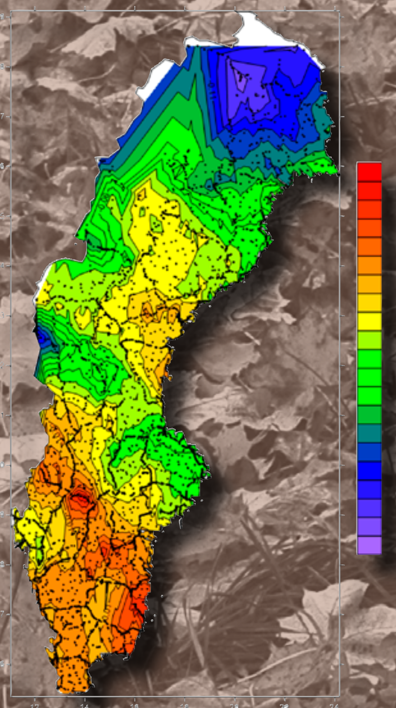
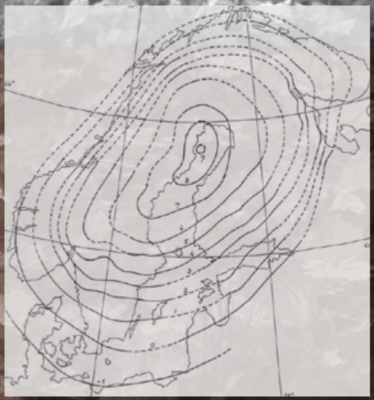
Kart & Bildteknik

Mapping and image Science

Scientific edition

2011:3





Kartografiska Sällskapet

Swedish Cartographic Socieyt

Kartografiska

Välkommen till Mapping and Image Science, det tredje vetenskapliga numret av Kart & Bildteknik. Det förra numret gavs ut i december 2006 men målsättningen är att ge ut ett vetenskapligt nummer varje år. Alla artiklar granskas av erkända forskare med internationellt erkänd vetenskaplig kompetens. Syftet med Mapping and Image Science är:

- Att sprida kunskap om studier som utförs i Sverige och övriga Norden inom Kartografiska Sällskapets ämnesområde.
- Att ge yngre forskare (t.ex. doktorander) möjligheter att få resultat från sina studier granskade och publicerade.

Instruktioner för att skriva artiklar till Mapping and Image Science finns på Kartografiska Sällskapets hemsida (http://www.geoforum.se/; se under KS-Publikationer och tidskrifter-Kart & Bildteknik-Vetenskapliga nummer).

I detta nummer publiceras fyra vetenskapliga uppsatser och en sammanfattning av ett arbetsliv som professor i kartografi. Jonas Ågrens och Runar Svenssons uppsats beskriver utvecklingen av en nordisk landhöjningsmodell och av det nya höjdsystemet för Sverige. Den beskriver på ett enkelt och lättförståeligt sätt framtagningen av dessa samt inordningen av dem i en geodetisk infrastruktur. Hanna Stigmars och Kirsten Rassmus-Gröhns uppsats är en del i EU-projektet Haptimap som styrs från Lund. Den behandlar metoder att mäta användbarheten av kartografiska hjälpmedel som navigationsstöd för gångtrafikanter med nedsatt syn. Den ingår i Stigmars doktorsavhandling. Nästa uppsats är också från Lund. Den behandlar metoder för att identifiera svårtydda ytor i topografiska kartor med meningen att få mer användbara kartor. Den fjärde uppsatsen är från Högskolan i Gävle. Den behandlar harmonisering av metadata och geodata i INSPIRE för nättjänster för att underlätta sökningen och användningen av data inom INSPIRE.

Den sista uppsatsen är skriven av Ferjan Ormeling när han lämnade sin tjänst som professor i kartografi vid Utrecht University. Det var redan då beslutat att tjänsten inte skulle utlysas vilket också drabbat andra professurer i geografi. Även om uppsatsen inte är vetenskaplig och granskad som en sådan ger den en inblick i vad som kommer att saknas då professuren försvinner.

Mycket nöje och ett stort tack till alla författare och anonyma granskare.

Glumslöv i juni 2011

Bengt Rystedt

TABLE OF CONTENT

Bilaga till Kart & Bildteknik 2011:3

Ansvarig utgivare:
Peter Wasström
Ordförande i Kartografiska Sällskapet
Tel: 026 - 63 32 37
E-post: peter.wasstrom@lm.se

Redaktör:
Bengt Rystedt
Professor
Tel: 070-6073429
E-post: bengt.rystedt@lm.se

Layout och produktion:
Malm Reklam & Bild i Gävle AB
Tel: 026-19 58 39
E-post: malm.reklam&telia.com

Tryck:
Gävle Offset
Tel: 026-66 25 00

The Height System RH 2000 and the Land Uplift Model NKG2005LU

Jonas Ågren, Runar Svensson

Lantmäteriet, Geodetic Research Department
SE-801 82 Gävle, Sweden
jonas.agren@lm.se, runar.svensson@lm.se

Abstract

The computation of the latest precise levellings in Finland, Norway and Sweden was made as a Nordic co-operation within the Working Group for Height Determination of the Nordic Geodetic Commission (NKG). This work includes the compilation of the Baltic Levelling Ring (BLR), which consists of precise levellings from all the Nordic and Baltic countries as well as Poland, Germany and the Netherlands. The main purpose of this paper is to describe the choice of system definition and computation of the postglacial land uplift model for the final adjustment of the BLR. The result of this adjustment constitutes the Swedish height system RH 2000.

Initially, it was decided that the new system should be a realisation of the European Vertical Reference System (EVRS) using the Normaal Amsterdams Peil (NAP) as zero level. The final land uplift model is a combination of the geophysical model of Lambeck, Smither and Ekman and the mathematical (empirical) model of Vestøl. It has later been adopted as a Nordic model with the name NKG2005LU. We describe the path leading to this model and analyse the consequences of the chosen definition and uplift model by comparing the resulting heights to Mean Sea Level (MSL) in the Nordic and Baltic Seas and to the old Swedish height system RH 70. We finally discuss the role of the land uplift model in the modern Swedish geodetic infrastructure.

1.Introduction

Sweden introduced the new National height system RH 2000 in 2005. It is the final result of the third precise levelling of Sweden. The computation was made in Nordic co-operation within the Working Group for Height Determination of the Nordic Geodetic Commission (NKG) and with EUREF (IAG Reference Frame Sub Commission for Europe). The adjustment includes observations from the whole Baltic Levelling Ring (BLR) consisting of precise levellings from the Nordic and Baltic countries as well as from Poland, Germany and the Netherlands.

