Usability Testing of Legibility Constraints

Hanna Stigmar^{1,2} and Lars Harrie¹

¹GIS Centre, Lund University, Lund, Sweden ² National Land Survey of Sweden, Gävle, Sweden hanna.stigmar@lantm.lth.se, lars.harrie@nateko.lu.se

Introduction

A major issue in cartography is the usability of maps, such as real-time maps for the web and mobile devices. These maps can be tailored for a specific purpose and even for a specific user (Reichenbacher 2004, Gartner 2004). This large freedom to tailor sets requirements on new analytical measures, or constraints, that describe the usability of the map.

This study is part of the project "the Swedish Planning Portal" (*Planeringsportalen*). By using this portal a user (employed at companies or governmental/local authorities) should be able to find planning information and especially geographic information related to physical planning. To enable the user to download and view the geographic information web services (following the OGC standards) will be set up. As for maps in general, it is for a web based map service like a planning portal important that the presented information is as usable as possible. This means that is should be easy for the user to read and comprehend the maps. In the first version of the Swedish Planning Portal there will not be any real-time generalisation. But there will be access to data layers in several resolutions, and a major issue is to select the proper layers for a certain application.

Usability of web services' maps can be seen from different perspectives. Arleth (2001) states that the *inner map design* and the *outer map design* are two equally important parts of a map service. The inner map design deals with the cartographic data and defines the elements' levels of detail, communication and aesthetics. The outer map design deals with the surrounding functions that are intended to be used on the map. Usability tests for maps should not differ from usability testing in general. Test persons, representative of the target population, should perform predefined tasks on the maps (Nielsen 1993, Rubin 1994). Due to the recent emergence of mobile map services, not much research has been performed concerning cartographic and usability aspects of these (Nivala 2005). Arguing for a usability approach in cartography, Nivala performed a usability tests for maps have been performed by Brodersen et al. (2001), Dillemuth (2005), and Arleth (2001).

In recent years, the generalisation research has tried to model the overall process of generalisation using constraints (Harrie and Weibel 2007). A constraint can be seen as requirements that should be obtained in the generalisation process. The constraints can be classified into *preservation constraints* (describing vital aspects of the map that should not be lost in the generalisation process) and *legibility constraints* (describing the readability of the map). It should be noted here that we are only concerned with the cartographic generalisation (i.e., generalisation aiming at the visualization of the information). Generalisation aiming at decreasing the amount of data for storage and distribution is not at all considered.

The aim of this study is to define and evaluate measures of legibility constraints of a map (inner map design) used for planning purposes. The purpose is to be able to select a layer with a good spatial resolution. In user tests the readability of maps (with different resolution) is examined. For each map several readability measures are calculated. By comparing the results of the user tests to the calculated measures it is possible to detect which measures are able to reflect the readability of the maps. These measures can then be used for selecting appropriate layers for the presentation.

The paper is organized as follows. It starts by describing the readability measures used in the study. Then we treat the methodology of the usability tests. The paper concludes with result of the first test and a discussion.

Readability measures

In this study we evaluate readability measures of a map. These measures are chosen to be useful for selecting layers of correct resolution and information content.

To enable a thorough evaluation of map properties and how to measure them, we need to examine many different types of measures. The measures should reflect different aspects of object and relations' complexities in order to provide a fuller and more detailed picture of the properties of the map. Based on the characteristics of the measures they can be subdivided into three *measure types*:

- amount of information,
- spatial distribution, and
- object complexity.

The map objects can be subdivided into *information types* based on their geometrical properties and if they concern the background/foreground of the map. In this study we use the following five information types (c.f. van Smaalen 2003, in Mackaness and Ruas 2007):

- *Minor objects* consisting of smaller stand-alone point, line or area objects. Symbols that are stored as points in the database are approximated with their minimum bounding rectangle in the study.
- *Line networks* consisting of line objects (such as roads and rivers) forming networks.
- Area objects forming tessellations.
- *Field-based data* consisting of e.g. contour lines.
- *Foreground data* such as text, icons and application specific data, e.g. planning information.

