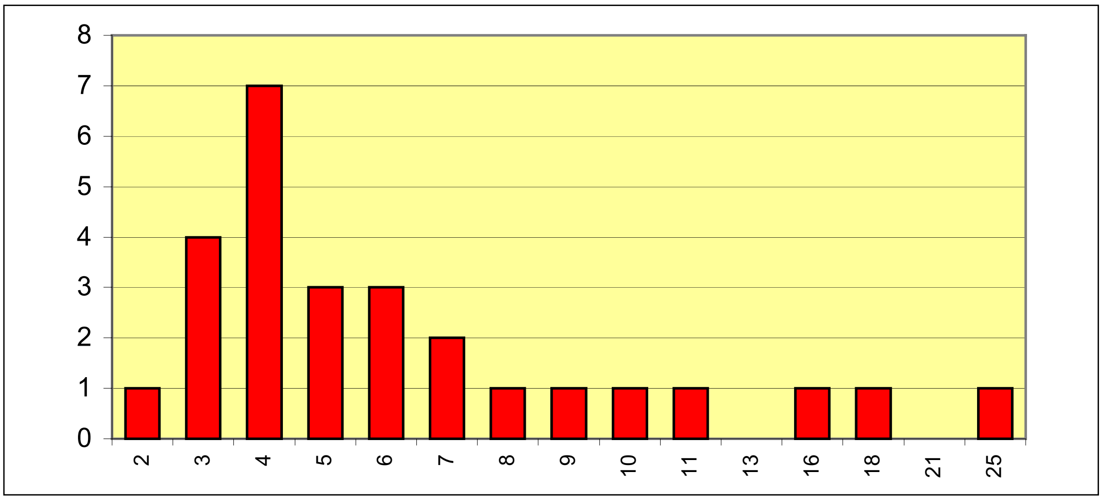
- 7 building shapes were degraded.
- On 8 areas (i.e. 40% of urban areas) the user noticed the loss of relative position between polygons (461 & 464). Most of the time such problems are linked with conflict 482: loss of relative density. It means that the building removal was done without preserving the building's relative position, and the local process, area by area, did not allow the maintenance of the differences of density between areas.
- Some proximity conflicts between polygons and lines could not be solved (340, 350).

5.5.5 Use of Algorithms

5.5.5.1 Douglas

The Douglas algorithm was used as the building simplification algorithm, although an 'artificial area' simplification algorithm is proposed. The choice is due to the complexity of tuning the parameter of this algorithm.

Amongst the possible parameters, only the 'offset distance' which corresponds to the classical Douglas parameter is used. The values are grouped around 4–5 meters. The number of values in the following study (graph 14) is under the number of times this algorithm was used because some values were not recorded.



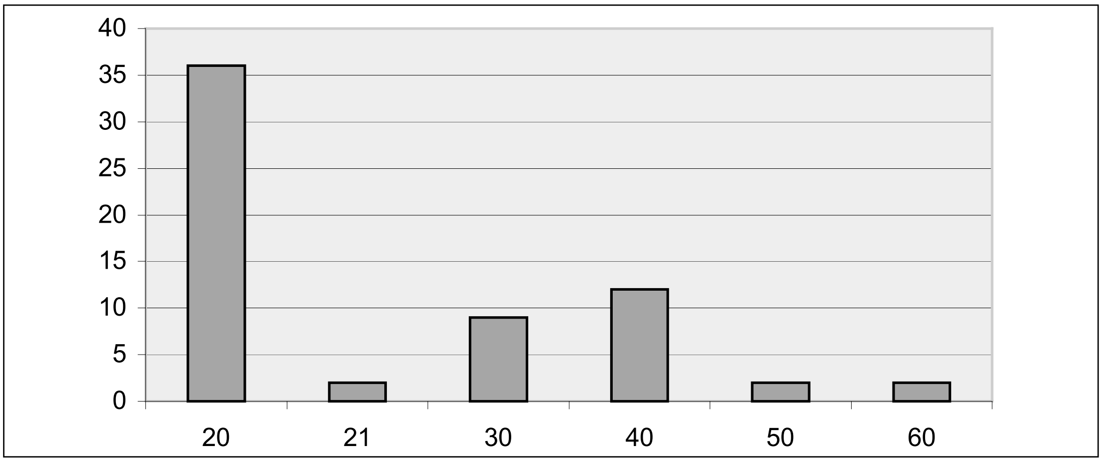
Average = 6,9 Median = 5 More frequent = 4 St dev = 5,3

Graph 14 – Distribution of Douglas and Peucker parameter value (test 3).

In order to correct shape distortion due to this algorithm, squaring was always performed afterwards: In 90% of the cases, simplification generates a 'shape not appropriate' conflict (130).

5.5.5.2 Exaggeration

This algorithm can be used for areas or for lines. For areas it works as a dilation algorithm. The parameter value is the percentage of the increase in area. A 100% exaggeration doubles the area of a polygon.



Average = 27,5 Median = 20 More frequent value = 20 Standard deviation = 10,8

Graph 15 – Distribution of exaggeration parameter value (test 3).

A percentage is not the best parameter value to guide this process. A target size would be more appropriate for cartographic generalisation purposes.

Actually, the analysis of the final building size shows that **35 buildings** are still under the minimum size threshold at the end. It is a classical effect of interactive generalisation where it is difficult to visually control object size, especially on the screen with zoom capacity.

5.5.6 Visual Result

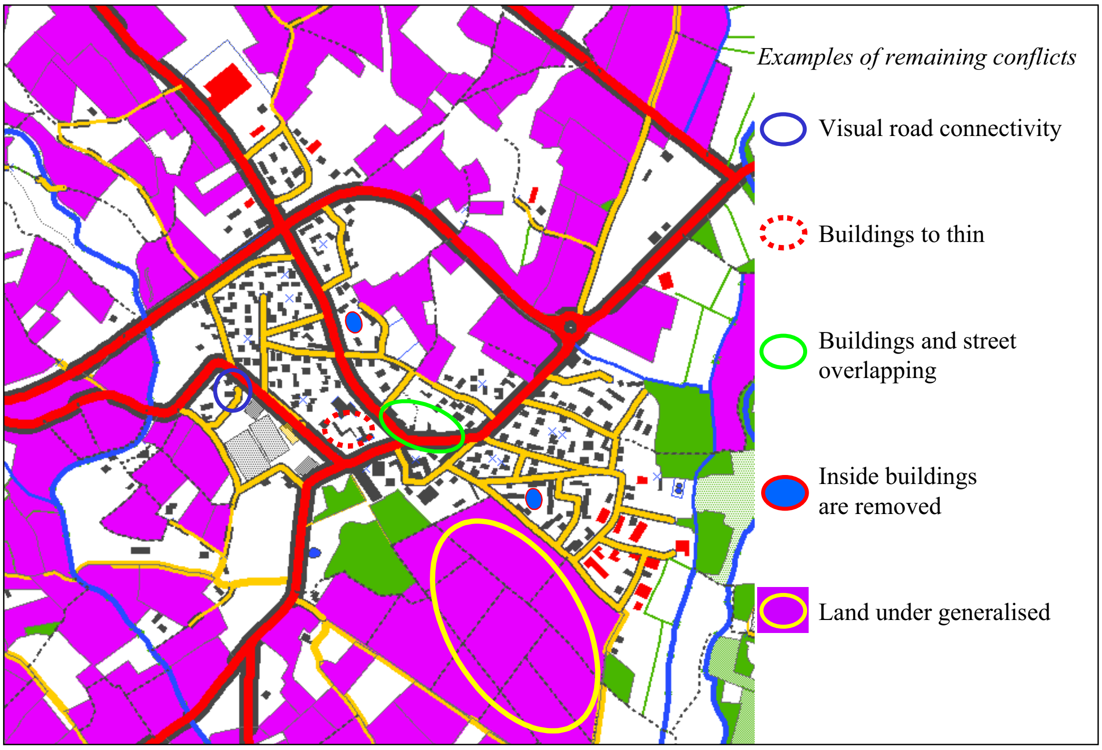


Figure 13 – Visual result (test 3).

Figure 13 analysis

- Generalised data are clearly more readable than non-generalised ones.
- Building density is quite homogeneous and some patterns were maintained, even though the final building pattern is less regular than the initial one.
- A lot of buildings are overlapped by streets, even though some free areas exist. Moreover, some buildings remain too close together.
- Buildings are slightly too small for the final scale.
- Some building shapes are badly affected.
- The village centre is not generalised as well as other areas. Perceptually, it disappeared.

5.6 Test 4: Generalisation on a Set of Packages – ICC

This test was performed at the ICC by Maria Pla and Blanca Baella.

The ICC is the National Mapping Agency of Cataluna which produces databases and maps. In terms of generalisation, they produce 1:10 000 from 1:5 000 and 1:100 000 from 1:50 000. For those productions they use two generalisation packages:

- MGE/MG from Intergraph for medium scale
- CHANGE from Hanover University for large scale.

They also have some internal software.

This generalisation test does not correspond to their production line. Consequently, they used different software programs from:

- CHANGE: ANGI (building generalisation in CHANGE)
- MGE: Douglas, simple average, aggregation, point typification, line typification
- their own library: area-to-point transformation (collapse).

The generalisation process follows the production philosophy, which is: trying to utilise batch processing as much as possible to reduce interactivity. We can make the hypothesis that this test follows the steps a production line would use if the set of algorithms were satisfactory. Operations are always performed on a large set of data. Consequently, we cannot follow the same analysis method used for the previous tests where individual steps were described separately.

5.6.1 Global Results

- Number of operations 35, divided into:
  - 19 with generalisation algorithms
  - 16 interactive (12 object removal sequences)
- All operations are performed on a set of objects.
- The sets of objects on which algorithms are performed are defined by their nature. For this purpose, when the initial classification is not appropriate, objects are shared by two classes.
  - On the 19 algorithms used:
  - 8 are considered as correct 42 %
  - 11 are considered as medium 58%

5.6.2 The Process

5.6.2.1 The General Sequence

For clarity purposes we have grouped the 35 operations into 11 sequential steps:

Step 1 Unimportant objects are removed. 10 classes of objects are removed in 10 operations: trees, paths, land lines, brushwood areas, etc.

Step 2 Buildings larger than 400m^2 are generalised by means of ANGI (building generalisation package in CHANGE) (around 25 buildings). This generalisation simplifies building shape and aggregates some close ones.

Step 3 Small buildings (< 400m^2) are transformed into symbols (as test number 1).

Step 4 Buildings which are symbols are typified

Step 5 Line objects are either simplified or smoothed

Step 6 Road typification is performed on some road sub-types: 1- street in the village centre; 2- unimportant roads.

Step 7 Pedestrian tracks are typified

Step 8 3 operations are performed on hydrographic polygons

Step 9 Walls are classified interactively and the unimportant ones are removed

Step 10 Too close wooded areas are aggregated

Step 11 Small orchard areas are removed

More generally the sequence is:

1. Unimportant object removal according to their type
2. Building generalisation: either dilation or change into symbols and typification
3. Network (road and river): geometric improvement
4. Road and track typification
5. Remaining treatment of rural objects: wall, wood, lake.