To reduce the observations to a common reference epoch, the postglacial land uplift model NKG2005LU was constructed. It is computed as a combination of the mathematical (empirical) model of Vestøl (2005) and the geophysical model of Lambeck et al. (1998); see also Vestøl (2006) for a full documentation of a later, modified version of the model. To obtain heights agreeing as closely as possible with the heights of other European countries, the system is defined as a realisation of the European Vertical Reference System (EVRS); cf. Ihde and Augath (2001) and EUREF (2005). It should be pointed out that at the time this work was made (2005) it had not yet been defined on the European level how the phenomena of postglacial rebound should be taken care of (cf. the end of this introduction).

The main purpose of this paper is to introduce the third precise levelling of Sweden, present the BLR, discuss the definition of RH 2000 and describe the path leading to the land uplift model NKG2005LU. Another purpose is to analyse the consequences of the definition and land uplift model by studying the height of the Mean Sea Level (MSL) along the Swedish coasts and by comparing RH 2000 to the old Swedish height system RH 70. The paper ends with a short discussion of the land uplift model as a part of the Swedish geodetic infrastructure, comprising the network of permanent reference stations SWEPOS®, the Swedish reference system SWEREF 99, the height system RH 2000, the gravity system RG 82 and the geoid model SWEN08_RH2000.

The main part of this paper was presented at the General Assembly of the Nordic Geodetic Commission (NKG) in Copenhagen 2006 and has been published in the corresponding proceedings; see Ågren and Svensson (2006). Consequently it is intended for professional geodesists. It is, however, believed that the content might also be of interest to readers with only a basic knowledge of geodesy and surveying, even though some parts are rather technical. The reader is referred to Ekman (2002) (in Swedish) for a good introduction to the subject.

Since RH 2000 was finalised in early 2005, the European Vertical Height System (EVRS) has changed so that the NAP level is now given by 13 benchmarks distributed over the stable parts of Europe; see Sacher et al. (2009). The latest common European height system, EVRF 2007, has further been computed using the same land uplift model as for RH 2000 (i.e. NKG2005LU) and using the same reference epoch. These developments are not treated in this paper.

2. The third precise levelling of Sweden

The fieldwork of the third precise levelling of Sweden started in 1979 and the last line in the network was observed in 2001. See Becker (1984) and Becker et al. (1998) for more details concerning the third precise levelling. The main reason for starting yet another levelling in 1979 was that the previous levellings were not sufficiently accurate at the same time as the coverage of the country was too poor (Fig.1). The aim of the new third precise levelling was thus to create a homogeneous network covering the whole country, dense enough to allow all local users to connect their measurements to easy accessible benchmarks. Another aim was to achieve a better estimation of the land uplift by comparing the new levelling with the first and second counterparts (Eriksson et al. 2002). The network of the third precise levelling is dense and homogeneous at the same time as all the levelling was carried out using the same technique and equipment. Considering the length of the project, the latter fact is quite remarkable.

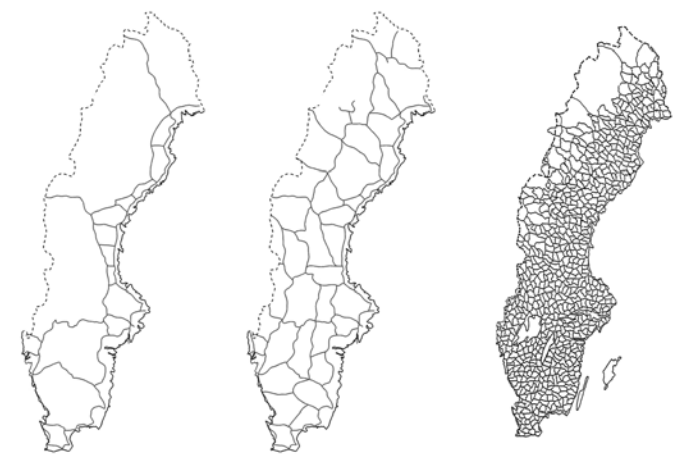


Fig. 1. The network of the first, second and third precise levellings of Sweden (in the order mentioned from left to right).

The network consists of approximately 50 000 km double-run levelling and 50 800 benchmarks. The distance between them is about 1 km. The network covers the whole country in closed loops, with a circumference of approximately 120 km, except for the mountain areas to the North-West where the roads are few, ´which makes it difficult to achieve the same density (Fig. 1). However, since the population is also very sparse in these areas, this is not too critical. The lines are connected to the levelling lines of the neighbouring countries, which results in closed loops along the borders. This also makes it possible to extend the network to the whole BLR (cf. Sect. 3). The network was planned in co-operation with the local users in order to increase the utility of the points for the connection of local measurements.

The observations were carried out by means of the motorised levelling technique using one instrument car and two rod cars (e.g. Becker et al. 1998; 2001). The instrument is a Zeiss Jena NI002 with a 360 degree rotating eyepiece and a reversible pendulum. The rods are 3.5 m invar staffs with double scales. A separate 3.0 m invar rod is used for the connection to the benchmarks. The average sight length is approximately 35 meters (50 meters allowed). To make it possible to determine the land uplift, accessible points from the former precise levellings were connected to the network. The rejection limit was $2\sqrt{L}$ mm (L in km) during the whole project, which corresponds to a 2-sigma limit. If the measurements of a double-run section differed more than this limit, then the section was relevelled. About 7 % of all sections have been remeasured for this reason. The whole production line, from the observations to the archive, is digital.

3. The Baltic Levelling Ring

As mentioned in the introduction, a large part of the processing of the precise levellings of Sweden, Finland and Norway has been made as a Nordic co-operation under the auspices of NKG. Denmark also contributed actively to the task, even though the Danish height system DVR 90 had already been finalised at the time (Schmidt 2000). To be able to connect to the Normaal Amsterdams Peil (NAP), which is the traditional zero level for the United European Levelling Network (UELN), and to be able to determine the relations to the neighbouring countries of Sweden, it was decided to extend the Nordic network with the precise levellings from the Baltic States, Poland, Northern Germany and the Netherlands. The non-Nordic data was provided by EUREF from the UELN-database.

The whole network, which has been named the Baltic Levelling Ring (BLR) is illustrated in Fig. 2. Unfortunately, it has not been possible to close the ring with levelling observations around the Gulf of Finland. However, by means of other information (sea surface topography or GPS in combination with a geoid model), closing errors may still be computed amounting to a valuable check of the adjustment. It should be noticed, though, that only levelling observations are included in the final adjustment. As can be seen in Fig. 2, the BLR extends over a quite large area. To be able to reduce all observations for the postglacial land uplift, the applied model should naturally cover the same area. This should be kept in mind in Sect. 6.