The readability measures should ideally be defined on the visual presentation of the data. Therefore we remove insignificant points in the object before the map is analysed. With insignificant points we mean points that do not affect the visual presentation of the data. The removal of the insignificant points is performed with Douglas-Peukers algorithm (Douglas and Peucker 1973) with a low threshold value.

In the following of this section a description of the measures in our study will be provided. In the end of the section Table 1 presents the measures and their applicability for the different information types.

Measures of amount of information

The measures of amount of information are based on the amount and size of the map objects.

Number of objects

The measure *number of objects* gives the number of all objects in the map (c.f. Harrie and Stigmar 2008). For point and area objects the definition of an object is obvious, while for line objects it can be somehow problematic. For this study we have decided to define one line object as the segment/segments situated between two nodes that are either end nodes or nodes where three or more links converge.

Number of points in the objects

Based on the ideas of Biederman (1985), the object points can be thought to reflect the amount of work we need to perform in order to perceive an image. Thus, the relatively simple measure *number of points in the objects* gives the total number of object points for all map objects (cf. Stigmar 2006).

Number of object types

The measure *number of object types* gives the number of all object types in the map. This measure will be used for all objects.

Object line length

The measure *object line length* gives the total line length of the map objects (for area objects the total length of the boundary is used).

Object area

The measure *object area* gives the total area of the map objects (cf. Harrie and Stigmar 2008). As point objects do not have an area we approximate this with an average area of their symbols. This because the symbols presented to the map user do have an area.

Measures of spatial distribution

The measures of spatial distribution are based on the density and distribution of the map objects.

Spatial distribution of objects

The measure *spatial distribution of objects* is used for minor objects and line networks. It is based on Li and Huang's (2002) geometric measure (and earlier work by Sukhov 1967, 1970), which gives the entropy (see Shannon and Weaver 1964) of the map objects' Voronoi regions using the ratio between each Voronoi region size and the total map size. The problem with Li and Huang's definition is that it cannot handle maps with different number of objects. To circumvent this shortage we normalise the entropy value with the entropy we would get if all Voronoi regions were of equal size (and the number of objects invariant). Hence, we obtain the following index (HI_{SD-Obi}):

$$HI_{SD_{-}Obj} = \frac{\sum_{i=1}^{n} p_i \log p_i}{\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n}} = \frac{\sum_{i=1}^{n} p_i \log p_i}{\log \frac{1}{n}}$$
(1)

where p_i is the ratio between the Voronoi area *i* and the total map area, and *n* is the number of objects.

Spatial distribution of points

The measure *spatial distribution of points* is, like spatial distribution of objects, used for minor objects and line networks. It gives an index for point distribution (HI_{SD_Poi}), with an analogous definition:

$$HI_{SD_Poi} = \frac{\sum_{i=1}^{k} p_i \log p_i}{\log \frac{1}{k}}$$
(2)

where p_i is the relative size of the Voronoi region for point *i* and *k* is the number of points.

Homogeneity

Homogeneity for a single object is defined as the proportion of the Voronoi neighbours that belong to the same object class. The homogeneity for the map is then simply defined as the mean value of the homogeneity of the objects. This measure is quite similar to Li and Huang's (2002) thematic measure, which gives an entropy based on objects' neighbours' types.

Number of neighbours

The measure *number of neighbours* is based on Li and Huang's (2002) "topologic" measure. The average number of neighbours for the map objects is computed from the objects' Voronoi regions' neighbours.

Individual density

The measure *individual density* describes the tightness of an object within its surroundings. It is defined as the ratio between each object area and the object's Voronoi area. The measure is taken from Hangouet (1999, as described in AGENT DC1 1999).

Proximity indicator

This indicator is defined as the number of object pairs where the shortest distance between the objects is less than a threshold value.