5.6.2.2 Algorithms on Set of Objects

Table 33 – Algorithms used on object type (test 4).

type	ANGI	Area to pt	Point typfic.	Line typific.	Douglas	Spl average	Aggregation	Total
buildings	1	1	1					3
roads, tracks				3	5			8
hydro-linear					2	2		4
hydro-area		1			1		1	3
vegetation							1	1
Total	1	2	1	3	8	2	2	19

Table 33 analysis

7 algorithms were used; they are often related to a type of object:

– ANGI simplifies and aggregates buildings

– Small buildings are transformed into symbols (area to point) and typified by point typification

– Roads and tracks are simplified by Douglas and typified by line typification

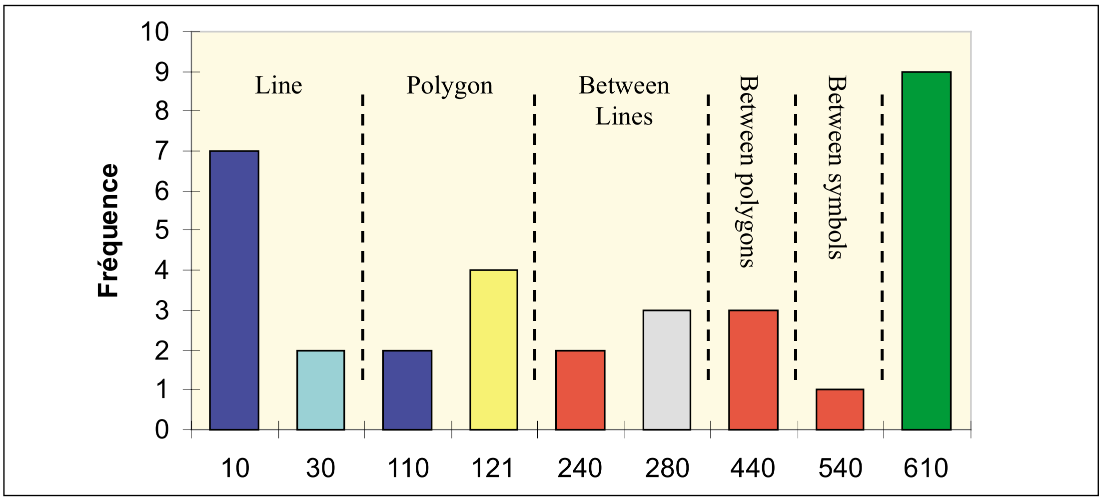
– Rivers are simplified by Douglas and smoothed by simple average

– One hydro polygon is transformed into a symbol (area to point), 2 polygons are aggregated
– Connected vegetation polygons are aggregated.

Other operations are interactive.

5.6.3 Conflicts and Conflict Resolution

5.6.3.1 Conflicts which Trigger Operations



10, 110	too detailed	30 Bad shape	240, 440, 540	too close	610 Non important
121	too small		280	too dense	

Graph 16 – Conflicts which trigger operations (test 4).

Graph 16 analysis

On 35 operations, 33 were justified by a specific conflict code.

– The number of conflicts is smaller than the other tests because objects are generalised by groups according to their semantics.
– Conflict codes are more generic.
– The ICC used a new code (610) to qualify unimportant objects for removal.

5.6.3.2 Conflicts, Operations and Algorithms

Table 34 – Operations used to solve conflicts (test 4).

	Simplif.	Smoothing	Simp+ Agg.	Collapse	Caricature	Typification	Aggreg.	Fusion	Classific.	Objj remova	Total	
Line too detailed 10	7										7	20,0%
line too angular 30		2									2	5,7%
olygon too detailed 110	1		1								2	5,7%
Polygon too small 121				2						2	4	11,4%
Line proximity 240					1					1	2	5,7%
Line over density 280						3					3	8,6%
Polygons too close 440							1	1			2	5,7%
Symbols too close 540						1					1	2,9%
Not characteristic 610										9	9	25,7%
lo specified conflict									3		3	8,6%
Total	8	2	1	2	1	4	1	1	3	12	**35**	
	22,9%	5,7%	2,9%	5,7%	2,9%	11,4%	2,9%	2,9%	8,6%	34,3%		

Table 35 – Algorithms used to solve conflict (test 4).

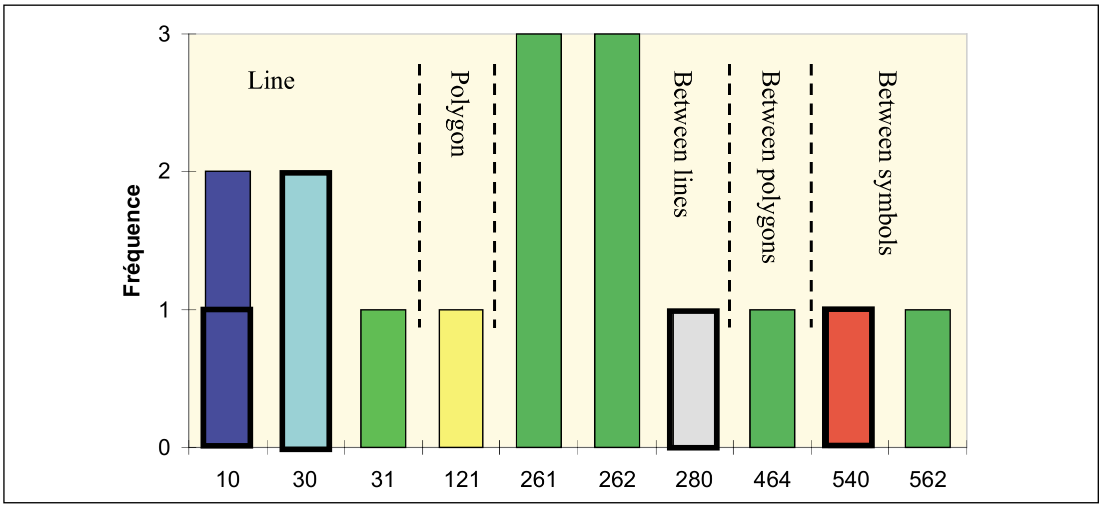
	Douglas	Sp average	ANGI	Area to pt	L. typifi.	Aggreg.	Pt typifi.	Total
Line too detailed 10	7							7
Line too angular 30		2						2
olygon too detailed 110	1		1					2
Polygon too small 121				2				2
Line over density 280					3			3
Polygons too close 440						2		2
Symbols too close 540							1	1
Total	8	2	1	2	3	2	1	**19**
	42,1%	10,5%	5,3%	10,5%	15,8%	10,5%	5,3%	

Table 34, table 35 analysis

– Both tables are nearly diagonalised, which shows that in general each algorithm is assigned to a single kind of conflict on a type of geometry.

– Object removal is mainly used to remove unnecessary objects and those that are too small (which can also be considered as unnecessary).

5.6.3.3 Generated and Unsolved Conflicts



10 to detailed 30 Shape not appropriate 121 too small 280 over density
31 los of shape 262 los of connectivity 261, 464, 562 loss of spatial distribution

Conflicts solved by other operations

Graph 17 – Generated and non-solved (unsolved?) conflicts (test 4).

Graph 17 analysis

- This graph shows that some conflicts were solved by algorithms:
  - conflict 540 (symbols over density) was solved by symbol typification
  - conflict 30 (line with too many angularities) was solved by smoothing
  - conflict 280 (line over density) was solved by line typification.
- Other conflicts are generated by using algorithms and cannot be corrected by post treatment. These conflicts are related to the loss of object distribution, as can be noticed on the final map.

5.6.4 The Use of Algorithms

5.6.4.1 Algorithm Assessment

Table 36 – Algorithm assessment (test 4).

Assessement	Douglas	ANGI	Spl average	Aggregation	Area to pt	Line typific.	Point typific	Total
Average	3	1	1	1	1	3	1	11
Good	5		1	1	1			8
Total	8	1	2	2	2	3	1	19

Table 36 analysis

- ANGI => Loss of polygon's spatial structure
- Point typification => Loss of symbol's spatial structure
- Douglas => Too many angularities
- Line typification => Loss of linear spatial structure and loss of network connectivity

5.6.4.2 Parameter Values

Since only a few operations are performed, no statistics are computable:

- Douglas is always used with 5m which corresponds to a visual threshold at this scale
- Point typification transforms a set of points into a new subset. It is used with 35m as the cluster tolerance value.
- Line typification: with MGE line typification, the 'conflict resolution' is used which is better for rivers (tree) than for roads (close graph). The minimum spacing parameter value depends on the importance of the object. The chosen values were: 70 meters for secondary roads, 100 meters for streets in village centres, 150 meters for pedestrian tracks.
- The threshold to collapse polygons to symbols is 400m^2 which corresponds to the minimum area value for this scale.

5.6.5 Visual Result

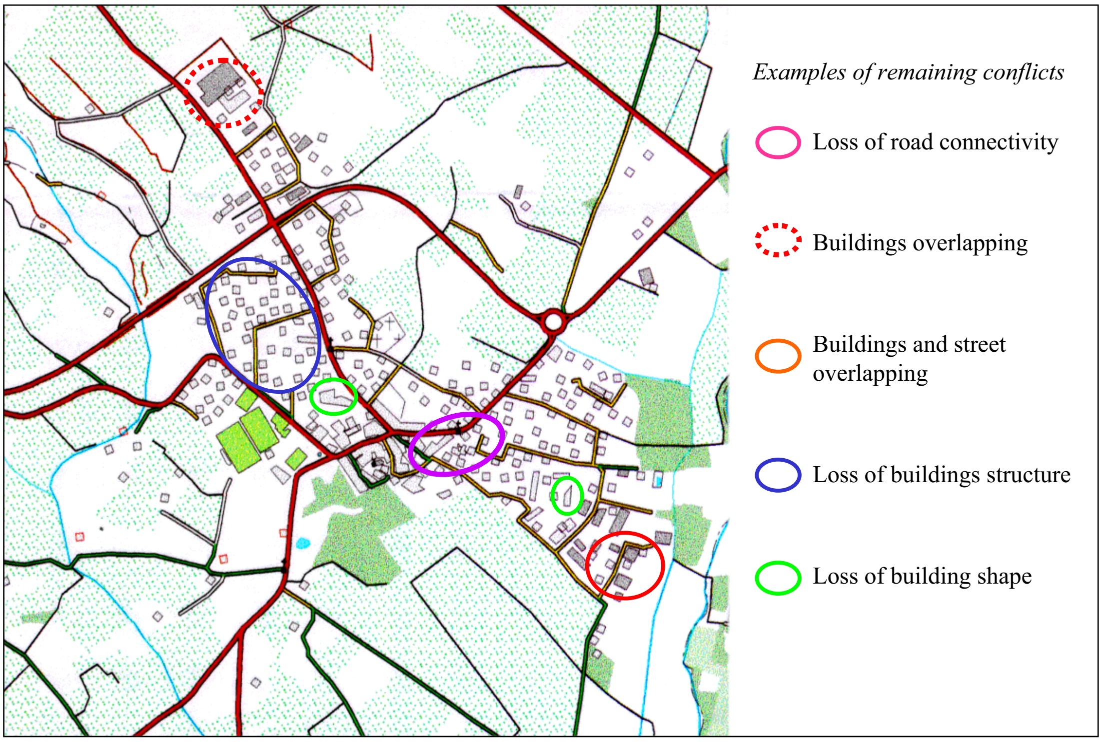


Figure 14 – Visual result (test 4).

Figure 14 analysis

- Batch processing degraded the geographical data significantly
- Road connection is not maintained. This is due to the inadequacy of the spanning tree approach (in Line typification) for street network removal
- Building overlaps are generated by ANGI and could not be solved by another algorithm
- Point typification applied on buildings degraded the overall distribution of buildings
- Buildings and streets are too close to one another.
- Some building shapes are degraded after simplification

5.7 Test 5: Generalisation on Stratège – IGN France

This test was performed by *Anne Ruas* from IGN on the Stratège platform (an Object Oriented prototype developed at the Cogit laboratory for contextual generalisation research).

For an independant analysis, this test was not analysed by *Anne Ruas* but by *Sébastien Mustière*. For homogeneity and explanation purposes the description of the analysis was completed by *Anne Ruas*.

5.7.1 Global Results

Two types of operations that were not performed in other tests appear here: object creation (because Stratège is an object-oriented system) and evaluation (because Stratège contains some evaluation tools used to guide the system).

- Number of operations — 308 operations were performed:
  - 199 generalisation transformations
  - 41 object creations (NB: in the object-oriented sense)
  - 68 automated evaluations
- Interactive evaluation after each operation. This evaluation checks whether the algorithm worked correctly according to what it is supposed to do.
  - a large amount of operations are considered of good quality (87 %)
  - some are considered of medium quality (12%)
  - and few of bad quality (1%)
- Interactive evaluation after a set of operations in an urban block surrounded by streets. This evaluation checks whether the working area is generalised well and sufficiently.
  - half of the generalised block still contains one or more conflicts.

5.7.2 Operations and Algorithms Used

Table 37 – Operation and algorithms used (test 5).