Fig. 2.
The Baltic Levelling Ring (BLR).

4. System definition of RH 2000

The choice of system definition was discussed within the NKG; see for instance Mäkinen et al. (2004). However, RH 2000 had to be finalised by Sweden alone in January/February 2005 (Ågren and Svensson 2007). At that time, the crucial aspects of the definition had already been considered in the NKG.

Then, it was decided that RH 2000 should be defined as the Swedish realisation of the European Vertical Reference System (EVRS). The main reason for this was to arrive at a system that agrees as well as possible with other European systems. From this decision it follows that (EUREF 2005; Ihde and Augath 2001):

- The **zero level** is given by the geopotential number from the latest official UELN-solution (EVRF 2000) for the Normaal Amsterdam Peil (NAP). It is true that NAP was not strictly a part of the 2005 definition of the EVRS, but at the time it was the only available alternative to realise EVRS (as for EVRF 2000).

- **Normal heights** are utilised. The normal height is one type of "height above sea level"; cf. Heiskanen and Moritz (1967) for a proper explanation.

- The **zero system** is utilised for the permanent tide. Different strategies exist for how to deal with the permanent, non-varying part of the tidal correction. The "zero system" implies that the permanent attraction of the sun and the moon is removed, while the corresponding deformation of the Earth is left as is; cf. Ekman (1989 and 2002).

One problem with the EVRS definition available in 2005 is that no advice was given of how the postglacial land uplift should be treated. This means that these matters had to be taken care of at the Nordic level. Since postglacial rebound is a very significant phenomena (cf. next section), the system definition for RH 2000 includes the following items concerning how the land uplift is handled:

- The **reference epoch** for the reduction of postglacial rebound is 2000.0. This was decided at the Nordic level within the NKG.

- The **postglacial land uplift model** is NKG2005LU. The construction of this model is described in Sect. 6.

It should be pointed out that the RH 2000 definition is applied for the adjustment of the whole BLR, even though RH 2000 is strictly only a Swedish system.

5. Observations of the postglacial rebound of Fennoscandia

The Nordic area (Fennoscandia) is still experiencing postglacial rebound after the melting of the Weichselian ice sheet which covered Northern Europe and terminated some 10 000 years ago. The maximum uplift is approximately 1 cm per year; see for instance Fig. 3 below. This means that in projects with geodetic measurements conducted over long time periods in high accuracy applications, the phenomena must be taken into account somehow. To correct all measurements to a suitable reference epoch, a model describing the uplift is needed.

The land uplift can be determined using different types of observations. For the construction of a land uplift model for the computation of the BLR, the most important types are:

- Tide gauge (mareograph) observations of the Mean Sea Level (MSL), from which the apparent uplift is determined by linear regression. The apparent uplift is the uplift with respect to MSL over a certain time period. In the present project, the uplift values derived by Ekman (1996) in 58 mareographs are used. Ekman estimates the standard error of these velocities to 0.2 mm/year.

- The absolute land uplift estimated from times series at permanent GPS stations. It should be noted that this type of uplift refers to a global reference frame, which is fixed with respect to the center of mass of the Earth. For the present project, the velocities from the BIFROST project as computed by Lidberg (2004) has been the main source of GPS derived land uplift. The standard error is estimated by Lidberg (ibid.) to 0.2–0.6 mm/year, but he concludes that these values are likely too low.

- The third type of observation is repeated precise levelling. In Finland and Sweden three levellings have been performed in each country. In Norway, the levelling has (more or less continuously) been going on during the last hundred years, but has not been systematically repeated at the same lines. Since the lines from different epochs have been connected in a network, the uplift can nevertheless be determined; cf. Vestøl (2006) and Ågren and Svensson (2007). The uplift from precise levelling is with respect to the geoid. It is often referred to as the levelled uplift.

Above, the main types of land uplift observations have been summarised. It should be stressed that they yield different types of land uplift. The apparent uplift $\dot{H}_a$ is related to the uplift with respect to the geoid $\dot{H}$ (levelled uplift) as

$$\dot{H} = \dot{H}_a + \dot{H}_e \qquad (1)$$

where $\dot{H}_e$ is the Mean Sea Level rise. In this paper, the approximation is used that the absolute uplift $\dot{h}$ is linearly related to $\dot{H}$ (Ekman and Mäkinen 1996a); cf. e.g. Sjöberg (1989) for more strict formulas. Furthermore, the formulation used in practice by Vestøl (2005) is preferred:

$$\dot{h} = \dot{H} + s \cdot \dot{H} \qquad (2)$$

where s is a scale factor. Notice that the scale factor here refers to the levelled uplift and not, as it is usual in this context, to the absolute uplift. Vestøl (2005) estimates $\dot{H}_e$ to 1.32 mm and to 6% using the above mentioned observations. This relationship is assumed for NKG2005LU.

6. Construction of a land uplift model

The first part of the work to construct a land uplift model was to evaluate the existing models to find out to what extent they are suitable for the task. After that, it was decided to combine them to reach the best possible result. Below, the three main evaluated models are first described. After that, the combination procedure leading to the final uplift model is summarised. It should be pointed out that more models than the ones presented here have been considered, but only these three candidates are treated in this paper.

6.1 The model of Ekman (1996)

The first model is the one by Ekman (1996), which has been constructed using repeated levelling, the apparent uplift estimated from long time series at 58 tide gauges and some lake-level observations. The main problems with this model for the present task are that: it only uses the first and second precise levellings in Sweden; it is based on no information in much of the interior parts of Sweden, as well as in Norway; and that it does not cover the whole area of the BLR. In addition, no uplift values from GPS were available to Ekman (ibid.). Due to these reasons, it was concluded that this model is not suitable for the task.