Measures of object complexity

The measures of object complexity are based on the shape and size of the map objects.

Object size

The measure *object size* defines the number of objects smaller (in map units) than a threshold value.

Line segment size

The measure *line segment size* defines the number of line segments smaller (in map units) than a threshold value.

Angularity

The *angularity* measure is defined as all the changes in direction of a line divided by the total length of a line (cf. *sinuosity* in João, 1998 and *total angularity* in McMaster, 1987).

Line connectivity

The measure *line connectivity* gives the mean degree of the nodes in the graph, where the degree is defined as the number of links incident to a node (Trudeau 1993, Mackaness 1999).

Polygon shape

The measure *polygon shape* gives the ratio between the object area and the area of its convex hull.

Application of the readability measures

Table 1 shows for which information types the readability measures are applied.

Table 1. A compilation of the measures and their application for the information types (rows) and measure types (columns).

	Measures of amount of information	Measures of spatial distribution	Measures of object complexity
Minor objects	 Number of objects Number of points in the objects Object line length Object area 	 Spatial distribution of objects Spatial distribution of points Number of neighbours Individual density Proximity indicator 	Object sizeLine segment sizeAngularityPolygon shape
Line networks	 Number of objects Number of points in the objects Object line length 	Proximity indicator	Line segment sizeLine connectivityAngularity
Area objects forming tessellations	 Number of objects Number of points in the objects Object line length Object area 	Number of neighbours	Object sizeLine segment sizeAngularityPolygon shape
Field-based data	 Number of objects Number of points in the objects Object line length 	• Number of neighbours	Line segment sizeAngularity
Foreground data	 Number of objects Number of points in the objects Object line length Object area 		
All objects	• Number of object types	• Homogeneity in a group	

Implementation of the readability measures

The computations of the readability measures are made by a Java program built on the open source packages *JTS Topology Suite* (JTS) and *JTS Unified Mapping Platform* (JUMP) (JUMP project 2007). In order to create Voronoi regions we use the c-program Triangle (Shewchuk 1996, 2002) integrated using Java native interface (Gordon 1998).

Method

The overall aim of the study is to develop guidelines for how the maps in the Swedish Planning Portal should be designed and presented. The study will be a mixture of a usability study and a preference study run in a computer environment. Two user tests will be performed, one so called expert test using planning professionals, and one main map test using average map users. The studies will be based on previous work reported in Stigmar (2006) and Harrie and Stigmar (2008). A pilot expert test has been performed with a planning professional in order to determine the design of the test procedure, the test tasks and the test maps. The overall test sequence is shown in Figure 1.

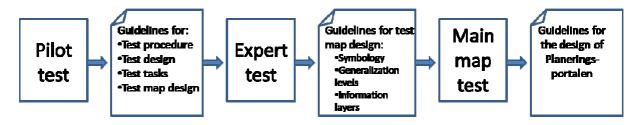


Figure 1. The test sequence in the study.

In all the tests static maps will be used. As the test subjects are not able to adjust the maps (e.g. zoom in or out, add or remove data layers) it is important that the maps are designed to be suitable for the tasks. Each map will also be generated in four generalisation levels. The maps used in the tests are created in advance using a standard GIS-program. For all maps the readability measures are computed.

Expert test

The expert test aims at finding guidelines to be used later for the main map test. The test is categorized as a preference study with some elements of a usability study. The test will use test subjects who work with planning tasks on a regular basis and are familiar with planning applications such as the Swedish Planning Portal. The intention is to find these test subjects in municipal planning departments. A total of six to eight experts will conduct the test.

The maps used for the expert test should be representative for the intended use of the Swedish Planning Portal. They will consist of "background" and "foreground" information. The background information consists of topographic map data from Lantmäteriet (the Swedish national mapping agency) and from the local municipality, while the foreground information consists of planning related data such as restricted and protected areas. Two cartographic symbologies will be used for the test: one common topographic map symbology as used by e.g. Lantmäteriet, and one symbology more adapted to web maps.

