

STATE-OF-THE-ART OF AUTOMATED GENERALISATION IN COMMERCIAL SOFTWARE



TABLE OF CONTENTS			
ACKNOWLEDGEMENTS	3		
ABSTRACT	4		
EXECUTIVE SUMMARY	5		
CHAPTER 1 PRESENTATION OF THE PROJECT	15		
1.1 Introduction	15		
1.2 Previous research related to map specifications for automated map generalisation	16		
1.3 Scope of the current study	17		
1.4 Project set up	17		
CHAPTER 2 METHODOLOGY	19		
2.1 Requirement analysis	19		
2.1.1 Selecting the test cases	19		
2.1.2 Formulation of NMA map specifications for automated generalisation	20		
2.1.3 Harmonising NMA map specifications for automated generalisation	21		
2.1.4 Analysing the test cases	23		
2.2 The test process	24		
2.3 Evaluation of system capabilities, test processes and constraint expressions	25		
2.4 Automated constraint-based evaluation	26		
2.4.1 Automated constraint-based evaluation	26		
2.4.2 Computing outputs	27		
2.4.3 Visually comparing focus zones	27		
2.4.4 Quantifying differences	28		
2.4.5 Expert evaluation	28		
CHAPTER 3 RESULTS AND INTERPRETATION	31		
3.1 Outputs of the tests	31		
3.2 Evaluation of the capabilities of the systems	32		
3.2.1 Details on the completed system templates	32		
3.2.2 Generalisation functionalities of the systems	33		
ArcGIS	33		
Main generalisation functionalities	33		
Summary of the system template provided by the testers for ArcGIS	36		
Advantages and limitations	39		
Change, Push and Typify (CPT)	39		
Main generalisation functionalities	39		
Advantages and limitations	40		
Push	40		
Main generalisation functionalities	40		
Advantages and limitations	41		
Typify	41		
Main generalisation functionalities	41		
Advantages and limitations	42		
Summary of the system template provided by the testers for Change, Push, Typify	42		
Advantages and limitations for CPT software	46		
Radius Clarity	46		
Main generalisation functionalities	46		
Summary of the system template provided by the testers for Radius Clarity	53		
Advantages and limitations	53		
arqand	53		
Main generalisation functionalities	53		
Summary of the system template provided by the testers for arqand	54		
Advantages and limitations	57		
3.2.3 Summary of the capabilities of the systems	57		
General capabilities	57		
Quality of the available operators	60		
3.2.4 Conclusions for the capabilities of tested software systems based on testers' information	60		
3.3 Evaluation of test processes	64		
3.4 Evaluation of constraint expressions	66		
3.5 Automated evaluation of generalised outputs: results and conclusions	68		
3.5.1 Automated constraint-based evaluation of interactively generalised data	69		
3.5.2 Automated constraint-based evaluation of automatically generalised data	70		
Automated evaluation of minimum area of buildings	71		
Minimum area of buildings: ICC test case	73		
Minimum area of buildings: IGN test case	76		
Minimum area of buildings: Kadaster test case	77		
Minimum area of buildings: OSGB test case	78		
Findings and conclusions of automated evaluation of minimum area constraint	79		
Automated evaluation of minimum distance between two buildings	80		
Minimum distance between two buildings: ICC test case	82		
Minimum distance between two buildings: IGN test case	83		
Minimum distance between two buildings: Kadaster test case	85		
Minimum distance between two buildings: OSGB test case	86		
Automated evaluation of minimum distance between buildings and roads	87		
Minimum distance between buildings and roads respecting different road type: ICC test case	88		
Minimum distance between buildings and roads respecting different road type: IGN test case	89		
Minimum distance between buildings and roads respecting different road type: Kadaster test case	90		
Minimum distance between buildings and roads respecting different road type: OSGB test case	91		
Findings and conclusions of automated evaluation of minimum distance between buildings and roads	91		
3.6 Evaluating generalisation outputs by comparing generalised outputs: results and conclusions	92		
3.6.1 ICC - Town centre blocks and street representation (selection, aggregation)	94		
3.6.2 ICC - Coastline generalisation	99		
3.6.3 ICC - Generalisation of complex junctions	105		
3.6.4 ICC - Generalisation of suburban buildings	108		
3.6.5 ICC - Parallelism between roads and buildings	112		
3.6.6 IGNF - Buildings (selection, intersection, area)	115		
3.6.7 IGNF - Mountainous roads (coalescence to break series)	118		
3.6.8 IGNF - Vegetation selection and generic simplification	121		
3.6.9 IGNF - Ski lifts representation	124		
3.6.10 Kadaster - Channel network selection	127		
3.6.11 Kadaster - Settled area, building selection	131		
3.6.12 Kadaster - Generalisation of parallel roads and cycle tracks	136		
3.6.13 Kadaster - Railway simplification	141		
3.6.14 OSGB - Adjacent buildings representation (aggregation, simplification, shape preservation)	143		
3.6.15 OSGB - Detached buildings and fences	146		
3.6.16 OSGB - Dual carriageway representation, roads parallelism preservation	149		
3.6.17 Main findings and conclusions of evaluation by comparing outputs	152		
3.7 Expert evaluation: results and conclusions	153		
3.7.1 Details on the responses of the expert evaluation	154		
3.7.2 Global expert evaluation	155		
Frequency of responses' scores	155		
Harmonised respondents' scores	160		
3.7.3 Detailed constraint-based expert evaluation	162		
Evaluation of constraints for individual objects	163		
Evaluation of constraints for two similar objects	166		
Evaluation of constraints for two different objects	168		
Evaluation of constraints for groups of objects	171		
3.7.4 Examples of work, badly and differently solved constraints	173		
3.7.5 Ranking the generalised outputs	174		
3.7.6 Main findings and conclusions of expert evaluation	176		
CHAPTER 4 VENDORS' SOLUTIONS	179		
4.1 Vendors' tests	179		
ArcGIS	179		
ChangePushTypify	182		
IGN test case	182		
OSGB test case	183		
ICC test case	184		
Kadaster test case	185		
Radius Clarity	186		
arqand	189		
Conclusions on the vendors' tests	189		
4.2 Developments since 2007 and references to examples from practice	198		
ArcGIS	198		
ChangePushTypify	199		
Radius Clarity	199		
arqand	199		
CHAPTER 5 CONCLUSIONS AND DISCUSSION	193		
5.1 Answers to research questions	193		
5.1.1 What are the possibilities and limitations of commercial software systems for automated generalisation with respect to NMA requirements?	193		
5.1.2 What different generalisation solutions can be obtained for one test case and why do they differ?	195		
5.2 Conclusions and further research	195		
5.2.1 Defining map specifications as constraints	196		
5.2.2 Formulating and evaluating generalisation specifications	196		
5.2.3 Computing-based generalisation	197		
5.2.4 Evaluating generalisation software beyond constraints	198		
5.2.5 Concluding remarks	198		
REFERENCES	200		
APPENDICES	204		
Appendix I Visualisations of initial data (Only digital)	205		
Appendix II Symbol descriptions of output maps to be obtained by the tests (Only digital)	205		
Appendix III Harmonised constraints for one object	206		
Appendix IV Harmonised constraints for two objects	208		
Appendix V Harmonised constraints for group of objects	210		
Appendix VI NMA specific constraint sets (Only digital)	212		
Appendix VII Visualisations of output data produced by project team (Only digital)	212		
Appendix VIII Visualisations of output data produced by vendors (Only digital)	212		
Appendix IX Completed system templates	213		
Appendix X Quality of available generalisation operators	219		
Appendix XI Responses and analysis of expert evaluation (Only digital)	220		