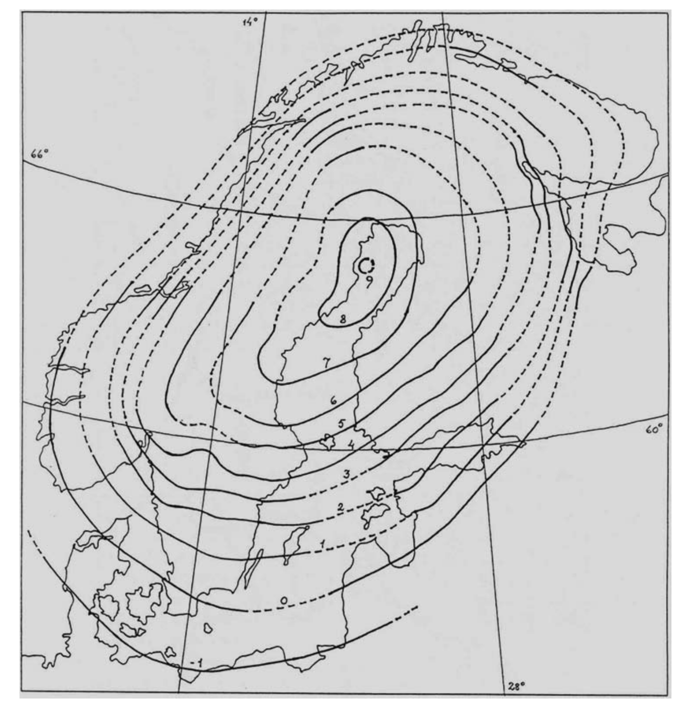


Fig. 3. Apparent uplift from the model of Ekman (1996). Unit: mm/year.

6.2 The geophysical model of Lambeck, Smither and Ekman

Another alternative is to use a geophysical model, which consists of a physical model of the lithosphere, the mantle and the ice sheet. One advantage of this kind of model is that it may provide a geophysically meaningful interpolation and extrapolation of the uplift phenomena. For instance, from the fact that the lithosphere is comparatively rigid, it follows that the land uplift model should be smooth. The lithosphere can sustain loads of smaller dimension.

The geophysical model of Lambeck et al. (1998) was the best geophysical model available at the time. This model has been tuned to the apparent uplift in the tide gauges referred to in the last subsection (Ekman 1996), some lake-level observations and ancient shorelines. The model was only available as a digital image from the publication, meaning that it had to be digitised for the purpose. The digitised version (NKG), here referred to as Lambeck's model, is presented in Fig. 4.

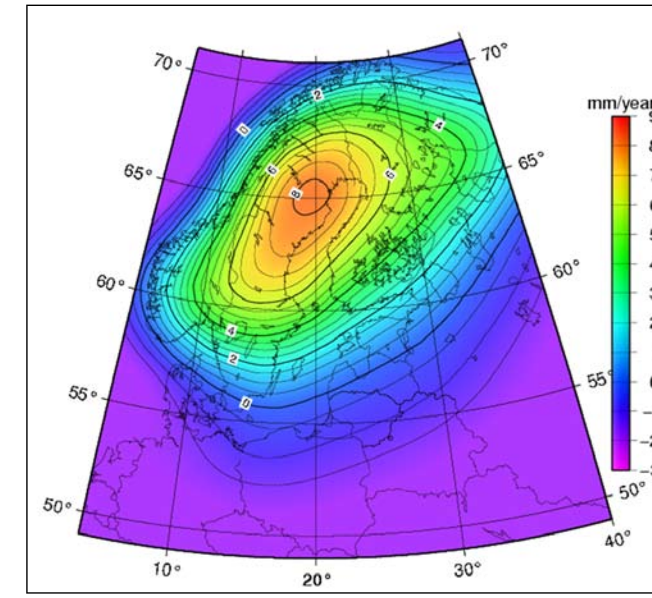


Fig. 4. Apparent uplift according to the model of Lambeck et al. (1998). Contour interval: 0.5 mm/year.

The evaluation of Lambeck's model was made by studying the residuals of tide gauge and permanent GPS observations, the latter being converted to apparent land uplift by the relationship described at the end of Sect. 5. The result is illustrated in Fig. 5

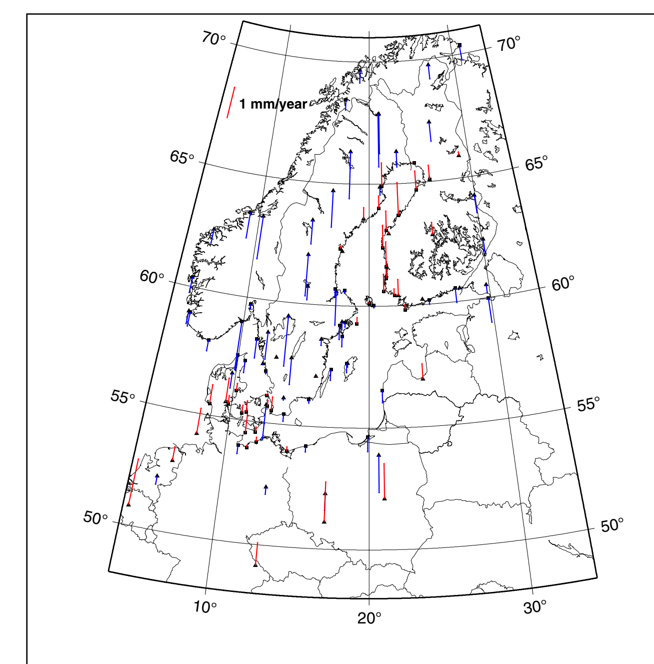


Fig. 5. Tide gauge and GPS residuals for the Lambeck model. The tide gauges are denoted by squares and the GPS stations by triangles. The scale is given by the 1 mm/year arrow to the North West.

It can clearly be seen that the model is very biased for the interior parts of Sweden. The errors are systematically as large as 1-1.5 mm/year. Mainly due to this reason, it was decided that the geophysical model of Lambeck et al. (1998) could not be used. Since it was out of the question to compute a new geophysical model, it was decided to either go for a mathematical (or empirical) model or to modify Lambeck's model for those areas in which better information is available.

6.3 Vestøl's Mathematical Model

A mathematical (empirical) land uplift model is a mathematically defined surface that has been constructed to fit the available land uplift observations in some suitable way. In 2005, one such model was developed by Olav Vestøl from the Norwegian Mapping Authority. Since then, Vestøl has continued to improve his model (Vestøl 2006), but here only the version available in January 2005 (at the RH 2000 deadline) is treated. It is presented by Vestøl (2005); cf. Ågren and Svensson (2007) for additional details. The land uplift observations are summarised in Sect. 5. By means of least squares collocation with unknown parameters using a polynomial trend surface of degree 5, the land uplift is estimated in the observation points. From these point uplift values, the land uplift is finally estimated in a grid by a simple gridding algorithm (inverse distance weighted mean of four observations, one in each search quadrant if closer than 120 km). The model is illustrated in Figs. 6 and 7.