The test subjects will be given different tasks, one task at a time, and asked to solve the task with the help of one map presented on a computer screen. The tasks will be divided into four categories representing possible tasks for the users of the Swedish Planning Portal:

- finding specific feature,
- describing feature,
- retelling specific feature, and
- preference test.

Only one of the generalisation levels will be presented. The test monitor will observe the subject while solving the task. Sound recording will also be done, as well as screen action recording, to enable post-processing of the test. When the task is completed, the test monitor will discuss the map with the subject. The discussion will focus on the topics:

- Was the map suitable for solving the task?
- What would make the map better/worse?
- How should the symbology be designed?
- Does the map contain too much information? If so, what type of information? And where in the map is there too much information?
- Are the objects in the map to complex? If so, which types of objects are too complex?

The discussion will be initiated from a set of discussion questions, which will be formulated in a way that the readability measures' effect will be considered. The subject will here be shown all four generalisation levels of the map, as well as the two symbologies. The outcome of the expert test will be used to improve the design of the main map test.

Main map test

The main map test aims at finding guidelines for the design of maps to be used in the Swedish Planning Portal. The test is categorized as a usability study using average map users as test subjects. Approximately 30 test subjects will be tested.

The maps of the main map test will originate from the same data as the maps of the expert test. The design of the maps, with regards to information layers, generalisation levels and symbology, will be based on the outcome of the expert test.

The test process will be more automated than for the expert test. The test subjects will be able to perform the test using a computer and writing answers and comments on paper. The subjects will be given one task at a time, which to solve with the help of a map presented on the computer screen. The tasks will be of the same types as the ones in the expert test. When having solved one task the subject will be asked to answer a set of questions regarding the task and the map design and readability. Thereafter the subject will perform the next task. Each task will be timed. As each test subject will get a different (concerning generalisation level) set of maps it will be possible to compare the results for the different generalisation levels. The results of the test will then be compared to the readability measures computed for the test maps. Finally, the measures' performance will be evaluated in order to determine which ones are appropriate to use for selecting appropriate layers with the right resolution and information content.

Pilot expert test

A pilot expert test has been performed in order to determine if the test procedure, maps, and test tasks are suitable for the "real" expert test. The test subject works as a municipal development engineer, which is one type of the intended users of the Swedish Planning Portal.

Three different maps were used in the test, one in scale 1:30 000, one in 1:10 000 and one in 1:5 000. Each of the maps was represented in three different generalisation levels. The test subject was first shown the most detailed map (generalisation level 1, see Appendix) and asked to solve the task "Where in the map is the largest occurrence of ancient remains". The subject was asked to think aloud while solving the task. When the task was completed a discussion was initiated by the test monitor. The discussion originated from a set of questions regarding the test task, the map scale, the symbology, the information layers, and the overall readability of the map. The readability was then discussed more thoroughly concerning different aspects, as reflected by the readability measures. The subject was first asked to grade the readability concerning the aspect in question, e.g. too few/ too many objects (*Number of objects*), and regularly/irregularly distributed objects (*Spatial distribution of points* and *Number of neighbours*). The effect of the particular readability aspect was then discussed.

Results of pilot test

The aim of the pilot expert test was to find out whether the test was functioning in a desired way. The main concerns were the overall test procedure, the test tasks, and the maps.

The test procedure was functioning fairly well. A more thorough introduction to the test was desired, though. The test tasks were regarded as easy to comprehend. The test subject did,

however, find it difficult to relate some of the tasks to her work as she would not use this sort of map application for solving this type of task. Currently the test subject did normally not utilize GIS functionality in her map work.