Final report available from EuroSDR website

www.eurosdrr.net

Objectives



- To study:
 - capabilities/limitations of commercial software systems for automated generalisation with respect to NMA requirements
 - what different generalisation solutions can be generated for one test case and why do they differ?

What did we do?



- Requirement analysis Oct 2006 till June 2007
- Testing June 2007 till Spring 2008
- Evaluation Summer 2008 till Spring 2009
- Finalising the project Autumn 2009

Test cases



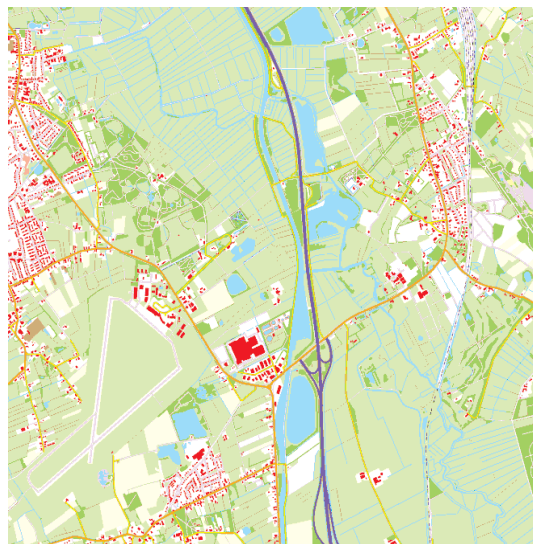
Area type	Source dataset	Target dataset	Provided by	Nr input	Main layers
Urban area	1:1250	1:25k	OS GB	37	buildings, roads, river, relief
Mountainous area	1:10k	1:50k	IGN France	23	village, river, land use
Rural area	1:10k	1:50k	Kadaster, NL	29	small town, land use, planar partition
Costal area	1:25k	1:50k	ICC Catalonia	74	village, land use (not mosaic), hydrography