The model agrees well with the observations but is not defined for the whole BLR, which is a drawback. The few observations outside the Nordic countries and the simple gridding algorithm lead to strange "staircase cylinders" at the outskirts of the model. This behaviour can be clearly seen in Fig. 7, especially to the South. Furthermore, the model looks a little too rough with zigzag contour lines. Consequently, some smoothing might be motivated. One problem with a mathematical land uplift model is that it is not evident how the algorithm should be "tuned".

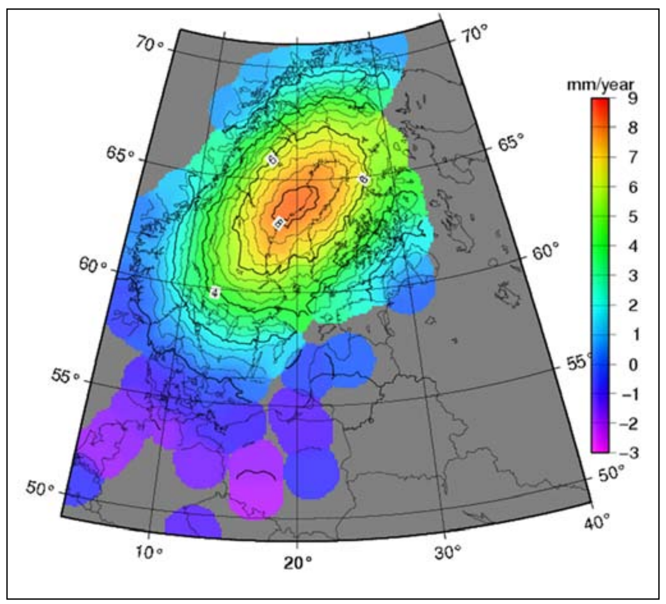


Fig. 6.Apparent uplift according to the model of Vestøl (2005). Contour interval: 0.5 mm/year.

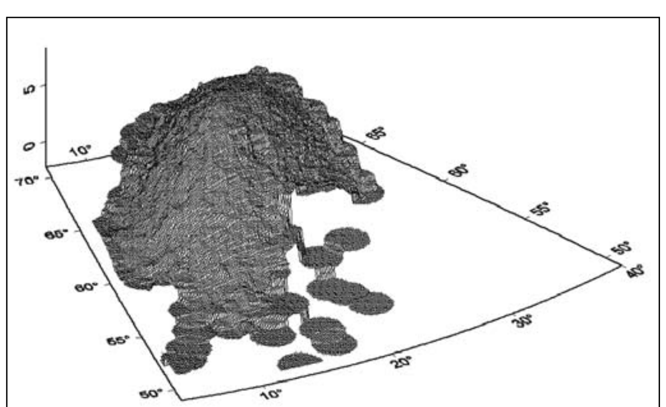


Fig. 7. Wireframe plot of the model of Vestøl (2005)

6.4 Combination of Vestøl's and Lambeck's Models (NKG2005LU)

One clear advantage of Lambeck's model is that it covers the whole BLR area in a reasonably realistic way. Vestøl's model does not. On the other hand, Vestøl's model fits much better with the observations over the Nordic countries compared to Lambeck's model. Therefore, a combination of the two models seems to be the best possible choice to cover the whole area, all the way down to NAP.

Many different ways to optimise the combination have been explored, however, these are not described in this paper. Interested readers can study Ågren and Svensson (2007). The final model, which was originally called RH 2000 LU (Ågren and Svensson 2007), is basically a smoothed version of the model of Vestøl (2005) in the central (Nordic) parts of the area. Outside this area, a smooth transition to Lambeck's model is accomplished. The RH 2000 LU model has later received a more official status within the NKG and has been named NKG2005LU. The model, as illustrated in Fig. 8, was applied in the RH 2000 adjustment of the BLR area. The tide gauge and GPS residuals are illustrated in Fig. 9.

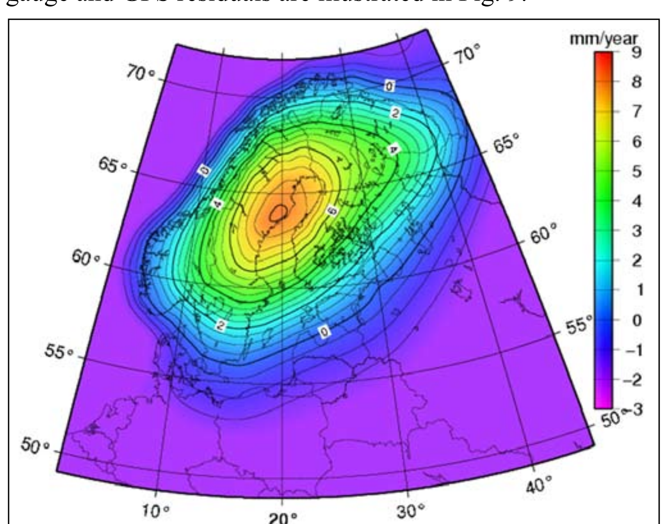


Fig. 8. Apparent land uplift according to the NKG2005LU model. Contour interval: 0.5 mm/year.

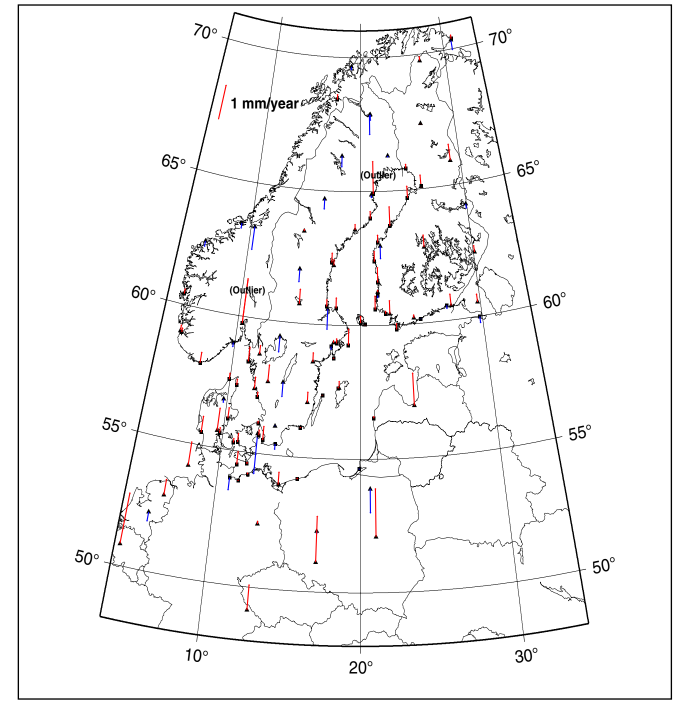


Fig. 9. Tide gauge and GPS residuals for the NKG2005LU model. The tide gauges are denoted by squares and the GPS stations by triangles. The scale is given by the 1 mm/year arrow to the North West.