Operation	**Number**	**%**	**Algorithm**	**Algorithm applied on…**	**Number**	**%**
Removal	65	21%	Remove and center	2 buildings	18	6%
			Street-typification	Whole city	1	0.3%
			Dead-end-removal	1 dead-end	29	9%
			Delete-small-build.	all buildings in a block	16	5%
			Remove-hole	1 hole	1	0.3%
Aggregation	26	8%	Aggreg-disp	2 polygons	7	2%
			Fusion-meso	2 blocks	19	6%
Typification	22	7%	Building-removal	all buildings in a block	22	7%
Dilation	44	14%	Dilation	all buildings in a block	38	12%
			Enlarge in Rectangle	1 building	6	2%
Displacement	42	14%	LDT-disp	all buildings in a block	35	11%
			Shift-dxy	1 building	6	2%
			Deform-curvilinear	1 line	1	0.3%
object creation	41	13%	Net-partition	Whole city	1	0.3%
			Create-object	Whole city	2	1%
			Micro-constraint	all buildings in a block	38	12%
Evaluation	68	22%	Compute-constraint	a block	68	22%
Total	**308**	**100%**	**Total**		**308**	**100%**

– **Remove-and-center:** replaces two polygons by the largest one at the centre of gravity. It is a cluster typification applied on two polygons.
– **Street-typification:** removes streets by aggregating urban blocks
– **Aggreg-disp:** aggregates two buildings while displacing them one towards the other
– **Building removal:** removes buildings within an urban block according to congestion, size and density criteria until the density exceeds a specific value.
– **LDT-disp:** displaces buildings within an urban block to avoid overlapping. The displacement is triggered by street symbolisation
– **Shift-dxy:** translates a polygon
– **Deform-curvilinear:** propagates and smoothes a local deformation of a line towards connected lines
– **Net-partition:** partitions the space according to (road) network hierarchy
– **Micro-constraint:** computes the constraint on a single object. It is now limited to size constraint
– **Dilation:** all buildings within a block dilate themselves up to the minimum building size.
– **Compute-constraint:** computes constraints at meso level. Proximity, density, component sizes and semantic constraints are computed.

Table 37 Analysis

– The algorithms used have different application areas:
  – 43 transformation operations (without evaluation and object creation) are applied on one object only
  – 44 involved two objects
  – 111 involved all the buildings contained within an urban block
  – 1 involved the whole city
– If we classify the operations according to the type of transformation they perform, we obtain:
  – 35 % of the operations (removal, typification, aggregation) are used to reduce the quantity of information
  – 14 % of the operations (dilation) are used to change the geometry of simple objects. The only algorithm used to simplify the geometry of an object is “enlarge to rectangle” (that transforms a building into a rectangle) because this is the only one available in StratËge.
  – 14 % of the operations (displacement) are used to displace the objects
  – 35 % of the operations (object creation and evaluation) are not transformation operations but operations necessary for guiding the process.

5.7.3 Operation on Objects

5.7.3.1 Global Sequence

Four operations were first been performed on the entire data set:

1. The first two operations duplicate initial object geometry for visualising and controlling the data evolution.
2. The third operation creates partitions according to road hierarchy. This partition is used for street selection.
3. The fourth operation removes streets. It is based on the creation and aggregation of urban blocks.

Areas surrounded by streets (urban blocks) are then treated one by one automatically. These areas are not always the same as the reference areas defined for the test purpose. Some reference areas may contain several blocks, and a block may contain several reference areas because of the street removal.

The chronological order of block treatment was:

1. the town centre blocks,
2. the peripheral blocks close to the town centre. These blocks contain several buildings concentrated in one part of the area.
3. Country areas containing isolated buildings or no buildings at all.

Some areas with no buildings were not treated at all.

Analysis

– Working areas are defined for generalisation purposes; they are different from initial areas given for the test.
– These working areas are generalised in a thematic order.

5.7.3.2 Sequence in a Block

Sequences of block generalisation use some operations on the whole block (typification, dilation, displacement), and some on more local ones (remove-and-center, eliminate dead-line, enlarge in rectangle, shift dxy, deform-curvilinear).

Local operations are used to perform local improvements in two cases:

– where the global algorithm did not produced acceptable results
– where individual cases need to be treated in particular ways

The sequence of global operations can be analysed further.

The first global operation used in a sequence is:

– a typification in 62% of the blocks. These blocks are the ones with high density (buildings have to be removed because of graphic space limitation) or medium density (buildings have to be removed to respect the difference of this density to that of a high density area).
– a displacement in 18% of the blocks. These blocks have low density where no buildings have to be removed and where buildings are close to each other
– a dilation in 21% of the blocks. These blocks have a low density where no buildings have to be removed and where buildings are quite far from one another.

Order of use of displacement and dilation:

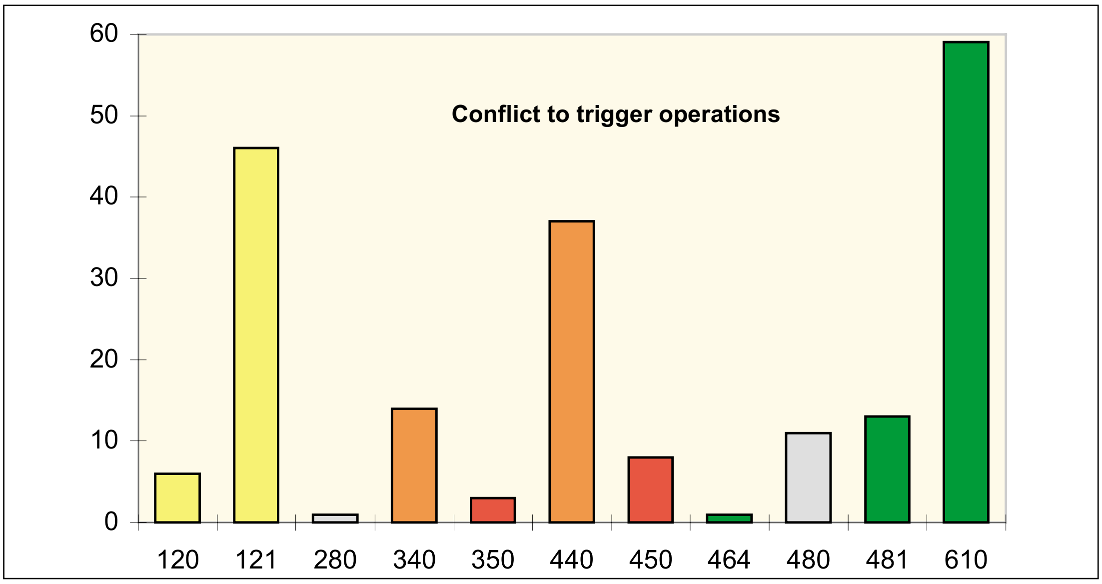
– In 65% of the blocks, the dilation is performed before the displacement. These blocks are the ones with medium density and medium proximity between buildings. If the dilation does not create overlapping conflicts between buildings, it is better to do this before displacing the buildings. Therefore, the sequence “dilation then displacement” is preferrable whenever possible.
– In 20 % of the blocks, the displacement is performed before dilation. These blocks are the ones with very high density where a dilation would create many overlaps between buildings, which is avoided by the displacement algorithm.
– In 15 % of the blocks, one of the operations, either displacement or dilation, is not performed. These areas have a very low density where almost no operation needs to be performed (typically areas with a few and large buildings).

Analysis

- In urban blocks, global and local operations are performed alternatively.
- The typical sequence of global algorithms in urban blocks is:
  - first, a typification when the density is high or medium
  - then, dilation and displacement when the density is medium. “Typification / dilation / displacement” is the most frequently used sequence (on half of the blocks)
  - or displacement and dilation when proximity is high
  - at least one of the three global algorithms (dilation, displacement, typification) is not used when the density is low
- Operations on rural areas are mainly local ones (removal or aggregation).

5.7.4 Conflicts and Conflict Resolution

5.7.4.1 Conflicts which Trigger Operations



Building	120 too thin	121 too small
Buildings and streets	340 too close	350 overlapping
Buildings	440 too close	450 overlapping
Density	280 too many streets	480 too many buildings
464 Loss of structure	481 to preserve density difference	610 Non important objects

Graph 18 – Conflicts used to trigger operations (test 5).

Graph 18 analysis

Classification of conflicts according to object character:

- Size too small 26,1 %
- Over proximity 31,2 %
- Over density 6 %
- To preserve density: 6,5 %
- To correct a structure 0,5 %
- Unimportant object 29,6 %

5.7.4.2 Links between Conflicts and Operations

Table 38 – Operation used to solve conflicts (test 5).

Conflict / Operation	dilation	removal	displacement	aggregation	typification	object creation	evaluation	Total
too thin	6							6
too small	38	8						46
too many streets		1						1
loss of structure		1						1
too close building/street			14					14
street/building overlapping			3					3
non characteristic		40		19				59
too close buildings		7	24	6				37
building overlapping		6	1	1				8
over density		1			10			11
homogeneity of density		1			12			13
None						41	68	109
Total	44	65	42	26	22	41	68	308

Table 38 analysis

- Removal is used to solve many types of conflicts, but especially to eliminate non-characteristic objects (i.e. unimportant).
- Too thin and too small objects are dilated or removed.
- Proximity problems are solved either by displacement, aggregation or removal.
- Density conflicts are solved by typification.
- Object creation and evaluation, which are not transformation operations, are not guided by conflicts.

Table 39 – Algorithms used to solve conflicts (test 5).

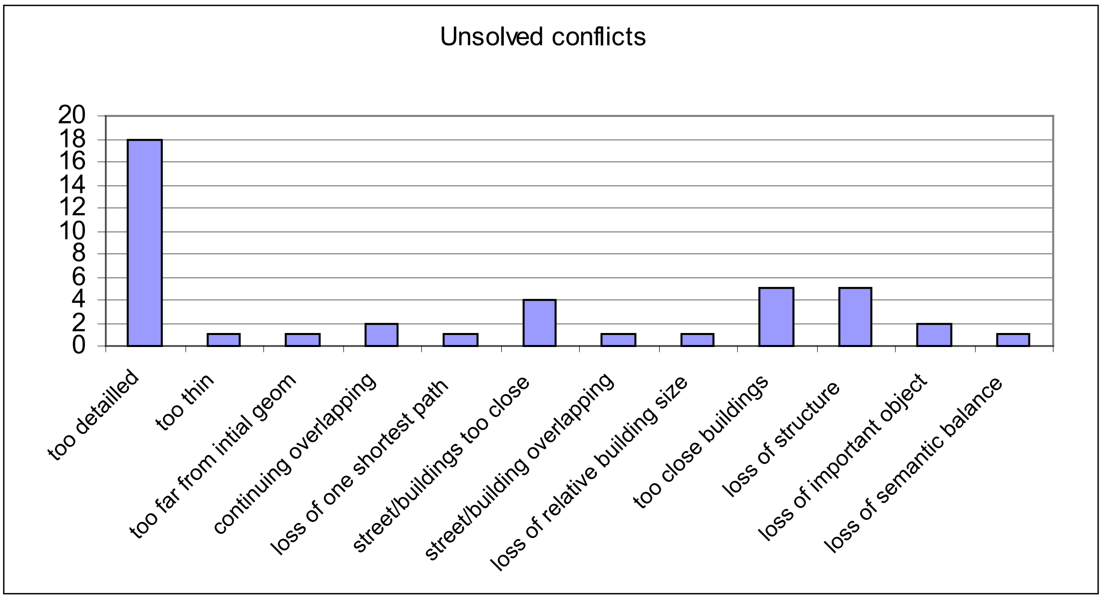
Conflict / Algorithm	Enlarge in Rectangle	street-typification	Remove and center	Dilation	delete-small-build.	dead-end-removal	remove-hole	fusion-meso	LDT-disp	aggreg-disp	shift-dxy	deform-curvline	building-removal	net-partition	create-object	compute-constraint	micro-constraint	Total
too thin	6																	6
too many streets		1																1
loss of structure			1															1
too small			2	38	6													46
non characteristic					10	29	1	19										59
too close building/street									14									14
too close buildings			7						21	6	3							37
building overlapping			6							1	1							8
street/building overlapping											2	1						3
over density			1										10					11
homogeneity of density			1										12					13
None														1	2	68	38	109
Total	6	1	18	38	16	29	1	19	35	7	6	1	22	1	2	68	38	308

Table 39 analysis

- Too thin buildings are transformed into rectangles.
- Too small buildings are dilated.
- Most of the displacements are done with the LDT-displacement algorithms. Very few buildings (3) are moved by a manually chosen shift.
- One algorithm can be used to solve between 1 and 6 types of conflicts.
- One type of conflict can be solved by 1 to 4 different algorithms.