ICC, 1:25k



IGN France, 1:10K



Kadaster, 1:10k



OS GB, 1:1250

One of the results: harmonised requirements

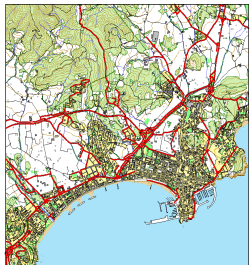


- 45 generic constraints:
 - 21 generic constraints on one object
 - 11 constraints on two objects
 - 13 constraints on group of objects
- About 300 constraints are defined as specialisations of generic constraints

Tests



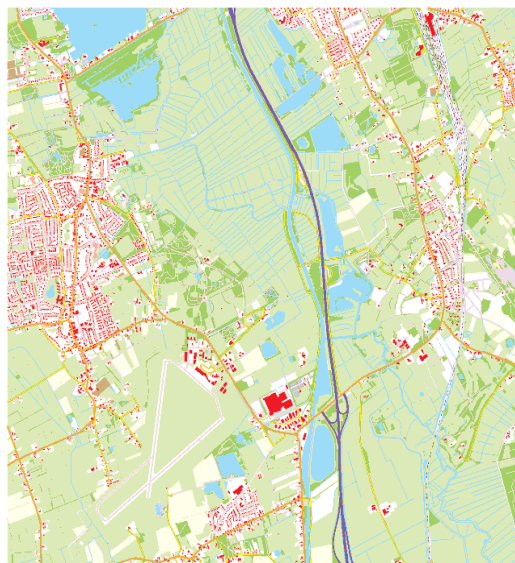
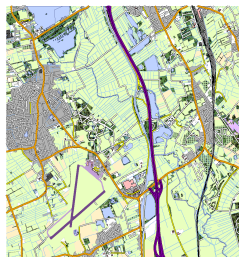
- Were performed:
 - by project team members on out-of-the-box versions
 - by vendors (1Spatial, ESRI, University of Hanover, Axes systems), possibly on improved and/or customized versions
- 35 test outputs were obtained (appr 700 thematic layers). NB: 1 test cost appr 1 week