The discussion questions were regarded as very difficult to understand and to put into practice. A problem here was that the test subject did not understand the cartographic terminology used to describe the map (such as *information amount* and *complexity of objects*). A challenge in the future tests is to find a vocabulary that is understood by the test subjects and that are possible to connect to the readability measures. Probably, the test phrases have to be reformulated for the "real" expert test in a less theoretical manner.

The maps were regarded as fairly ok. The symbology was regarded as good in most aspects. Some of the information layers were, however, regarded as too prominent, while some should be more enhanced. It was pointed out that the presented information layers were good, but the map would be better if additional layers were also presented. Concerning the generalisation levels, the medium generalized maps (generalisation level 2) were regarded as the best. The least generalized maps (generalisation level 1) were regarded as containing an excess amount of irrelevant information for solving the task. The most generalized maps (generalisation level 3) were regarded as containing too little information.

Appendix B contains graphs of some of the computed readability measures for the maps shown in Appendix A. Clearly it is quite complicated to interpret the readability measures and relate them to the maps. It is our intention that the tests in this study at least will provide an indication, e.g. an interval, of where the measures reflect a usable map. It should be noted that the maps used in the pilot study where rather similar. In the forthcoming studies more diverse maps must be used (using e.g. municipality data) to better reflect different values of the measures.

The readability measures are on a syntactic level (computed from the geometric objects), with some simple semantic information (object types) (cf. definition of syntactic and semantic in MacEachren 1995). The test subject did however comment the map on a more advanced semantic level e.g.: since the ancient remains overlay this open area it is likely to be a cultural protected area. We could conclude that the readability measures used in this test are far from supporting this type of more advanced map reading.

Discussion

There are several shortcomings of the readability measures such as:

• A map is often read in different hierarchical levels. First we get an overview and then we focus on the area of interest (cf. MacEachren 1995). Ideally, we should have measures that describe the map property on the different hierarchical levels (cf. the concepts of micro, meso and macro objects in e.g. Ruas 2000).

- For presenting property of the whole map we can use the measures for the amount of information. However, these measures really concentrate on the amount of information defined for single objects (or features within objects). When studying the overview of a map we tend to group objects into meaningful entities (see description of Gestalt theory in e.g. MacEachren 1995). For our studies we might have to introduce map properties that consider this grouping of objects.
- For those cases the map reader focuses on an area, we need measures that describe the property of that area of the map. As the properties of the map differ greatly from one part of the map to another, these measures will also vary. There are a number of techniques that could be used to divide the map. A simple way is to use a regular grid. The size of the grid cells needs to be given some consideration. Ideally, the size of the grid cells should be based on the size of the eye focus in map reading (see e.g. MacEachren 1995). A disadvantage with regular grid is that it is insensitive of the distribution of objects in the map. By using e.g. spatial clustering techniques (Murray and Estivill-Castro 1998, Yan et al. 2008) we can form meaningful groups of objects and define the measures of these groups.
- In the pilot text we did not put any attention to the foreground data, e.g. text, icon and planning information. This is something that we need to consider in the remaining tests.

However, the main challenge in the future tests is probably not to improve the readability measures. The main challenge is to make the connection between the usability testing and the readability measures.

Acknowledgement

We would like to thank Joakim Eriksson, the Design Centre, Lund University, for valuable comments on the usability/performance testing; Marie-Louise Svensson, Helsingborg stad, for performing the pilot study; and Lars Sävermarker, Lantmäteriet, for proving us with cartographic data. Financial support from the Vinnova project Planeringsportalen and from Lantmäteriet are acknowledged.

References

AGENT, 2000. Project AGENT, http://agent.ign.fr/summary.html (accessed 2008-02-26).

- AGENT DC1, 1999. Selection of Basic Measures. Available at: http://agent.ign.fr/deliverable/DC1.html (accessed 2008-02-26).
- Arleth, M., 2001. Design af skærmbilledkort. Doctoral Thesis, Aalborg Universitet, Denmark.
- Biederman, I., 1985, Human image understanding: recent research and a theory. *Computer Vision, Graphics, and Image Processing*, Vol. 32, No. 1, pp.29-73.
- Brodersen, L., Andersen, H. K. and Weber, S., 2001. *Applying eye-movement tracking for the study of map perception and map design*, National Survey and Cadastre, Denmark.