5.7.4.3 Unsolved Conflicts

Half of the generalised areas still contain one or more conflicts after treatment. The following figure lists them.



Graph 19 – Unsolved conflicts after treatments (test 5).

Graph 19 analysis

- Many "too detailed" conflicts are detected after treatment. This is due to the lack of a simplification algorithm in Stratège.
- Some proximity problems (11) remain after treatment.
- Some loss of structure (6) and loss of semantic problems (3) appeared.

5.7.4.4 Automated Detection of Conflicts

Some measures were used to automatically detect the conflicts. Table 40 summarises which conflicts were automatically detected as conflicts during the generalisation process.

Table 40 – Automatic / Interactive identification of conflicts (test 5).

Initial Conflict / Identification	Interactive (Number and %)		Automatic (number and %)		Total (Nb)
too thin	6	100%	0	0%	6
too small	7	15%	39	85%	46
too many streets	1	100%	0	0%	1
too close building/street	0	0%	14	100%	14
street/building overlapping	1	33%	2	67%	3
too close buildings	12	32%	25	68%	37
building overlapping	2	25%	6	75%	8
loss of structure	0	0%	1	100%	1
over density	1	9%	10	91%	11
homogeneity of density	4	31%	9	69%	13
non-characteristic	59	100%	0	0%	59
Total	93	47%	106	53%	199

Table 40 Analysis

- 53 % of the initial conflicts before a generalisation operation were detected automatically (80 % if we do not consider the "non-characteristic", "too thin" and "too many streets" conflicts).
- "Non-characteristic", "too thin" and "too many streets" are never detected automatically.

5.7.5 Use of Algorithms

5.7.5.1 Assessments on Algorithms

Table 41 – Algorithm visual assessment (test 5).

Asses. / Algo	Remove and center	Dilation	Enlarge in Rectangle	aggreg-disp	LDT-disp	shift-dxy	deform-curvline	street-typification	building-removal	dead-end-removal	delete-small-build.	remove-hole	fusion-meso	net-partition	create-object	compute-constraint	micro-constraint	Total
Bad quality					1					1						1		3
Medium Quality		7		4	11			1	5							10		38
Good quality	18	31	6	3	23	6	1		17	28	16	1	19	1	2	57	38	267
Total	18	38	6	7	35	6	1	1	22	29	16	1	19	1	2	68	38	308

Table 41 Analysis

- Algorithms are generally considered efficient (87 %).
- The most criticised algorithms are those which work on one set of objects (LDT-displacement and building removal).
- Evaluation of conflicts (with compute-constraint) is not always considered perfect (because proximity conflicts are over-evaluated and no measure for "Gestalt" evaluation exist).
- Dilation creates side effects (overlapping). It always works well intrinsically, but it would require a 'between buildings' displacement to generate no conflict.

5.7.5.2 Parameter Values

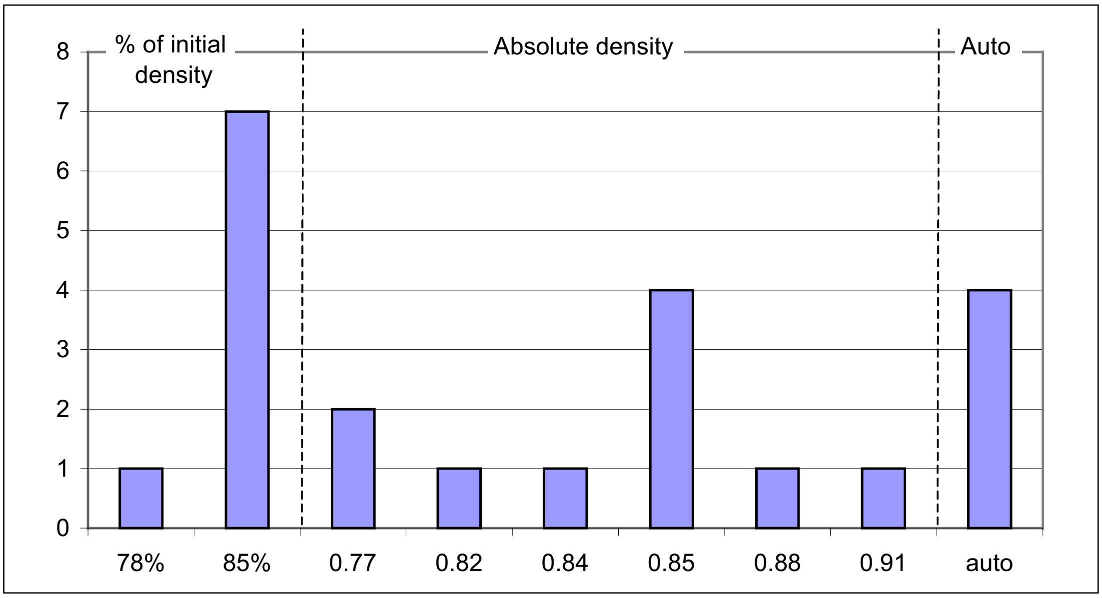
Most algorithms have only one parameter and most of the time it is set to a fixed value related to the minimum size constraints at the 1 : 50.000 scale.

Few algorithms can therefore be analysed along their parameter setting.

5.7.5.3 Parameter Values of “Building Removal”

The parameter in “building-removal” is the expected density after treatment.

This parameter can be set with three different definitions:
- as a percentage of the initial density
- as an absolute value (a black/white ratio on paper)
- as a automatically determined value, using the formula: min(0.84, 0.95*inital density). That means it is the minimum of an absolute density of 0.84 and a relative density of 0.95 times the initial density.



Graph 20 – Parameter of “building-removal” (test 5).

During the process, these three definitions were used to choose the conflict:
- They were used with an absolute value in high density areas for solving the “too high density” conflict.
- They were used as a percentage of the initial density or determined automatically in medium density areas for solving the “homogeneity of density” conflict, that is, to keep or enhance the difference of density between high density and medium density areas.

5.7.6 Visual Result

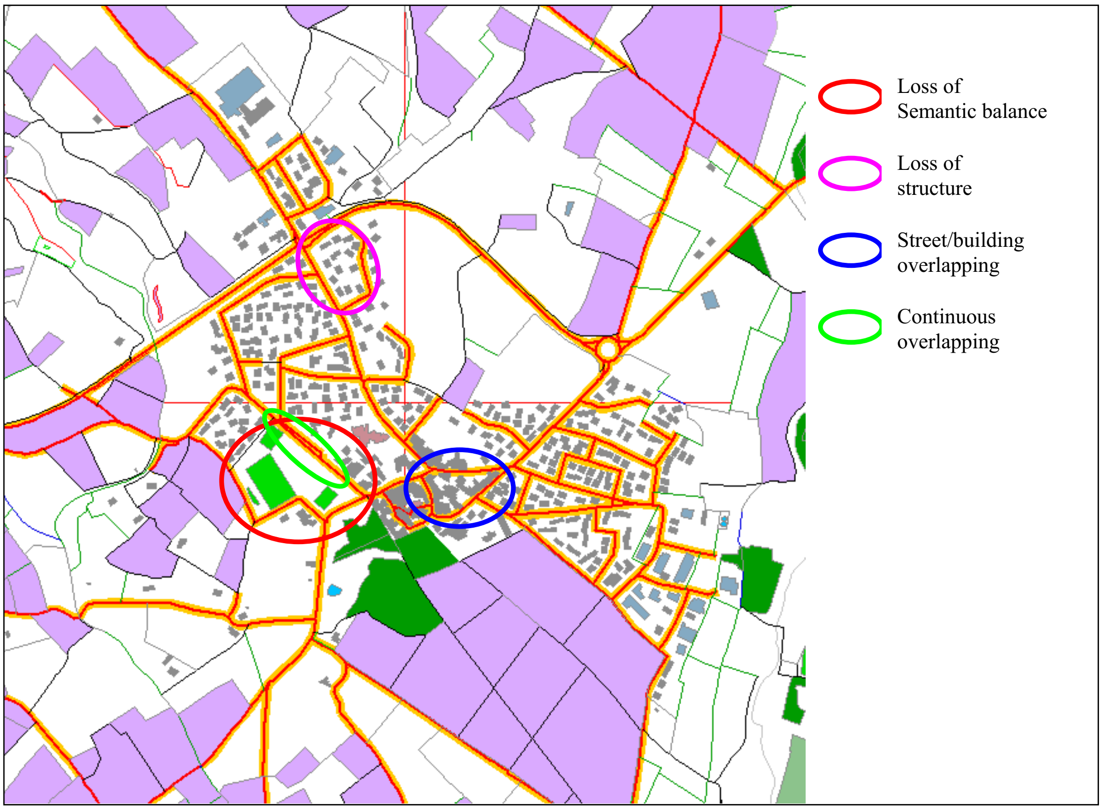


Figure 15 – Visual result (test 5).

Figure 15 analysis

– Overlapping conflicts remain in city center.
– Some buildings are still too close. The building removal could have been more stronger.
– In some blocks, the final distribution of buildings could be better.
– Rural areas are under-generalised (too many tracks).
– Dead-end removal is not perfect.

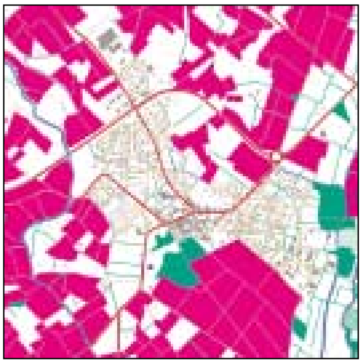
5.8 Global Analysis

The aim of a common analysis is not to compare the quality of the different GISs or generalisation for two reasons:

– Each platform could have produced better results if it had been used interactively.
– This test did not focuse on the results but on the method.

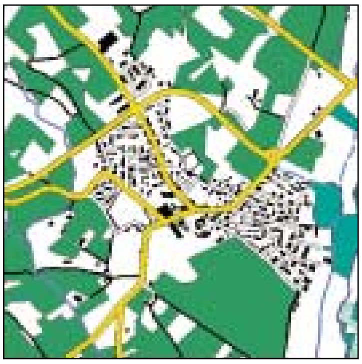
5.8.1 Visual Results

Initial data

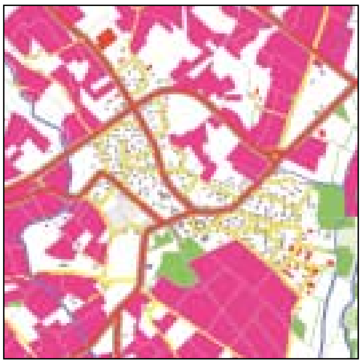


Test 1

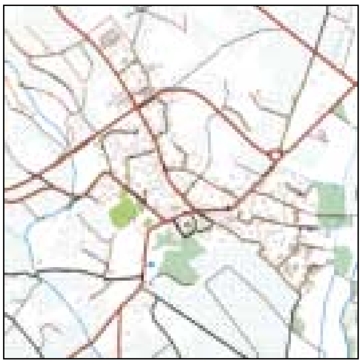
Test 2



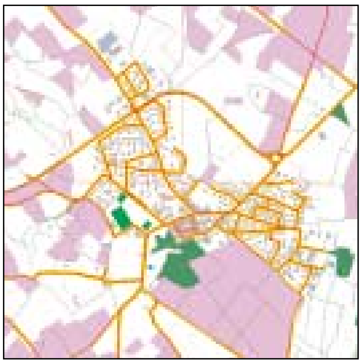
Test 3



Test 4



Test 5



5.8.2 Operations and Algorithms

We can first notice that the number of operations used is very different for such a small area. It does, however, show the quantity of interactive work required.

Table 42 – Comparison between tests. B /M: Bad / Medium.

	Test 1		**Test 2**		**Test 3**		**Test 4**		**Test 5**	
Nb of operations	405		122		300		35		308	
Nb of algorithms	268	66%	89	73%	205	68%	19	54%	199	65%
Nb of different algo	11		7		6		7		13 (+4)	
Gen. Assessment B | M	54%	24%			15%	10%	-	-	-	50%
Algo assessment B | M	2%	12%	-	49%	-	-	-	58%	1%	12%

The assessment was not filled out identically by all users, but overall we can consider that for 50% of the areas, the generalisation is not perfect. These areas are urban ones.