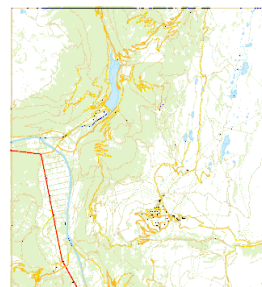
1:50K, derived from 1:25K, ICC



1:25K, derived from 1:1250, OSGB



1:50K, derived from 1:10K Kadaster



1:50K, derived from 1:10K, IGN, France

Evaluation



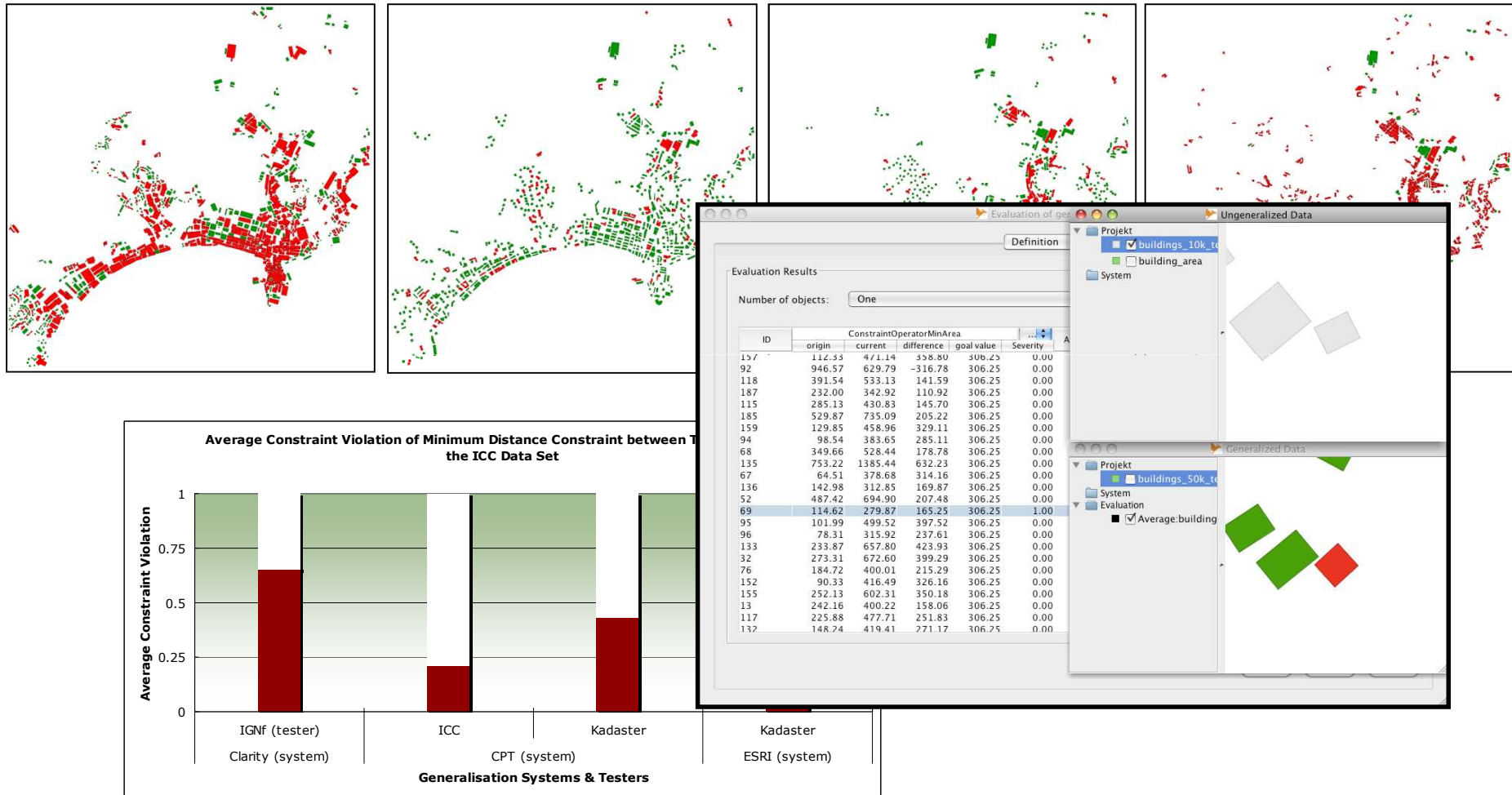
- Evaluation of:
 - System capabilities (based on completed system templates)
 - Processing (based on actions templates)
 - Constraint expression (based on constraint expression templates)
- Evaluation of generalised data:

Evaluation of generalised outputs, three methods



- Automated constraint-based evaluation
Dirk Burghardt, Stefan Schmidt, University of Zurich
- Evaluation which visually compared different outputs for one test case
Cecile Duchene, IGN France
- Qualitative evaluation by cartographic experts
Connie Blok, Jantien Stoter, ITC