- Dillemuth, 2005. *Human-map Interaction for Mobile Cartography*. Masters Thesis, University of California, Santa Barbara, USA. Avaliable at: http://www.geog.ucsb.edu/~julie/MastersThesis_Dillemuth.pdf.
- Douglas, D. H., and T. K. Peucker, 1973. Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature. *The Canadian Cartographer*, Vol. 10, No. 2, pp 112-122.
- Gartner, G., 2004. Location-based mobile pedestrian navigation services the role of multimedia cartography. International Joint Workshop on Ubiquitous, Pervasive and Internet Mapping, Tokyo. Available at: http://www.ubimap.net/upimap2004/html/papers/UPIMap04-B-03-Gartner.pdf (2007-03-29).
- Gordon, R., 1998. Essential JNI Java Native Interface. Prentice-Hall.
- Hangouet, J. F., 1998. Voronoi Diagrams on Segments Properties and Tractability for Generalization Purposes. Technical report for Agent, Cogit, IGN.
- Harrie, L. and Stigmar, H., 2008. An evaluation of measures for quantifying map information, Submitted.
- Harrie, L. and Weibel, R., 2007. Modelling the Overall Process of Generalisation In: Mackaness, W. A., Ruas, A. and Sarjakoski, L. T. (eds.), *Generalisation of Geographic Information: Cartographic Modelling and Applications*,, Elsevier Science.
- João, E. M., 1998. Causes and Consequences of Map Generalisation, Taylor & Francis.
- JUMP project, 2007. The JUMP project, http://www.jump-project.org/ (accessed 2007-04-04).
- Li, Z. and Huang, P., 2002. Quantitative measures for spatial information of maps, *International Journal of Geographical Information Science*, Vol. 16, No. 7, pp. 699-709.
- MacEachren, A.M., 1995. *How Maps Work, Representation, Visualization, and Design*. The Guilford Press, New York, USA.
- Mackaness, W. A. and Mackechnie, G. A., 1999. Automating the Detection and Simplification of Junctions in Road Networks, *GeoInformatica*, Vol. 3, No. 2, pp. 185-200.
- Mackaness, W. A. and Ruas, A., 2007. Evaluation in the Map Generalisation Process. In: Mackaness, W. A., Ruas, A. and Sarjakoski, L. T. (eds.), *Generalisation of Geographic Information: Cartographic Modelling and Applications*, Elsevier Science., pp. 89-111.
- McMaster, R. B., 1987. Automated Line Generalisation. Cartographica, Vol. 24, No. 2, pp. 74-111.
- Murray, A. T., and, V. Estivill-Castro, 1998. Cluster Discovery Techniques for Exploratory Spatial Data Analysis. International Journal of Geographical Information Science, Vol. 12, No. 5, pp. 431-443.
- Nielsen, J., 1993. Usability Engineering. Academic Press, San Diego, USA.
- Nivala, A.-M., 2005. User-Centred Design in the Development of a Mobile Map Application. Licentiate Thesis, Helsinki University of Technology, Finland.
- Nivala, A.-M., 2007. Usability Perspectives for the Design of Interactive Maps. Doctoral Dissertation, Helsinki University of Technology, Finland.
- Reichenbacher, T., 2004. Mobile Cartography Adaptive Visualisation of Geographic Information on Mobile Devices. PhD Dissertation, Technical University of Munich, Germany.
- Ruas, A., 2000. The Roles of Meso Objects for Generalisation. *Proceedings of the 9th Spatial Data Handling Symposium*, Beijing, pp. 3b.50-3b.63.
- Rubin, J., 1994. Handbook of Usability Testing, John Wiley And Sons Ltd, New York, USA.
- Shannon, C. E. and Weaver, W., 1964. *The Mathemathical theory of Communication*. Urbana: The University of Illinois Press.
- Shewchuk, J. R., 1996. Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator, in Lin, M. C. and Manocha, D. (eds), *Applied Computational Geometry: Towards Geometric Engineering*, Vol. 1148 of Lecture Notes in Computer Science, Springer-Verlag, Berlin, pp. 203-222.
- Shewchuk, J. R., 2002. Delaunay Refinement Algorithms for Triangular Mesh Generation, *Computational Geometry: Theory and Applications*, Vol. 22, No. 1-3, pp. 21-74.