Table 43 – Visual operator assessment.

	CLASSIFICATION	OBJECT REMOVAL	TYPIFICATION	AGGREGATION	FUSION'	SIMPLIFICATION	SIMP + AGGREG	SMOOTHING	SQUARING	DILATION	DISPLACEMENT	ROTATION	COLLAPSE	SEGMENTATION	EVALUATION	TOTAL	
Bad result		1		4		8				55	4				1	73	7%
medium		21	20	28	1	45	1	1	30	20	14	1	4	2	10	198	18%
good	3	263	88	48	25	23		1	85	68	30	61	72		57	824	75%

– The assessment shows that even classical algorithms (such as simplification, squaring, dilation or aggregation) do not provide good results, which is surprising.

If we analyse the final conflict code, main conflicts are related to:

– Over proximity mainly between objects, between buildings and streets.
– Loss of relative position between buildings. The distribution of buildings is poorly maintained.
– The shape of objects after generalisation are not perfect. Squaring, building simplification and aggregation algorithms are either missing or imperfect.

Table 44 – Number and percentage of operations used.

Number /%	Test 1		Test 2		Test 3		Test 4		Test 5	
Classification			1	1			3	9		
Object removal	111	27	33	27	85	28	12	34	46	14
Aggregation	45	11	9	7			1	3	7	2
Fusion	24	6	1	1			1	3	19	6
Typification	71	17	11	9			4	11	41	13
Orientation	59	15			4	1				
Simplification	12	3	18	15	38	13	9	32		
Smoothing							2	6		
Squaring	7	2	15	12	97	32				
Caricature							1	3		
Dilation	2	0,5	31	25	67	22			44	14
Displacement					7	2,3			42	14
Collapse	74	18					2	5,7		
Object creation									41	13
Evaluation									68	22

In terms of operations and algorithms used, a large range were used globally, but unfortunately no GIS contains all of them. Processes are adapted according to the capacity of each system.

An analysis of each table which associates operations and algorithms shows that there is not a perfect association between what the user wishes to do (operation) and what she/he uses to achieve this (algorithm). This underlines the lack of appropriate generalisation algorithms, especially contextual generalisation algorithms.

Quantitatively, the main operations aim at reducing the quantity of objects, in particular buildings. This reduction of quantity is performed in different ways:

- either interactive removal
- or removal by means of size or classification filtering
- or 'typification' (see the remarks on typification)

The distribution of algorithms used for the 5 tests is shown in the following. More details are given in annexe 2. Some functions were grouped to allow comparison.

- Interactive operations: 23,7 % 86,3 % of interactive operations are object removal.
- Generalisation performed by algorithms:
- Object removal 18,1 % (including typification)
- Aggregation 8,2 %
- Simplification 17,2 %
- Dilation 16,1 %
- Squaring 13,1 %
- Displacement 5,4 %
- Orientation 6,6 %
- Collapse 8,5 %
- Data enrichment 12,2 %

Typification algorithms are used more for contextual object removal than for true typification.

1. The distribution is not preserved.
2. Cluster typification is usually performed on a small number of objects (generally 2): Point typification, cluster typification, remove and centre. This implies that it is not very contextual.

Because typification needs symbols with MGE, a lot of buildings collapse into symbols (tests 1 and 4). With Lamps2 (test 2) this operation is performed by an aggregation.

- In most of the building reduction operations the result is a reduction 2 → 1. The advantage of such operations is that it solves proximity conflicts and allows for a building dilation. Of course, this kind of local operation is not contextual, therefore, building patterns cannot be maintained.

Most of the GISs have no displacement algorithm, which explains the unsolved proximity conflicts. Even with StratËge, a complementary displacement algorithm between buildings is missing.

5.8.3 General Process

- The processes used are different, partly due to the incomplete algorithms of each system.

However, we can notice that:

- First operations are related to 'easy' object removal by means of size or type filtering (classification). Some automated fusions are performed, which corresponds to an initial unimportant object removal.
- Generally, operations are performed on working areas.
- Early operations aim at removing 'unnecessary objects' and reducing the quantity of buildings by means of typification.

– Then local geometric problems are solved (dilation, simplification).
– Finally, object removal is performed, mostly to correct problems, though sometimes to balance object quantities.
– Vegetation is generalised at the end of the process (aggregation, fusion).

Street removal is performed in a different order in different tests:

– At the beginning of the process	test 1; test 5
– Within urban areas, but before buildings	test 2; test 3
– At the end of the process, before vegetation	test 4

In terms of results, the use of street removal at the end of the process (test 4) is not really significant as the algorithm used (line typification based on Minimum Spanning Tree) is not contextual as it does not include other feature classes (such as buildings). The result would have been the same, had it been performed at the beginning.

Building generalisation

– Most of the time, important (large and thematic) buildings are generalised before houses.
– Building typification often damages their distribution. This is partially solved by some local orientations.
– Building dilation is always performed, either by a direct dilation or by changing polygons into symbols (solution with MGE). Often dilation provokes proximity or overlapping conflicts which are usually solved by an object removal.
– If building simplification and dilation are required, dilation is normally performed first.
– When squaring is available, it is the last generalisation operation on a building.

5.8.4 Conclusions

The different tables (table 32, table 35, table 38, table 39) presenting the operators and algorithms used to solve a conflict show that generalisation can be automated: the matrices are 'nearly' square. The global table mixing all operators, algorithms and conflicts does not give more information. In these tables we should remember that as some algorithms are missing, object removal is often the only solution available: we noticed that the associations are often better between conflicts and operations (what the user wishes to do) than between conflicts and algorithms (what the user did).

Our assumption is that an identical test performed on a GIS with all the necessary algorithms would prove the correlation between conflict/object and solution.

During the accurate analysis of the process templates, we noticed that whenever the user described several conflicts instead of one, the correlation between problems and solution was even higher.

However, these tests – assessed by traditional cartographers – would presumably be much more heavily criticised than by our own visual assessment. Certainly we know the difficulty of automation and we are evaluating the algorithms rather than the results. We also assume that not enough attention has been given to the loss of initial information in terms of density, distribution, orientation and shape.

Without relative comparisons, the size, shape and density become uniform. The building distributions (balance between quantity, proximity and orientation) is poorly maintained.

6 The ICC Test

The aim of this last test, described below, is to generalise buildings from ICC large scale database to produce a 1 : 25.000 scale map.

6.1 Participating Institutions

Five tests were performed on four platforms:

- MGE/MG (Intergraph Corporation) by *Monika Jordan* from the University of Munich
- Lamps 2 (Laser-Scan Limited) by *Lars Harrie* from the University of Lund
- PlaGe (prototype developed by the Cogit Laboratory) by *François Lecordix* from the Cogit laboratory.
- CHANGE (University of Hanover) by *Brigitte Hunsen* from the University of Hanover
- On CHANGE, MGE and ICC soft by *Maria Pla* and *Baella Blanca* from the ICC.

6.2 General Analysis

6.2.1 Number and Kind of Operations

- Hanover and the ICC used batch processes which explain the limited number of operations. The building generalisation was performed with the ANGI package. The ICC also generalised the others objects (wall, vegetation, etc.) which explains why they performed 40 operations.
- Lund generalised small groups of buildings, mainly with Corner-flipping
- Munich used mainly Clarification (and some aggregation and squaring)
- The IGN generalised buildings one by one (with Building-simp) and then displaced the buildings with Nickerson.

Table 45 – Number of operations and algorithms used.

Test	Number of operation	Main operations	Secondary operation
Hanover (CHANGE)	1	ANGI	
ICC (CHANGE MGE)	40	ANGI	Ortho-aggregation Clarification (MGE)
Lund (Lamps2)	143	Corner-flipping (87%)	Douglas Dispike
Munich (MGE)	543 (189 algorithms)	Clarification (73%)	Aggregation, Squaring
IGN (PlaGe)	660 (457 algorithms)	Building-simp (48%) Nickerson (41%)	Douglas++ (11%)

6.2.2 Algorithm Parameter Value

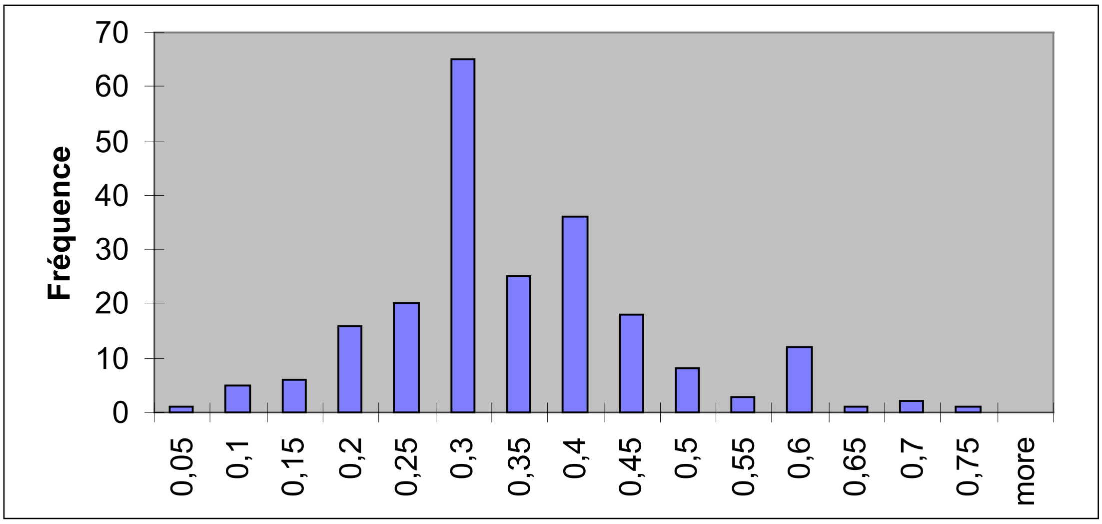
ANGI (CHANGE)

ANGI is the algorithm developed by the University of Hanover and used by Hanover and the ICC to simplify, square and caricature building shape. The parameter values are the same for all buildings and they integrate initial and final scales.

ANGI Parameter value
CONTOUR_1=Y
COMPREHEN=N
CONTOUR_2=N
E_SCALE=5000
TV_IDENT=0.02
TV_DIST=0.3
D_SCALE=25000
TV_LENGTH=0.5
TV_AREA=0.30
TV_COMPRE=0.30
TV_SHIFT=1.0
TV_ANGLE=40.0
Building-simp (PlaGe)

Building-simp is an algorithm developed by the COGIT laboratory to simplify square and caricature building shape. It has three parameter value:

- The minimum length value is used to remove smallest shape details and to emphasise medium ones. Parameter value unit is in map millimetre, at final scale. In graph 21 we can see that most of the value are between 0.2 and 0.6 mm (i.e. between 5 and 15 field metres)
- Two angle parameter values allow to flat lines and to remove small spikes. The following default values are generally used (in more than 95% of the cases)
  - Angle 1: 2,628 (95%) min = 0,523 max = 3,14
  - Angle 2: 0,392 (97%) min = 0,2 max = 0,645



Graph 21 – Building-simp parameter: minimum length value (mm at final scale).