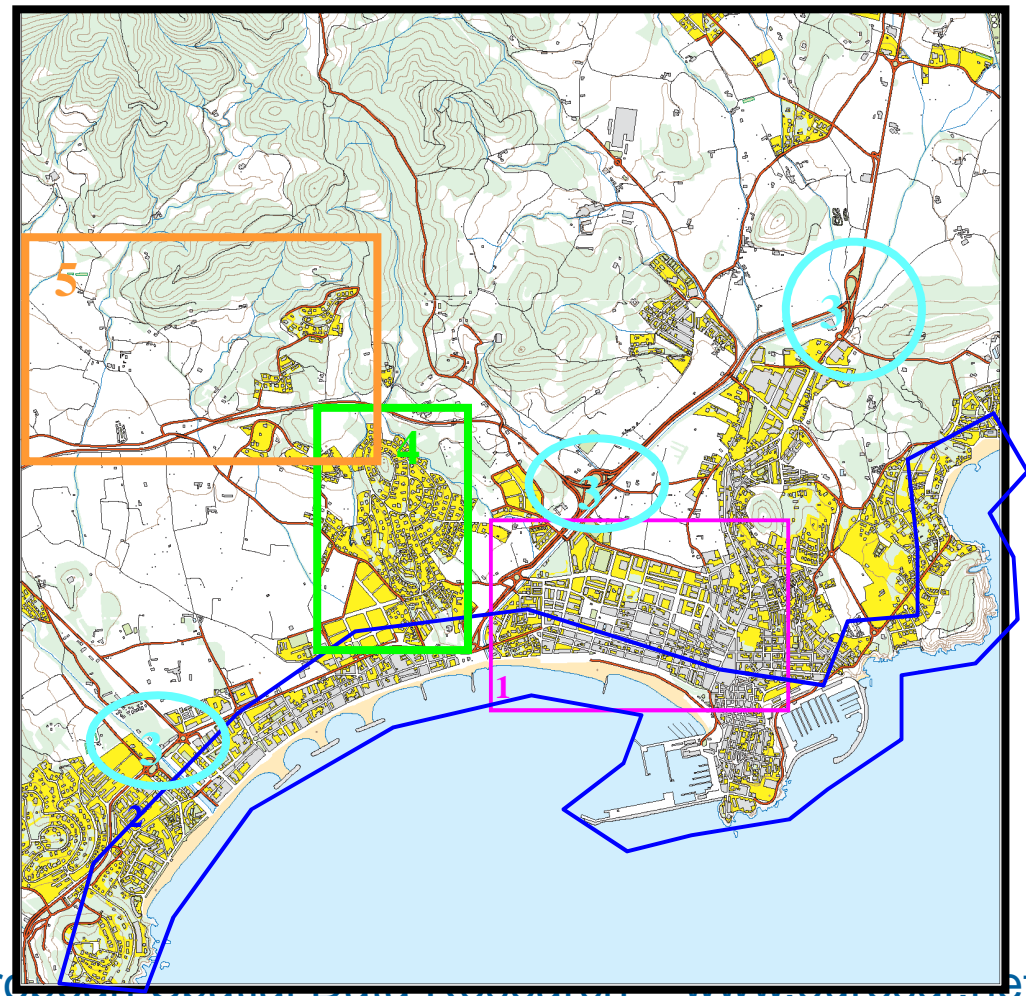
Automated constraint based evaluation of generalised data



Comparison evaluation of 16 focus zones



1. Town centre blocks and streets representation (selection, aggregation)
2. Coastline simplification
3. Conflicts in road interchanges
4. Generalization of suburban buildings (namely: preservation of buildings spatial distribution, buildings alignments)
5. Parallelism between roads and buildings



Descriptive sheet of each focus zone (16) included in final report



3.6.3 ICC - focus zone 3

Problem focused on:
Generation of complex junctions

Initial data:

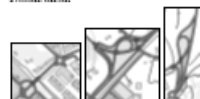


Figure 3.6.3.1 Initial data

What was expected according to the specifications:

Concerned data:

- Roads (line)

Defined constraints and rules (summary):

ICC-2-3, ICC-3-9.

Between two roads	: Minimum distance
Interchange (composed of roads)	: Minimum distance; displacement required; or simplification of shape if not enough space

For information, extract of the ICC paper map

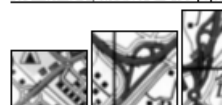


Figure 3.6.3.2 Extract of ICC paper map

Origins on this zone:

See next page.

Comments/questions with respect to the origins:

None.

Results of comparison and evaluation:

Interchange generation: CPI enables to displace roads away from each other thanks to the last query adjustment of PUSH. It results, in all the outputs obtained with CPI, in generated interchanges where the proximity conflict have been solved by displacement. This results in big displacement and parallelism for from each other (sometimes even further than the required

0.2mm). In some cases, in which probably have been better to remove some range (e.g. lower right ramp on the final extract as done in the ICC paper map indeed), but CPI doesn't enable to type interchanges. Some self-intersections have also been generated by CPI, due to a limitation of PUSH. The underestimation decreasing the density of vertices to avoid it but then it appears a small part of ramp where the vertices density causes the shape to keep it round aspect.

Clarity and ArcGIS do not provide tool for roads displacement neither tool for interchange typification, thus no changes have been done. Besides interchange generation in the result obtained by the Kadaster test on ArcGIS some road segments are missing resulting in important losses of connectivity in the network. It might be due to the elimination of short road segments performed thanks to an SQL query to satisfy a constraint on minimum length of a cut-de-ramp road segment (constraint ICC-1-45). This constraint was expected to only concern cut-de-ramp, but the SQL query may have selected these segments as well for any scenario (e.g. segment connected to road segment from another class, thus wrongly assessed as cut-de-ramp).