Stigmar, H., 2006. Amount of Information in Mobile Maps: A Study of User Preference. *Mapping and Image Science*, No. 4, pp. 68-74.

Sukhov, V. I., 1967. Information capacity of a map entropy, Geodesy and Aerephotography, Vol. 10, pp. 212-215.

Sukhov, V. I., 1970. Application of information theory in generalization of map contents, *International Yearbook of Cartography*, X, pp. 41-47.

Trudea, R. J., 1993. Introduction to Graph Theory, Dover

- van Smaalen, J. W. N., 2003. Automated Aggregation of Geographic Objects: A New Approach to the Conceptual Generalisation of Geographic Databases, Doctoral Dissertation. Wageningen University and Research Centre, The Netherlands.
- Yan, H., Weibel, R., and Yang, B., 2008. A Multi-parameter Approach to Automated Building Grouping and Generalization, Geoinformatica, Vol. 12, pp. 73-89.

Appendix A – Maps in the pilot expert test



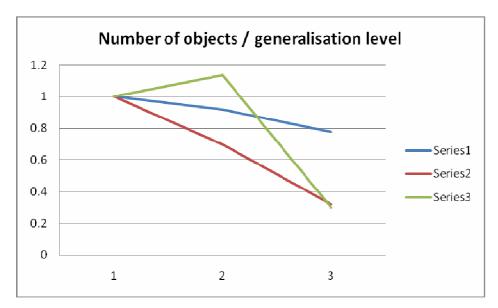
Generalisation level 1



Generalisation level 2

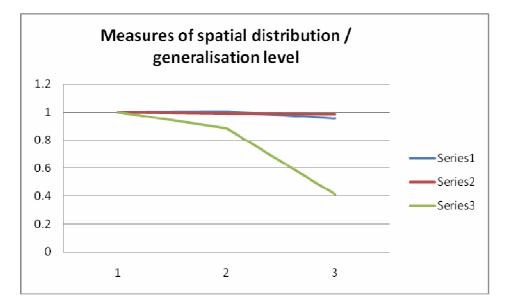


Generalisation level 3

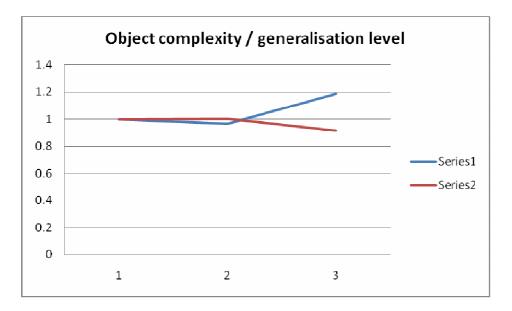


Appendix B – Computed readability measures

Normalized measures of amount of information. Series1 shows *Number of objects* for minor objects, Series2 shows *Number of objects* for line networks, and Series3 shows *Number of objects* for area tessellations.



Normalized measures of spatial distribution. Series1 shows *Conflict indicator* between objects less than 1mm, Series2 shows *Spatial distribution of objects*, and Series3 shows *Individual density* (of the 90% quantile).



Normalized measures of object complexity. Series1 shows *Polygon shape* (20% minimal ratio of object area and its convex hull area) for minor objects, and Series2 shows it for area tessellations.