If Figs. 5 and 9 are compared, it can be seen that NKG2005LU fits considerably better with the observations than the model of Lambeck et al. (1998). At the outskirts of the area the two models are similar, which is only what could be expected.

7. Final Adjustment of RH 2000

All levelling data from the whole BLR were included in the final adjustment. It should be stressed that only levelling observations were utilised. In the first step, the levelled height differences were converted to geopotential differences by multiplication with gravity (Heiskanen and Moritz 1967). A least square adjustment was then made of the geopotential differences between a total of 7 400 nodal points, of which 5 132 are Swedish. The national data sets in the BLR were given the weights determined by Karsten Engsager on behalf of NKG. The Swedish aposteriori standard error of unit weight is approximately 1 $\text{mm}/\sqrt{\text{km}}$. The estimated standard errors with respect to the NAP are illustrated in Fig. 10. As can be seen, they are approximately 2 cm in Sweden. In case the standard errors are transformed so that they refer to a fixed station in Sweden, for instance Gävle, they become smaller than 1 cm for the whole country, increasing approximately as the square root of the distance. The relative standard errors inside Sweden are thus below 1 cm, which is what matters in practice.

The result from the first adjustment, which is geopotential numbers at the nodal benchmarks, was then used as known values in the adjustment of all other benchmarks. In total, about 50 000 points have been determined in RH 2000. In the final step, the geopotential numbers were converted to normal heights.

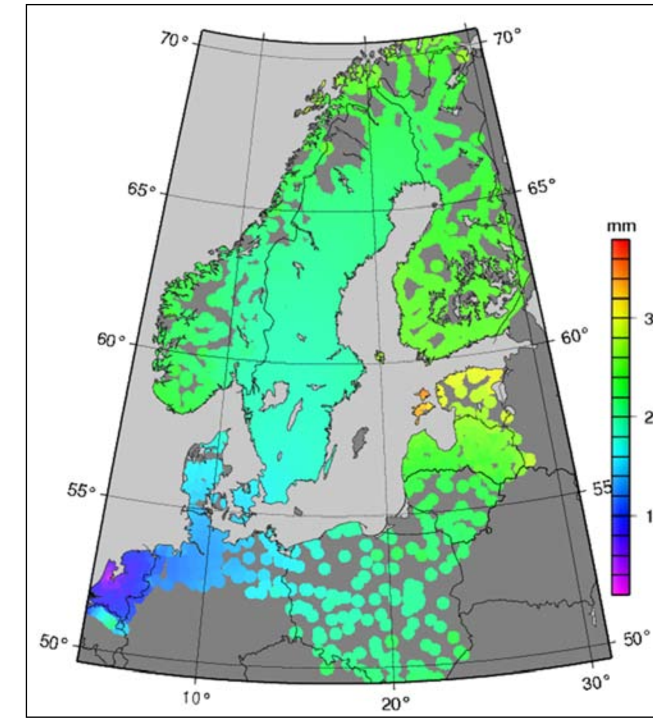


Fig. 10. Estimated standard errors relative to NAP. Unit: m.

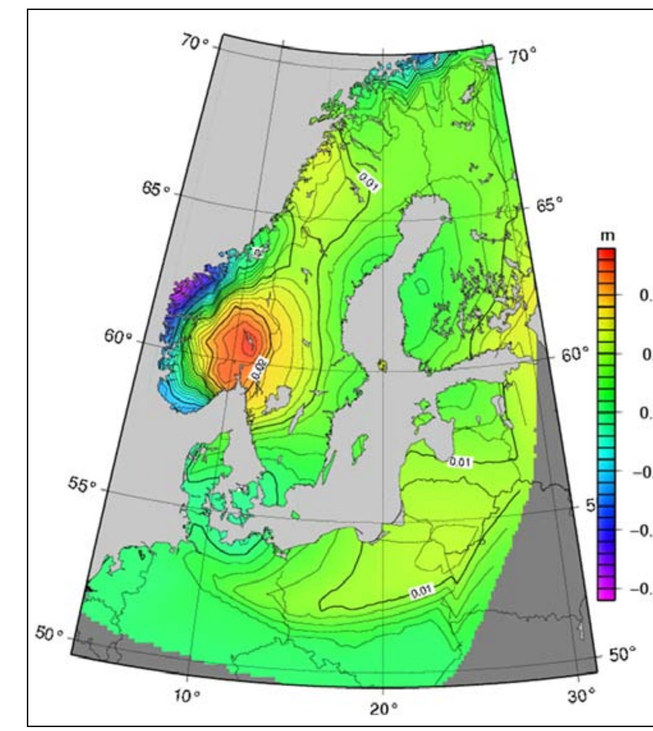


Fig. 11. Adjusted height differences between using the land uplift model of Lambeck et al. (1998) and the NKG2005LU model. Contour interval: 0.01 m.

The construction of the land uplift model NKG2005LU was described in Sect. 6. The model is similar to the model by Lambeck et al. (1998) at the outskirts of the area, but is an improvement of the latter for the central, Nordic parts. An important question at this point is how much this means in practice, as judged by the difference for the adjusted heights. The height differences between applying the two models are illustrated in Fig. 11. As can be seen, the deviations are largest in Norway, but they are definitely significant also in Sweden. Here the adjusted heights differ approximately 1-2 cm. Another conclusion is that the difference changes the geometry of the heights. It is concluded that the choice of land uplift model yields significantly different heights. Since the NKG2005LU model agrees best with the observations, it is believed that the choice of this model for the RH 2000 adjustment is warranted.

8. Comparison of RH 2000 with MSL and RH 70

Above, the choice of system definition and land uplift model for RH 2000 has been described. It is important to notice that these choices have not been performed blindly. Naturally we have studied the resulting RH 2000 heights and compared them both with the Mean Sea Level (MSL) along the Swedish coast and with other height systems. We would, for instance, not have accepted NAP as zero level in case the resulting MSL was completely inappropriate in the Baltic Sea. It is the main purpose of this section to study the MSL in RH 2000 at the Swedish coasts. Another aim is to compare RH 2000 with the older Swedish height system RH 70.

The MSL in RH 2000 at the epoch 2000.0 for four Swedish mareographs is illustrated in Fig. 12. The computation was made as a linear regression using 90-120 years of observations lasting until 2001. The data were obtained from the Swedish Meteorological and Hydrological Institute (SMHI). No corrections were applied to the sea level observations.