Less complex junctions: The generation of roundabout and branching roads is bad in all the outputs (last version) because, possibly, they have not been generated at all. This might be because of the form of the notion of "interchange" although in the specifications the term "interchange" was intended for all kind of junctions as soon as they are not simple crossroads, some tests might have interpreted as concerning "complex interchanges" only. As the interchanges are not made explicit under the form of "interchange object" in the ICC law, the doubt was possible. On they have not generated them because no tool was adapted to their generation. In some cases (both tests on Clarity, Kadaster tests on ArcGIS) the roundabout suffer from a loss of shape, probably due to the simplification done on road segments in order to decrease their granularity.

A remark on how the testers interpret the constraints and their translation into the system. The four CPI tests obtain similar results, but do not respect the translation of the constraints on interchanges (ICC-3-9) in the system in the same way. The Kadaster and ICC tests consider they have partially expressed the constraints and not fully, because no use of the typification can be planned, whereas the CPI tests consider they have not expressed the constraints in the system at all, which would lead to opposite results on interchanges in the corresponding output. In the same time, the Kadaster and ICC tests consider they have not expressed the generic constraint on minimum distance between objects at all (ICC-2-3), while the CPI tests consider to be fully expressed it. So, in the Kadaster and ICC output the results on interchanges seems to be due to effort of the tester to generate interchanges, whereas in the CPI output the interchanges have been handled as a particular case of a generic proximity constraint.

Main conclusions:

- CPI proposes a displacement tool for roads (PUSH) while Clarity and ArcGIS do not. PUSH sometimes generates self-intersections locally.
- No software proposes tool for interchange detection and typification.
- No software proposes tool for the simplification of less complex junctions.
- Similar result do not mean that the constraints have been handled in the same way by the testers.

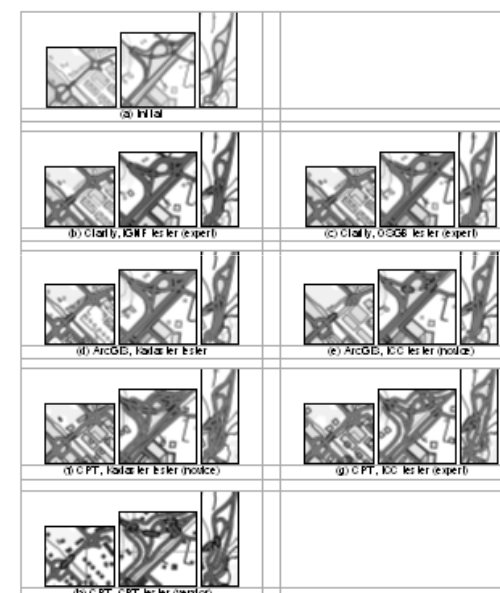
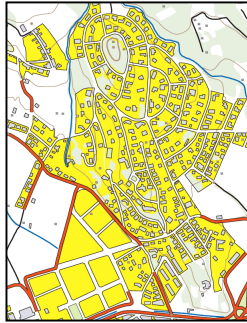


Figure 3.6.3.3 Test outputs of focus zone

ICC dataset – buildings in suburban areas



(a) Initial



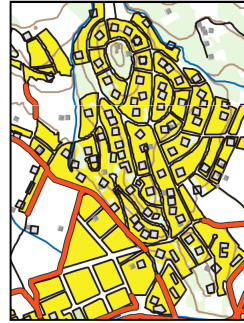
(f) CPT, TDK tester (novice)



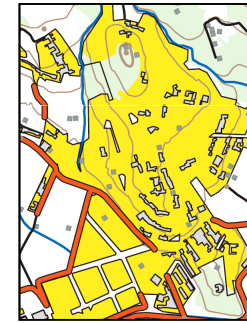
(d) ArcGIS, TDK tester (novice)



(b) Clarity, IGNF tester (expert)



(g) CPT, ICC tester (expert)



(e) ArcGIS, ICC tester (novice)



(c) Clarity, OSGB tester (expert)



(h) CPT, CPT tester (vendor)