Nickerson (PlaGe)

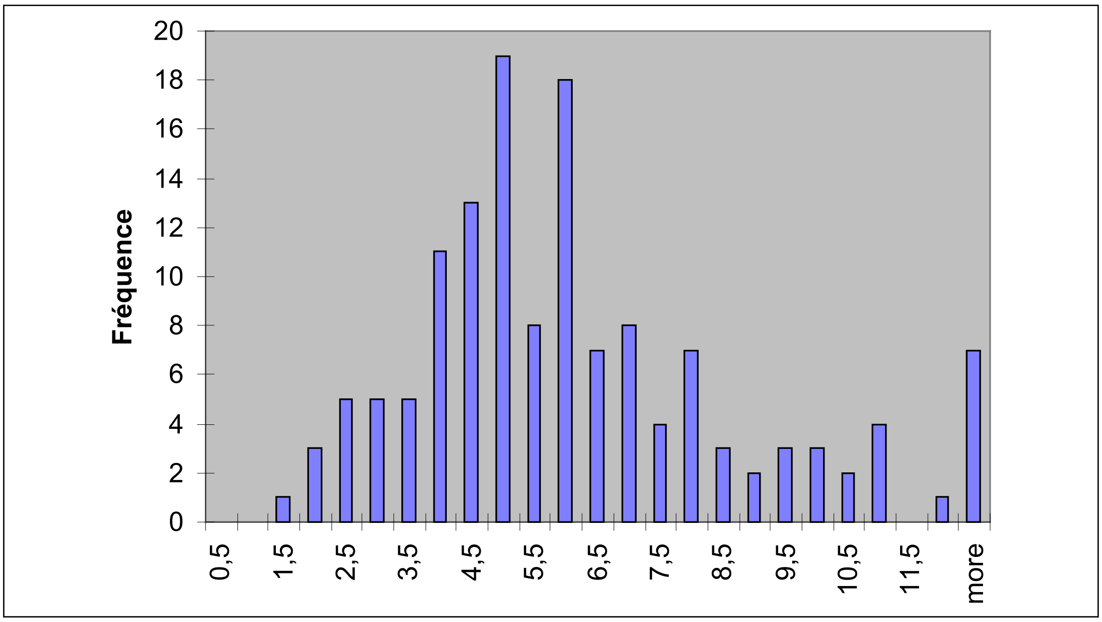
Nickerson algorithm has been used to 'displace' buildings from each other. Initially Nickerson algorithm is done for lines displacement. The parameter value is the minimum required distance between buildings so it is nearly always the same. The algorithm pushes the building boundary while absorbing the displacement perpendicularly and preserves each building own connectivity. For the building, it corresponds to a local erosion.

Here the aim of the operation is to create a larger free space that represents streets, even if the streets are not represented in the database. To automate the operation, it would have been necessary to create the streets (for example by means of the application of a Delaunay triangulation between buildings) and to use these streets to push automatically all the building boundary at a minimum distance to the street centre line.

Clarification (MGE)

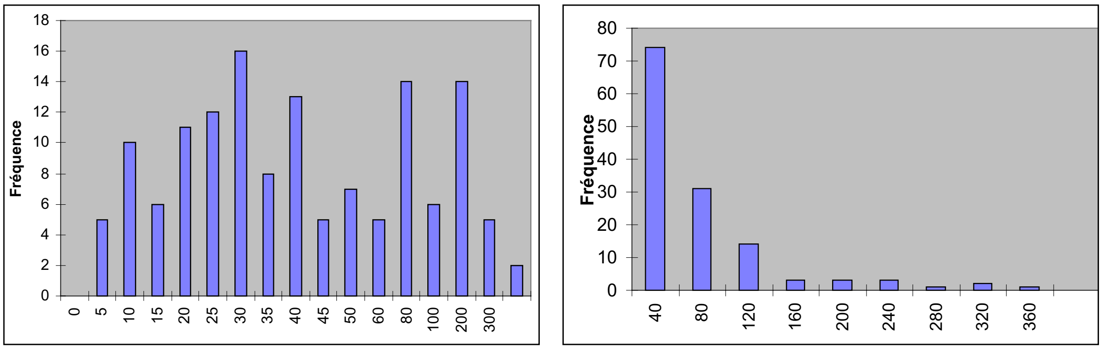
This algorithm has two possible behaviours: contour simplification or transformation into a rectangle according to both parameter values.

- The **length threshold** value is coherent to the final scale (around 5m, i.e. 0.2 mm). Moreover, the dispersion of values is rather low (graph 22).



Graph 22 – Clarification parameter: Minimum length value (meter).

- The **area threshold parameter** is used to replace the polygon by a rectangle if the required value is larger than the initial area. If this parameter value is low, it is nearly inactive and only the first parameter is significant. Consequently the distribution (graph 23 left view) shows that the user needed to reset each time the parameter value to avoid changing the polygon into a rectangle.



Graph 23 – Clarification parameter: minimum area value (m^2).

Corner-Flipping (Lamps2)

Three parameters (angle, long line and short line) allow the removal of shapes such as hats or stairs.

If one shape's longest dimension is smaller than 'long line' and the shortest dimension of the shape is smaller than 'short line', the shape is removed, otherwise the shape is preserved.

Table 46 – Description of Corner flipping algorithm statistical values.

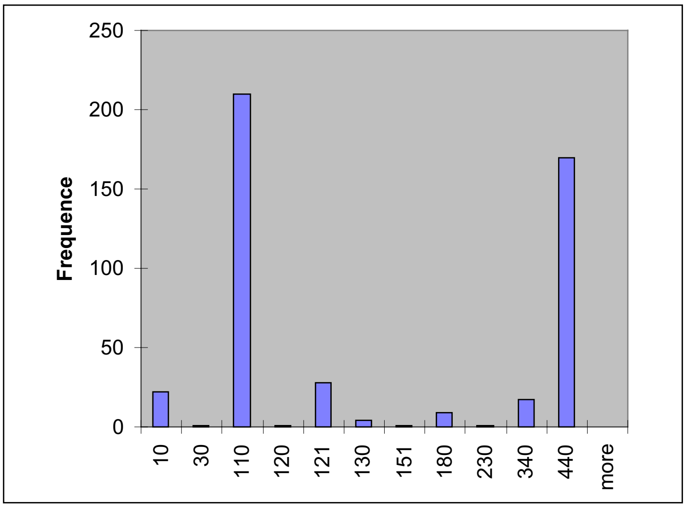
	Angle	long line	short line
Average	24,3	198	67
Median	30	185	69,3
More frequen value	30	101	87
Standard deviation	8,5	120	29,8
Max - Min	25,7	845	181
Minimum	4,3	50	0
Maximum	30	895	181

In table 46 we can see that if the angle is rather stable, the long line and short line values vary a lot. Consequently this algorithm is difficult to tune.

6.2.3 Conflicts

Conflicts that initiated generalisation

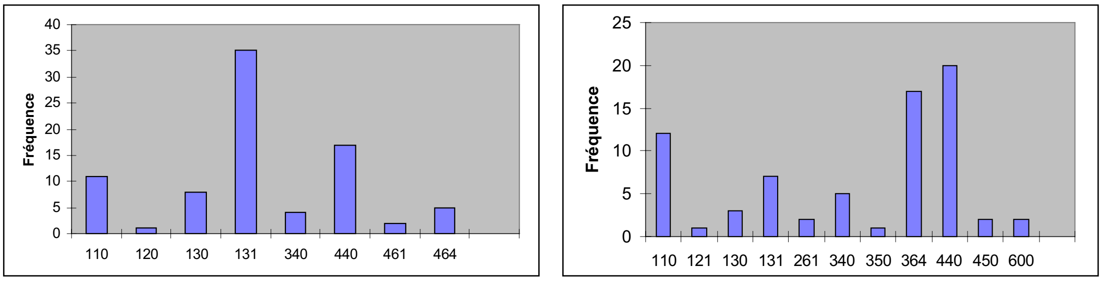
The conflicts that generated generalisation were the 'too high level of detail' of buildings (110) and their 'over proximity' (440) (graph 24 as an example).



Graph 24 – Initial Conflict code for IGN Test.

Remaining conflicts

Amongst critics, the users noticed that some buildings have not been enough generalised (110) or they have lost some of their shape properties (131). Moreover, whenever no displacement algorithms were available, the buildings remain too close to each other (440). In some tests some topological inconsistencies occurred (364 Munich and Lund).



Graph 25 – Final conflict code 1: IGN; 2: Munich.

- 110 Polygon too detailed
- 121 Polygon too small
- 131 Loss of shape property
- 340 Line and polygon too close
- 364 Loss of sharing geometry btw line and polygon
- 440 Polygons too close
- 461 Bad relative position between polygons
- 600 Semantic inconsistency

120 Polygon too thin
130 Polygon not squared enough
261 Loss of linear structure
350 Line and polygon overlapping
450 Polygon overlapping
464 Loss of polygon spatial structure

6.3 Visual Results

For this test we had four snapshots at final scale. At first sight we see that the level of generalisation is different even if the same initial criteria were given. The Lamps generalisation (Lund) is the strongest while the PlaGe (IGN) is the lightest. These differences are coming from cartographers own interpretation.

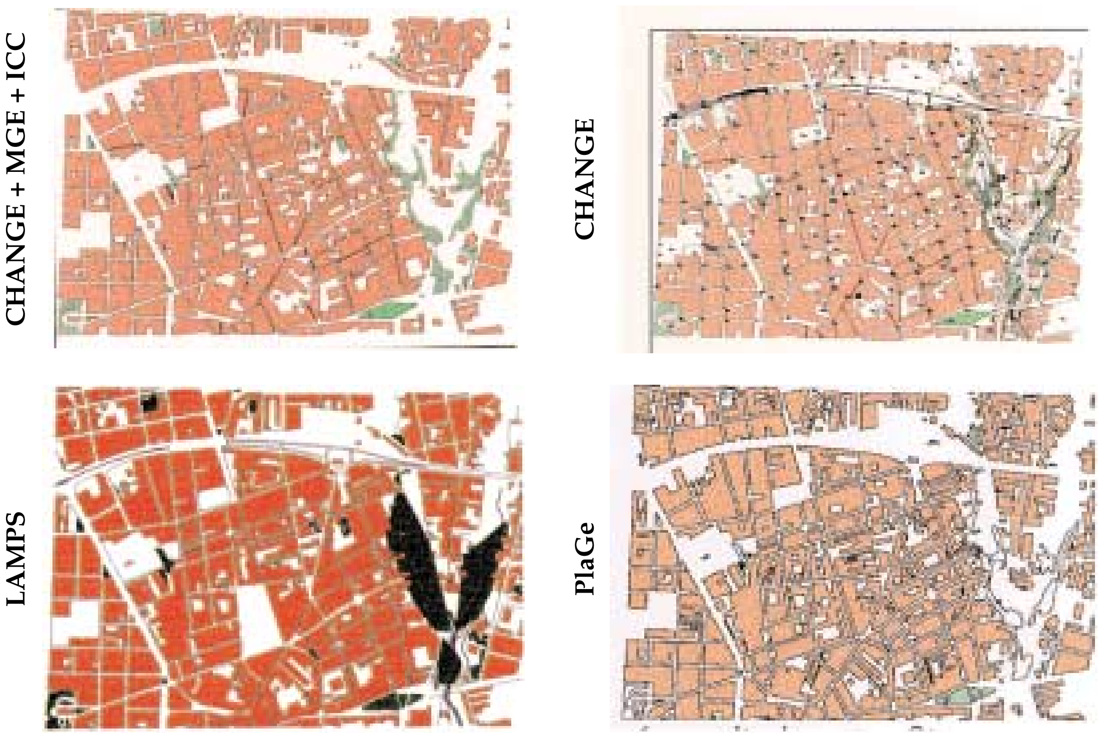


Figure 16 – The final generalised data on different platforms.

Globally, the results are rather correct, the main critic is the too high proximity between buildings even on the PlaGe generalisation, although an algorithm of displacement was used.

Main remaining problems

A more accurate analysis of the generalised data shows some bad quality building generalisation. Whatever the software used, a repetitive sequence of shapes is never well generalised because their is no identification of these sequences (figure 17). To generalise correctly such buildings, a shape typification operator would be necessary.

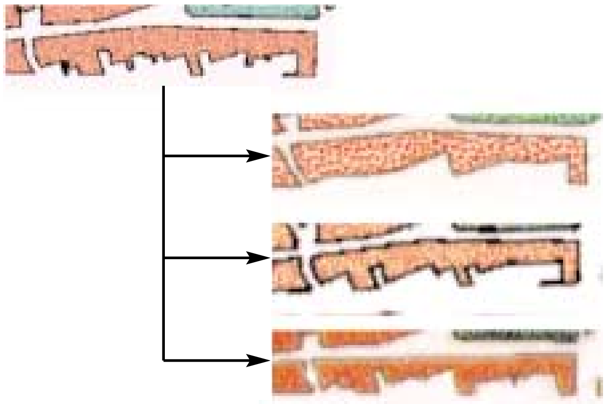


Figure 17 – Bad generalisation of repetitive sequence of shapes.