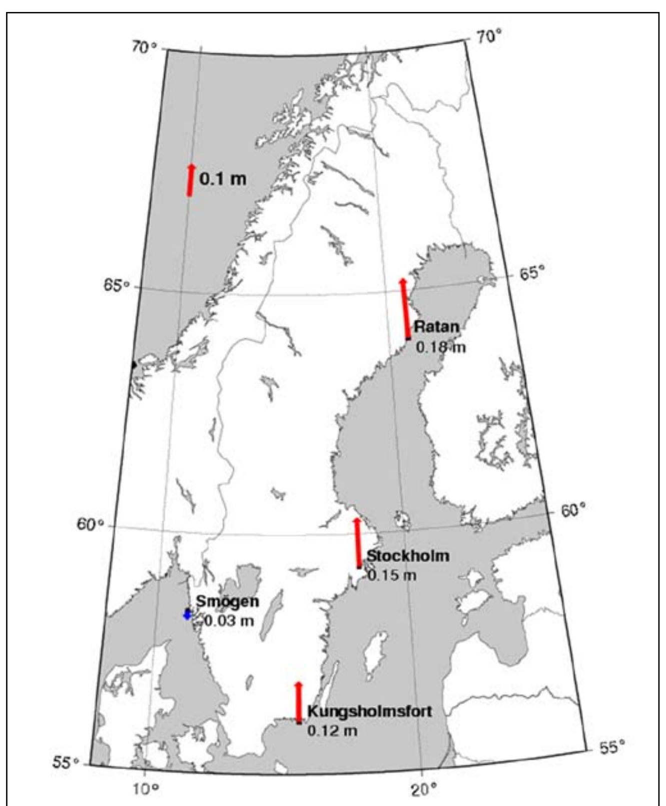


Fig. 12. Mean Sea Level (MSL) in RH 2000 at 4 tide gauges. Epoch: 2000.0.

By studying Fig. 12, it can be seen that the MSL is reasonably close to zero in the Western parts of Sweden and that the magnitude increases the further North one moves in the Baltic Sea. The main deviation is due to the sea surface topography and the fact that a zero permanent tide system is used for RH 2000; see Ekman and Mäkinen (1996b). Due to the mentioned effects, it is not possible to choose a zero level for RH 2000 so that the MSL becomes zero everywhere. Seen in this light, the obtained result seems pretty good. The MSL is almost zero at the West coast and increases the further one moves in the Baltic Sea, which is appropriate considering the sea surface topography (Ekman and Mäkinen 1996b). It should also be noticed that what is discussed here is the MSL at the reference epoch 2000.0. As times moves on, the sea level will reduce due to the land uplift. This means that the MSL in RH 2000 will become smaller and become even closer to zero. It is concluded that the choice of NAP as zero level yields a system with heights agreeing reasonably well with the MSL at the Swedish coasts. There is no reason to define RH 2000 using the mareographs along the Swedish coasts.

Let us turn now to the comparison of RH 2000 with the old RH 70 (epoch 1970.0 using a non-tidal permanent tide system). The difference in heights varies between 7 and 32 cm, which can mainly be explained by the different land uplift epochs and by the different permanent tide systems; cf. Ågren and Svensson (2007) for more details and a longer discussion. If all the known effects are corrected, the differences as

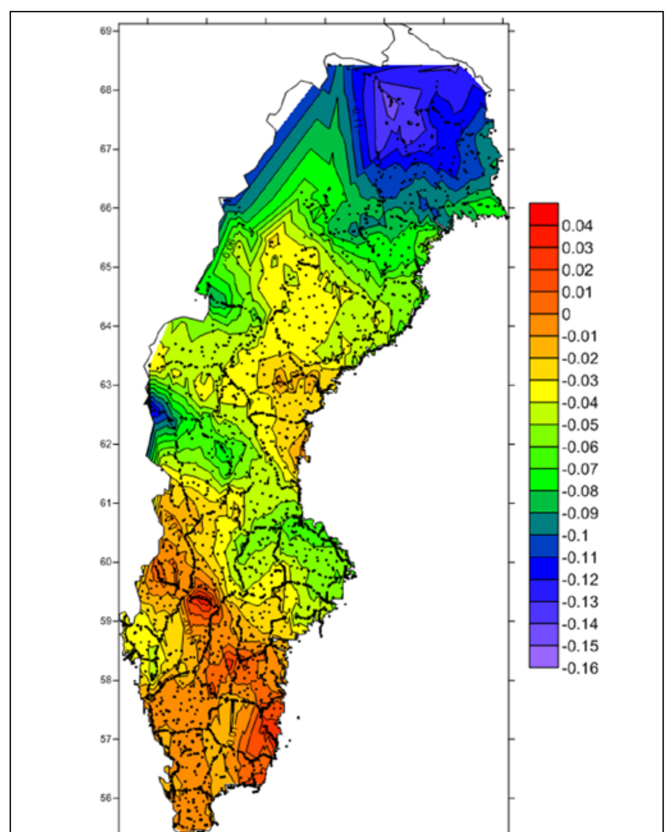


Fig. 13. Difference between RH 2000 computed with NKG2005LU and RH 70 with corrections applied for the land uplift epochs and permanent tide systems. Unit: m. Contour interval: 0.01 m.

It is believed that the differences in Fig. 13 are more or less what can be expected, considering the quality and denseness of RH 70; see Ågren and Svensson (2007). Since RH 2000 can be expected to be considerably better than RH 70, it is believed that Fig. 13 mainly illustrates the errors of RH 70.

9. The land uplift model as a part of the modern geodetic infrastructure.

It has been the purpose of this paper to describe how the new Swedish height system RH 2000 is defined as a realisation of EVRS, to present the construction of the land uplift model NKG2005LU and to study the final RH 2000 heights. Besides this, the third precise levelling of Sweden, the BLR measurements and the RH 2000 adjustment has been described in some detail.

Let us finish the paper with a short discussion of the role of the land uplift model in the Swedish geodetic infrastructure. The components considered here are the three dimensional reference system SWEREF 99 (Jivall 2001), the height system RH 2000 and the geoid model SWEN08_RH2000 (Ågren 2009) The latter relates SWEREF 99 and RH 2000 according to,

$$H_{\text{RH 2000}} = h_{\text{SWEREF 99}} - N_{\text{SWEN08\_RH2000}} \qquad (3)$$

where $h_{\text{SWEREF 99}}$ is the height above the ellipsoid in SWEREF 99, $H_{\text{RH 2000}}$ is the normal height in RH 2000 and $N_{\text{SWEN08\_RH2000}}$ is the geoid height (or more strictly, the height anomaly) from SWEN08_RH2000.