Expert evaluation: methodology



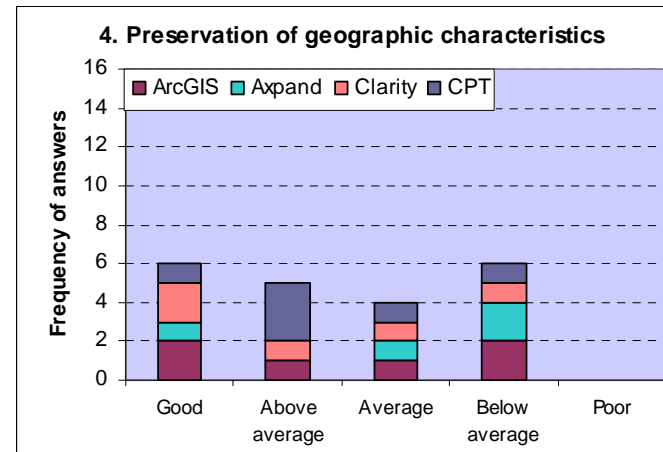
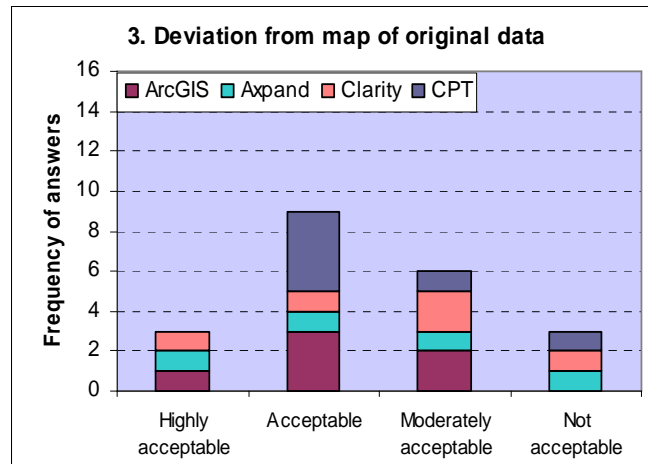
Global indicators
Level of manual editions required to meet the constraints
Deviation from initial (ungeneralised) data
Preservation of the geographic characteristics of the test area
Legibility
Seriousness and frequency of main detected errors
Number of positive aspects
Information reduction (undergeneralisation / overgeneralisation)

Individual constraints assessed in expert survey		
Constraints on one object	Constraints on two objects	Constraints on group of objects
minimal dimensions	spatial separation between features (distance)	quantity of information (e.g. black/white ration)
granularity (amount of detail)	relative position (e.g. building should remain at the same side of a road)	spatial distribution
shape preservation	consistencies between themes (e.g. contour line and river)	

Expert evaluation: example results



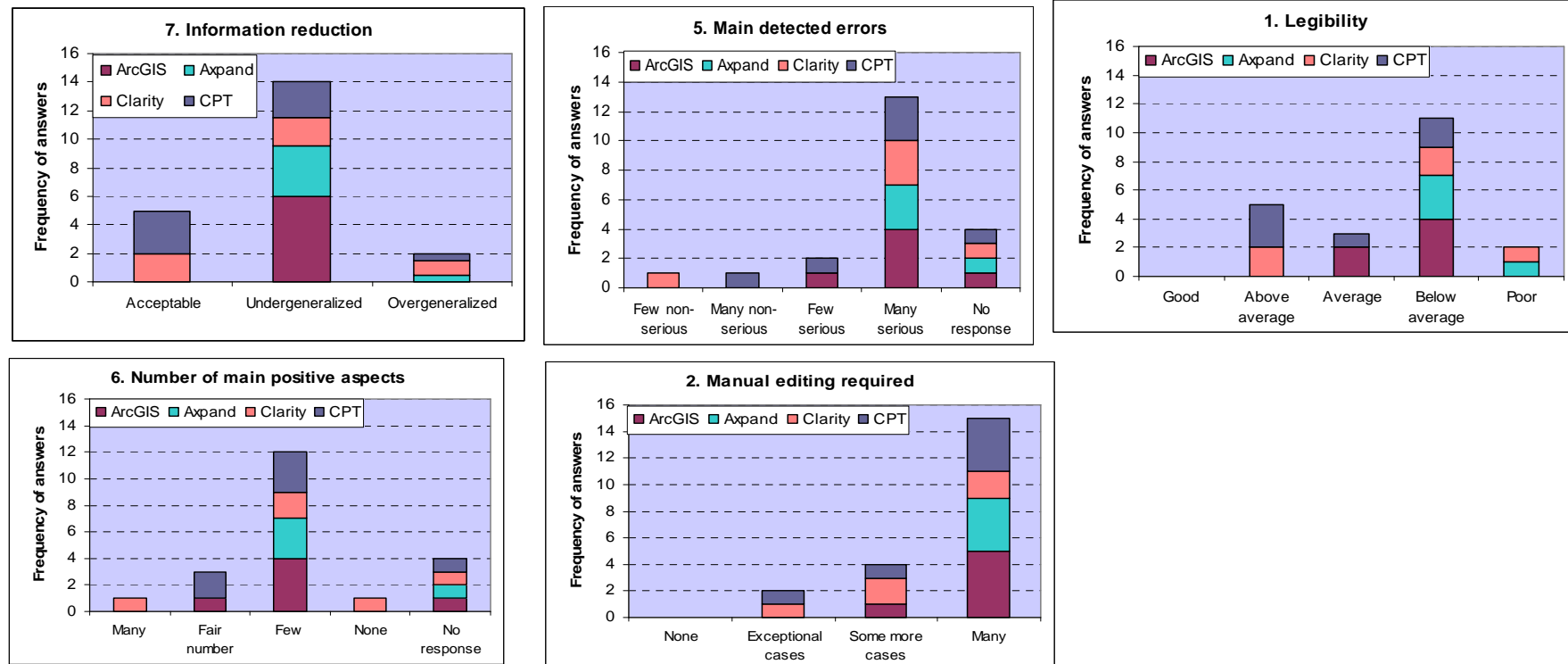
- Good scores for:



Expert evaluation: example results



- Lower scores for:



- Interesting if interactively generalised data would have been included

Conclusions capabilities of systems (1/4)



Discussed with vendors at IGN, Paris at 22
September 2009

- All systems offer potentials for automated generalisation, especially for single objects

Conclusions capabilities of systems (2/4)



- No generalisation problems are fully solved by the out-of-the-box systems
 - Some are close to being solved:
 - buildings and roads
 - Some are far from being solved: e.g.
 - apply different algorithms/parameters in different contexts (either not supported and/or detection measures are missing)
 - operations that concern more than one object (e.g. network typification)
 - terrain generalisation (relief)
 - displacement only in CPT and axpand

Conclusions capabilities of systems (3/4)



- For other problems solutions do exist (e.g. building simplification), but:
 - algorithms are difficult to parameterise; a direct match between parameters and constraints was often missing
 - detection tools are missing
 - controlling the effects of parameter values is difficult

Conclusions capabilities of systems (4/4)



- Satisfying complete NMA requirements requires customisation, progress should focus on:
 - Good customisation tools
 - Generic solutions (includes default parameterisation and default tools)
- Shortcomings have been solved by research (e.g. detection tools), and by vendors in parallel tests (e.g. displacement in Clarity and ArcGIS)

Conclusions on different results for one test case



- Specifications:
 - are sometimes fuzzy
 - do not fully express NMA requirements (focus on common/well known situations)
- Difficulties to parameterize the systems (once testers have understood the specifications):
 - Specification expression and parameters expected by the system often don't match
 - Differences between novice and expert testers of the system, or of the test case (even if expert of the system)
- Differences between testers:
 - Avoiding many errors versus striving for very good results for certain constraints or areas

Considerations on results



- Results are not that bad as they may look:
 - High expectations of the project (constraints, selection of complex/known problems, high quality paper maps)
 - Some missing functionalities have been fixed in vendors' parallel tests
 - Not a surprise that out-of-the box versions are not capable of fulfilling NMAs requirements; customization is definitely required
 - Systems are used more satisfactory in practice
 - Project is well received by vendors to push internal developments

Topics for future research



- Completing/refining constraints set
- Formalising/evaluating preservation constraints
- Constrained based evaluation:
 - Weighting & prioritizing
 - Interaction between constraints
 - Ignoring constraints for satisfying others
 - Constraint satisfaction values in ranges

Future project



- Testing on criteria beyond constraints
 - User-friendliness for parameterisation
 - Scalability and performance
 - Customisation!
 - Preservation of topology
 - Creation of links between initial and output data
 - Generalisation of incremental updates

Only if significant improvements are achieved on criteria tested in this project!

Many thanks to:

vendors:

***Axes systems, ESRI,
University of Hanover, and
1Spatial***



Core project team	Temporal members	Testers
Jantien Stoter (TUD&KAD) Dirk Burghardt (Zurich) Blanca Baella (ICC) Cécile Duchêne (IGNF) Maria Pla (ICC) Nicolas Regnauld (OS GB) Guillaume Touya (IGNF) Connie Blok (ITC)	Karl-Heinrich Anders, Jan Haunert (Hanover) Nico Bakker (Kadaster, NL) Francisco Dávila (IGNS) Peter Rosenstand (KMS, DK) Stefan Schmid (Zurich) Harry Uitermark (Kadaster, NL)	Magali Valdepérez (IGNS) Patrick Revell (OS GB) Stuart Thom (OS GB) Sheng Zhou (OS GB) Willy Kock (ITC, NL) Annemarie Dortland (Kadaster, NL) Maarten Storm (formerly Kadaster, NL) Patrick Taillandier (IGNF)