Moreover, proximity inside buildings due to small entrances or courtyards are never well generalised (figure 18). All the tested algorithms are based on consecutive points analysis so they can not identify such problems. Local displacement inside boundaries would be necessary. The only available solution is always to over generalise such shapes.

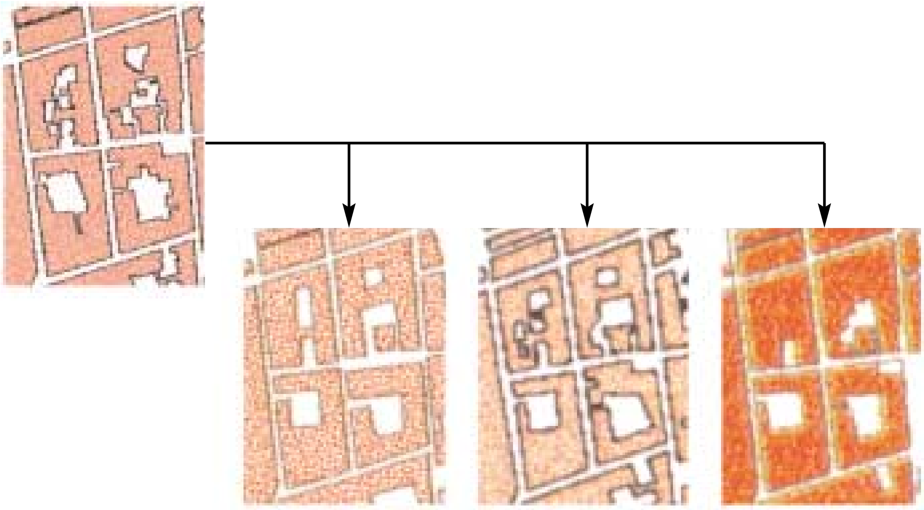


Figure 18 – Bad management of building self proximity.

6.4 Conclusion

Due to this scale range, it was not possible to analyse the sequence of generalisation that concerned only buildings. Buildings were generalised in the same way, sometimes with the same parameter values, sometimes not. In the last case, we can notice that parameter values are grouped around a value close to the readability threshold for this scale. Results can be considered as globally acceptable, which means that this scale change is not too difficult to generalise.

However, some main problems remain:

- some complex shapes composed of a sequences of small shapes (figure 20) are not generalised properly: either main directions are lost or the buildings are over generalised
- the minimum distance inside and between buildings are not respected (figure 21):
  - courtyards and small entrances are not open up,
  - buildings are not displaced from each other.

Certainly, an improvement would consist in adding local displacement inside and between building boundaries constrained by wall orientation.

7 Conclusions and Outlook

7.1 Main Conclusion of the Tests

The analysis of all these tests, together or independently, offers a large information on generalisation process.

In terms of quality, the best generalisation is the building one, because it is not very contextual. Moreover, buildings have more or less the same geometrical properties, which explains why batch solutions are possible and why parameter values are rather grouped. The only remaining problems are related to proximity conflicts between buildings due to the lack of displacement algorithms.

For the two other tests (urban and roads) the results are not so good, even interactively. This is due to a lack of algorithms.

For roads, it seams difficult to use a same set of algorithms on sinuous and heterogeneous lines. Best processes segment the line and apply locally a specific algorithm according to the segment shape and conflict. Moreover, some packages would need to add caricature or 'bend extend' algorithms in order to maintain the important road bends: a smoothing can not solve symbol overlapping while maintaining the shape of the bend.

For urban generalisation, the quality is rather poor. The absence of contextual algorithms to remove objects and to displace them made good generalisation nearly impossible, even if the users always managed to find tricks to solve as many problems as possible. The building patterns are lost, objects overlap each other.

In terms of processes we noticed a strong link between a kind of conflict and a proposed solution, which means that the automation is theoretically possible if we do manage to develop measurements to identify and qualify conflicts. However for urban generalisation, an object removal was often used either without conflict description or to solve a local conflict in the absence of other solution (by default). It seems that the over density conflict is not always perceived although it is a main conflict in generalisation. In the same way, 'non important object' was rarely used. Certainly some more work on contextual conflict description is necessary. We also enlightened some typical sequences used for road and urban generalisation that could be transformed into generalisation rules. However, the notion of constraints would have help to distinguish between solutions (consequently to enforce the rules). For example to solve a conflict, if two algorithms can theoretically be used, the best solution is the one that best preserves the geographical properties of the objects.

Finally during the analysis, I gave when necessary the parameter value of the algorithms. Whenever the values are grouped, it could be used either for automation or to provide the user with a meaningful default value. Whenever the values are not grouped, either it means that the choice depends on a property of the objects (and not only on scale change) that should be computed, or its means that the algorithm is not very well designed as the tuning is not predictable. These algorithms should be replaced by new ones.

We wish that these tests would lead to a general awareness of the difficulties encountered in automatic generalisation and the urgency of studying and solving these problems:

- a better description of the data content (spatial analysis),
- developing conflict detection tools with respect to user needs and the property of geographical data,
- developing more contextual algorithms such as object contextual removal (according to flexible criteria) and object displacement,
- automating processes by learning techniques based on character and conflicts, and experimental tests.

7.2 Perspective

We believe that the methodology used for this test analysis is very cumbersome and after this first attempt, we should try to develop a more flexible and digital method to evaluate experiments. This would allow a more frequent and easier exchange of experiments.

Manual tracing is far too heavy and time consuming. If tracing automatically the sequences of action is easy (with few developments), it is still difficult to define a common way of describing objects or groups of object. It is also difficult for the user to clearly justify each of his decision during the process, and to record it. We believe that the research community should develop and share intelligent trace capacity in order to speed up research in automation.

To improve the efficacy of this test two points should have been changed:

- The symbolisation and minimum size should have been imposed to all users,
- We should have distinguished conflicts and constraints to explain the motivation of choosing an algorithm.

Moreover to speed up the test and the test analysis (3 years were necessary), an OEEPE seminar of a week would have been better to perform the test, but the difficulty would have been to move the platforms in a single place.

Amongst perspectives to learn more about generalisation we can propose different approaches:

- Comparing the algorithms related to one action only, on common sample data sets (see [*Duchêne* 2000]). This method could help improving algorithms and learning faster on appropriate algorithm parameter values,
- Proposing common data sets to be generalised in a very well defined way. The generalised data would be sent back with an optional file describing the sequences used in order to allow a process analysis too.
- Proposing generalised data sets to be evaluated by different students and comparing the evaluation methods and results.

All these data sets (samples, real data or generalised data) could be available on a web site and results could be analysed by the OEEPE working group. One of the priorities is really to convince teachers to put some human effort on evaluation by means of student projects and phD. It seems difficult to work on automation without working on evaluation: a system which does not contain appropriate measurement (to characterise and to evaluate) can not be successful. Ongoing researches on automatic learning for generalisation which are very promising (e.g. [*Zucker* et al 2000]) also rely on (and are limited by) the quality of the measures and the algorithms.

8 Literature

8.1 Reports of the Working Group

- *Baella, B.; Colomer, J. L.; Pla, M.* 94: Map Generalizer, A pratical experience. OEEPE internal report ICC.
- *Baella, B.; Colomer, J. L.; Pla, M.* 94: Change: automatic cartography. OEEPE internal report ICC.
- OEEPE 97: Generalisation algorithms templates. Rapport Technique Working group on generalisation.
- *Rousseau, D.; Rousseau, T.* and *Lecordix, F.* 94: Expertise de Map Generalizer: logiciel de généralisation interactive d'Intergraph. Rapport technique COGIT.
- *Weibel, R.* and *Ehrliholzer, R.* 95: An evaluation of MGE Map Generalizer. Rapport technique GIUZ Zurich Suisse.

8.2 References

- *Dassonville, L.* 99: Quality control, data quality and users, metadata for geographical information. Proceedings of the International Symposium on Spatial Data Quality'99. Hong Kong, *Shi, Ed. W.; Goodchild, M. F.; Fisher, P. F.*, 18–20 July 99, pages 132–142.
- *Duchêne, C.* 2000: Methods of Evaluating Generalisation Systems. GISScience 2000 & ACI 2001.
- *João, E. M.* 98: Causes and consequences of Map Generalisation. Thèse de Doctorat. London, School of Economics, *Taylor, Ed.* and *Francis.*
- *McMaster, R. B.* and *Veregin, H.* 97: Visualizing cartographic generalization. Autocarto'15. Seattle Etats-Unis, Volume 5, pages 174–183.
- *McMaster, R. B.* 83: Mathematical measures for the evaluation of simplified lines on maps. Thèse de doctorat Université du Kansas. Etats-Unis.
- *Weibel, R.* 96: A typology of constraints to line simplification. Spatial Data Handling'96. Delft Pays-Bas. Advances in GIS Research II, *Taylor, Ed.* and *Francis*, pages 9 A.1–9A.14.
- *Weibel, R.* and *Dutton, G.* 98: Constraint-based automated map generalization. Spatial Data Handling'98, Vancouver, Canada, pages 214–224.
- *Zucker, J.-D.; Mustière, S.; Saitta, L.* 2000: Learning Abstraction and Representation Knowledge: an Application to Cartographic Generalisation. 5ème International Workshop on Multi-Strategy Learning (MSL). Guimaraes-Portugal. 5–7 juin 2000.

Appendix 1: Operators and algorithms used for the 5 BDTopo tests.

	NO ALGORITHM	classification	remove-and-center (Strat)	Street-typification (Strat)	Dead-end-removal (Strat)	Small-area-removal (Strat)	Remova-holes (Strat)	point-to-point (MGE)	centroid (MGE)	line typification (MGE)	Building-removal (Strat)	Aggregation (LSC)	Merge (LSC)	Ortho-aggreg (MGE)	Aggreg-and-displace (Strat	Aggregation (MGE)	Meso-Fusion (Strat)	Boudary extend (MGE)	Douglas	Douglas-dispike (LSC)	Corner-flipping (LSC)	Clarification (MGE)	Angi (Change)	Simple-avergae	Squaring (MGE)	Squaring (LSC)	exagerrate (LSC)	Dilation (Strat)	Enlarge-in-rectangle (Strat)	Cluster-displacement (LSC	LDT-displacement (Strat)	Translate-dxy (strat)	Curvilinear-deform (Strat)	Point orientation (MGE)	Point / Line orientation (MG	Area-to-point (MGE)	Area-to-point (ICC)	Create-partition (Strat)	Create-reference (strat)	Evaluate-micro-const (Strat	Evaluate constrainte (Strat)	
			OBJ REMOVAL / TYPIFICAT.									AGGREGATION							SIMPLIFICATION						SQR		DILATION			DISPLACE				ORI.		COL.		ENRICHMNT				
classification		6																																								6
obj removal	240	1	18	1	29	16	1																																			306
typification								3	69	3	22	10	1																													108
aggregation	1											2	7	8	7	9	19	29																								82
fusion	24											1				1																										26
simplification	1																		10	43	12	10				1																77
simp +aggreg																							1																			1
smoothing																								2																		2
squaring	2																			1					7	109																119
caricature	1																																									1
dilation	2																										98	38	6													144
displacement	1																													6	35	6	1									49
rotation	4																																	35	24							63
collapse																																				74	2					76
segmentation	2																																									2
cree-meso																																						1				1
cree-objet																																							2	38		40
evaluation																																									68	68
Total	278	7	18	1	29	16	1	3	69	3	22	13	8	8	7	10	19	29	10	44	12	10	1	2	7	110	98	38	6	6	35	6	1	35	24	74	2	1	2	38	68	1171

LIST OF THE OEEPE PUBLICATIONS

State – March 2001

Official publications

1 *Trombetti, C.:* „Activité de la Commission A de l'OEEPE de 1960 à 1964" – *Cunietti, M.:* „Activité de la Commission B de l'OEEPE pendant la période septembre 1960 –j anvier 1964" – *Förstner, R.:* „Rapport sur les travaux et les résultats de la Commission C de l'OEEPE (1960–1964)" – *Neumaier, K.:* „Rapport de la Commission E pour Lisbonne" – *Weele, A. J. v. d.:* „Report of Commission F." – Frankfurt a. M. 1964, 50 pages with 7 tables and 9 annexes.

2 *Neumaier, K.:* „Essais d'interprétation de »Bedford« et de »Waterbury«. Rapport commun établi par les Centres de la Commission E de l'OEEPE ayant participé aux tests" – „The Interpretation Tests of »Bedford« and »Waterbury«. Common Report Established by all Participating Centres of Commission E of OEEPE" – „Essais de restitution »Bloc Suisse«. Rapport commun établi par les Centres de la Commission E de l'OEEPE ayant participé aux tests" – „Test »Schweizer Block«. Joint Report of all Centres of Commission E of OEEPE." – Frankfurt a. M. 1966, 60 pages with 44 annexes.

3 *Cunietti, M.:* „Emploi des blocs de bandes pour la cartographie à grande échelle – Résultats des recherches expérimentales organisées par la Commission B de l'O.E.E.P.E. au cours de la période 1959–1966" – „Use of Strips Connected to Blocks for Large Scale Mapping – Results of Experimental Research Organized by Commission B of the O.E.E.P.E. from 1959 through 1966." – Frankfurt a. M. 1968, 157 pages with 50 figures and 24 tables.

4 *Förstner, R.:* „Sur la précision de mesures photogrammétriques de coordonnées en terrain montagneux. Rapport sur les résultats de l'essai de Reichenbach de la Commission C de l'OEEPE" – „The Accuracy of Photogrammetric Co-ordinate Measurements in Mountainous Terrain. Report on the Results of the Reichenbach Test Commission C of the OEEPE." – Frankfurt a. M. 1968, Part I: 145 pages with 9 figures; Part II: 23 pages with 65 tables.

5 *Trombetti, C.:* „Les recherches expérimentales exécutées sur de longues bandes par la Commission A de l'OEEPE." – Frankfurt a. M. 1972, 41 pages with 1 figure, 2 tables, 96 annexes and 19 plates.

6 *Neumaier, K.:* „Essai d'interprétation. Rapports des Centres de la Commission E de l'OEEPE." – Frankfurt a. M. 1972, 38 pages with 12 tables and 5 annexes.

7 *Wiser, P.:* „Etude expérimentale de l'aérotiangulation semi-analytique. Rapport sur l'essai »Gramastetten«." – Frankfurt a. M. 1972, 36 pages with 6 figures and 8 tables.

8 „Proceedings of the OEEPE Symposium on Experimental Research on Accuracy of Aerial Triangulation (Results of Oberschwaben Tests)" *Ackermann, F.:* „On Statistical Investigation into the Accuracy of Aerial Triangulation. The Test Project Oberschwaben" – „Recherches statistiques sur la précision de l'aérotriangulation. Le champ d'essai Oberschwaben" – *Belzner, H.:* „The Planning. Establishing and Flying of the Test Field Oberschwaben" – *Stark, E.:* Testblock Oberschwaben, Programme I. Results of Strip Adjustments" – *Ackermann, F.:* „Testblock Oberschwaben, Program I. Results of Block-Adjustment by Independent Models" – *Ebner, H.:* Comparison of Different Methods of Block Adjustment" – *Wiser, P.:* „Propositions pour le traitement des erreurs non-accidentelles" – *Camps, F.:* „Résultats obtenus dans le cadre du project Oberschwaben 2A" – *Cunietti, M.;*

Vanossi, A.: „Etude statistique expérimentale des erreurs d'enchaînement des photogrammes" – *Kupfer, G.:* „Image Geometry as Obtained from Rheidt Test Area Photography" – *Förstner, R.:* „The Signal-Field of Baustetten. A Short Report" – *Visser, J.; Leberl, F.; Kure, J.:* „OEEPE Oberschwaben Reseau Investigations" – *Bauer, H.:* „Compensation of Systematic Errors by Analytical Block Adjustment with Common Image Deformation Parameters." – Frankfurt a. M. 1973, 350 pages with 119 figures, 68 tables and 1 annex.

9 *Beck, W.:* „The Production of Topographic Maps at 1 : 10,000 by Photogrammetric Methods. – With statistical evaluations, reproductions, style sheet and sample fragments by Landesvermessungsamt Baden-Württemberg Stuttgart." – Frankfurt a. M. 1976, 89 pages with 10 figures, 20 tables and 20 annexes.

10 „Résultats complémentaires de l'essai d'«Oberriet» of the Commission C de l'OEEPE – Further Results of the Photogrammetric Tests of «Oberriet» of the Commission C of the OEEPE"

Hárry, H.: „Mesure de points de terrain non signalisés dans le champ d'essai d'«Oberriet» – Measurements of Non-Signalized Points in the Test Field «Oberriet» (Abstract)" – *Stickler, A.; Waldhäusl, P.:* „Restitution graphique des points et des lignes non signalisés et leur comparaison avec des résultats de mesures sur le terrain dans le champ d'essai d'«Oberriet» – Graphical Plotting of Non-Signalized Points and Lines, and Comparison with Terrestrial Surveys in the Test Field «Oberriet»" – *Förstner, R.:* „Résultats complémentaires des transformations de coordonnées de l'essai d'«Oberriet» de la Commission C de l'OEEPE – Further Results from Co-ordinate Transformations of the Test «Oberriet» of Commission C of the OEEPE" – *Schürer, K.:* „Comparaison des distances d'«Oberriet» – Comparison of Distances of «Oberriet» (Abstract)." – Frankfurt a. M. 1975, 158 pages with 22 figures and 26 tables.

11 „25 années de l'OEEPE"

Verlaine, R.: „25 années d'activité de l'OEEPE" – „25 Years of OEEPE (Summary)" – *Baarda, W.:* „Mathematical Models." – Frankfurt a. M. 1979, 104 pages with 22 figures.

12 *Spiess, E.:* „Revision of 1 : 25,000 Topographic Maps by Photogrammetric Methods." – Frankfurt a. M. 1985, 228 pages with 102 figures and 30 tables.

13 *Timmerman, J.; Roos, P. A.; Schürer, K.; Förstner, R.:* On the Accuracy of Photogrammetric Measurements of Buildings – Report on the Results of the Test "Dordrecht", Carried out by Commission C of the OEEPE. – Frankfurt a. M. 1982, 144 pages with 14 figures and 36 tables.

14 *Thompson C. N.:* Test of Digitising Methods. – Frankfurt a. M. 1984, 120 pages with 38 figures and 18 tables.

15 *Jaakkola, M.; Brindöpke, W.; Kölbl, O.; Noukka, P.:* Optimal Emulsions for Large-Scale Mapping – Test of "Steinwedel" – Commission C of the OEEPE 1981–84. – Frankfurt a. M. 1985, 102 pages with 53 figures.

16 *Waldhäusl, P.:* Results of the Vienna Test of OEEPE Commission C. – *Kölbl, O.:* Photogrammetric Versus Terrestrial Town Survey. – Frankfurt a. M. 1986, 57 pages with 16 figures, 10 tables and 7 annexes.

17 *Commission E of the OEEPE:* Influences of Reproduction Techniques on the Identification of Topographic Details on Orthophotomaps. – Frankfurt a. M. 1986, 138 pages with 51 figures, 25 tables and 6 appendices.

18 *Förstner, W.:* Final Report on the Joint Test on Gross Error Detection of OEEPE and ISP WG III/1. – Frankfurt a. M. 1986, 97 pages with 27 tables and 20 figures.

19 *Dowman, I. J.; Ducher, G.:* Spacelab Metric Camera Experiment – Test of Image Accuracy. – Frankfurt a. M. 1987, 112 pages with 13 figures, 25 tables and 7 appendices.

20 *Eichhorn, G.:* Summary of Replies to Questionnaire on Land Information Systems – Commission V – Land Information Systems. – Frankfurt a. M. 1988, 129 pages with 49 tables and 1 annex.

21 *Kölbl, O.:* Proceedings of the Workshop on Cadastral Renovation – Ecole polytechnique fédérale, Lausanne, 9–11 September, 1987. – Frankfurt a. M. 1988, 337 pages with figures, tables and appendices.

22 *Rollin, J.; Dowman, I. J.:* Map Compilation and Revision in Developing Areas – Test of Large Format Camera Imagery. – Frankfurt a. M. 1988, 35 pages with 3 figures, 9 tables and 3 appendices.

23 *Drummond, J.* (ed.): Automatic Digitizing – A Report Submitted by a Working Group of Commission D (Photogrammetry and Cartography). – Frankfurt a. M. 1990, 224 pages with 85 figures, 6 tables and 6 appendices.

24 *Ahokas, E.; Jaakkola, J.; Sotkas, P.:* Interpretability of SPOT data for General Mapping. – Frankfurt a. M. 1990, 120 pages with 11 figures, 7 tables and 10 appendices.

25 *Ducher, G.:* Test on Orthophoto and Stereo-Orthophoto Accuracy. – Frankfurt a. M. 1991, 227 pages with 16 figures and 44 tables.

26 *Dowman, I. J.* (ed.): Test of Triangulation of SPOT Data – Frankfurt a. M. 1991, 206 pages with 67 figures, 52 tables and 3 appendices.

27 *Newby, P. R. T.; Thompson, C. N.* (ed.): Proceedings of the ISPRS and OEEPE Joint Workshop on Updating Digital Data by Photogrammetric Methods. – Frankfurt a. M. 1992, 278 pages with 79 figures, 10 tables and 2 appendices.

28 *Koen, L. A.; Kölbl, O.* (ed.): Proceedings of the OEEPE-Workshop on Data Quality in Land Information Systems, Apeldoorn, Netherlands, 4–6 September 1991. – Frankfurt a. M. 1992, 243 pages with 62 figures, 14 tables and 2 appendices.

29 *Burman, H.; Torlegård, K.:* Empirical Results of GPS – Supported Block Triangulation. – Frankfurt a. M. 1994, 86 pages with 5 figures, 3 tables and 8 appendices.

30 *Gray, S.* (ed.): Updating of Complex Topographic Databases. – Frankfurt a. M. 1995, 133 pages with 2 figures and 12 appendices.

31 *Jaakola, J.; Sarjakoski, T.*: Experimental Test on Digital Aerial Triangulation. – Frankfurt a. M. 1996, 155 pages with 24 figures, 7 tables and 2 appendices.

32 *Dowman, I.*: The OEEPE GEOSAR Test of Geocoding ERS-1 SAR Data. – Frankfurt a. M. 1996, 126 pages with 5 figures, 2 tables and 2 appendices.

33 *Kölbl, O.*: Proceedings of the OEEPE-Workshop on Application of Digital Photogrammetric Workstations. – Frankfurt a. M. 1996, 453 pages with numerous figures and tables.

34 *Blau, E.; Boochs, F.; Schulz, B.-S.*: Digital Landscape Model for Europe (DLME). – Frankfurt a. M. 1997, 72 pages with 21 figures, 9 tables, 4 diagrams and 15 appendices.

35 *Fuchs, C.; Gülch, E.; Förstner, W.*: OEEPE Survey on 3D-City Models.

Heipke, C.; Eder, K.: Performance of Tie-Point Extraction in Automatic Aerial Triangulation. – Frankfurt a. M. 1998, 185 pages with 42 figures, 27 tables and 15 appendices.

36 *Kirby, R. P.*: Revision Measurement of Large Scale Topographic Data.

Höhle, J.: Automatic Orientation of Aerial Images on Database Information.

Dequal, S.; Koen, L. A.; Rinaudo, F.: Comparison of National Guidelines for Technical and Cadastral Mapping in Europe ("Ferrara Test") – Frankfurt a. M. 1999, 273 pages with 26 figures, 42 tables, 7 special contributions and 9 appendices.

37 *Koelbl, O.* (ed.): Proceedings of the OEEPE – Workshop on Automation in Digital Photogrammetric Production. – Frankfurt a. M. 1999, 475 pages with numerous figures and tables.

38 *Gower, R.*: Workshop on National Mapping Agencies and the Internet

Flotron, A.; Koelbl, O.: Precision Terrain Model for Civil Engineering. – Frankfurt a. M. 2000, 140 pages with numerous figures, tables and a CD.

The publications can be ordered using the electronic orderform of the OEEPE website

www.oeepe.org

or directly from

Bundesamt für Kartographie und Geodäsie
Abt. Geoinformationswesen

Richard-Strauss-Allee 11
D-60598 Frankfurt am Main