As explained above, the height system RH 2000 has been reduced to the reference epoch 2000.0 by means of the NKG2005LU model. No land uplift model was utilised to compute SWEREF 99, which is motivated by that only six weeks of GPS-observations were used to establish this system. The land uplift epoch is here 1999.5, which refers to the middle of the observation campaign. Consequently, the epochs of SWEREF 99 and RH 2000 differ by only 0.5 years. To take care of this small difference, a land uplift correction was included in SWEN08_RH2000 (Ågren and Svensson 2005). Thus, NKG2005LU has been used in the computation of both RH 2000 and SWEN08_RH2000.

At the present, no land uplift correction is needed for standard relative positioning in Sweden. In the near future, however, such corrections will be required when new points are determined in SWEREF 99 relative to the permanent GPS stations in the SWEPOS® network. The method to be used then is to first transform the known coordinates in SWEREF 99 to the observation epoch using a suitable land uplift model. After that, the new points are computed. The land uplift model is then applied backwards to get the new points in SWEREF 99 with the epoch 1999.5. This may be arranged in the way of a black box, so that the user does not notice that the correction is applied. It should be mentioned that also the horizontal components are affected by the postglacial rebound, which should be taken into account in this scheme. One should further consider plate tectonic motions. However, this is not the place to consider these developments in detail. It suffices to say that the land uplift model will be a very important part of the Swedish geodetic infrastructure in the future, which will make it possible to keep SWEREF 99 for a much longer time compared to if no such model is used. NKG2005LU is derived from the observations available i 2005. As time passes, more observations will continuously be collected, which will result in better land uplift models.

Acknowledgements

We thank all members of the NKG Working Group for Height Determination, especially Karsten Engsager, Jaakko Mäkinen, Veikko Saaranen and Olav Vestøl. We also thank Johannes Ihde and Martina Sacher, EUREF (IAG Subcommission for Europe) and the responsible agencies for providing the levelling data of the non-Nordic countries to the project.

References

Becker J-M (1984) Uppbyggandet av Sveriges nya riksnät i höjd. LMV-Rapport 1984:01, Gävle

Becker J-M, Lilje M, Eriksson P-O (1998) The Establishment of the Third Precise Levelling Network of Sweden. FIG, XXI International Congress, Brighton.

Becker J-M, Lilje M, Olsson P-A, Eriksson P-O (2001) Motorised Levelling – The Ultimate Tool for Production of Classic National Height Networks. In Drewes H, Dodson A H, Fortes L P S, Sanches L, Sandoval P (Eds.) Vertical Reference Systems, IAG Symposium Cartagena, Columbia

Ekman M (1989) Impacts of geodynamic geodynamic phenomena on systems for height and gravity. Bulletin Géodésique, 63/3

Ekman M (1996): A consistent map of the postglacial rebound uplift of Fennoscandia. Terra Nova 8: 158 165.

Ekman M (2002) Latitud, longitud, höjd och djup. Referenssystem och kartprojektioner inom geodesi, hydrografi och navigation. Kartografiska sällskapet, Gävle.

Ekman M, Mäkinen J (1996a) Recent postglacial rebound, gravity change and Mantle flow in Fennoscandia. Geophys. J. Int. 126: 229-234.

Ekman M, Mäkinen J (1996b) Mean sea surface topography in the Baltic Sea and its transition area to the North Sea: A geodetic solution and comparisons with oceanographic models. Journal of Geophysical Research 101: 11993-11999.

Eriksson P-O, Lilje M, Olsson P-A, Svensson R (2002): Validation and preparation of data for the computation of a new height system in Sweden. FIG, XXII International Congress, Washington DC.

Heiskanen WA and Moritz H (1967) Physical Geodesy. WH Freeman & Co, San Francisco

Ihde J, Augath W (2001) The vertical reference system for Europe. In JA Torres and H Hornik (Eds.) Report on the Symposium of the IAG Subcommission for Europe (EUREF) held in Tromsö, 22-24 June, 2000. Veröffentlichungen der Bayerischen Kommision für die Internationale Erdmessung, Astronomisch-Geodätische Arbeiten 61: 99-101.

Jivall L (2001): SWEREF 99 – New ETRS 89 coordinates in Sweden, analysis report. Lantmäteriet, Reports in Geodesy and Geographical Information Systems, 2001:6, Gävle, Sweden.

Lambeck K, Smither C, Ekman M (1998): Test of glacial rebound models for Fennoscandia based on instrumented sea- and lake-level records. Geophys. J. Int. 135: 375-387.

Lidberg M (2004) Motions in the Geodetic Reference Frame – GPS observations. Licentiate thesis, Department of Radio and Space Science with Onsala Space Observatory, Chalmers University of Technology, Göteborg

Mäkinen J. Lidberg M, Schmidt K, Takalo M, Lilje M, Engsager K, Eriksson P-O, Jepsen C, Olsson P-A, Saaranen V, Svensson R, Vestøl O (2004) Future height systems in the Nordic countries, their relation to the EVRS2000 and to INSPIRE GIS standards. Publication No. 13, Report on the Symposium of the IAG Sub-commission for Europe (EUREF) held in Toledo, 4-7 June, 2003. Verlag des Bundesamtes für Kartographie und Geodäsie, Frankfurt am Main.

Sacher M, Ihde J, Liebsch G, Mäkinen J (2009): EVRF2007 as Realization of the European Vertical Reference System. Bulletin of Geodesy and Geomatics No.1 2009.

Schmidt K (2000) The Danish height system DVR 90. Kort & Matrikelstyrelsen. Publications 4. series 8.

Sjöberg LE (1989) The secular change of gravity and the geoid in Fennoscandia. In S Gregersen and PW Basham (Eds.) Earthquakes at North-Atlantic Passive Margins: Neo-tectonics and Postglacial Rebound, pp. 125-139, Kluwer Ac. Publ.

Vestøl O (2005): Land uplift model, documentation – An empirical model based on levelling, tide gauges and GPS-rates, January 2005.

Vestøl O (2006) Determination of Postglacial Land Uplift in Fennoscandia from Levelling, Tide-gauges and Continuous GPS Stations using Least Squares Collocation. J Geod 80: 248-258.

Ågren J, Svensson R (2006) Land uplift model and system definition used for the RH 2000 adjustment of the Baltic Levelling Ring. In Knudsen (ed): Proceedings of the 15th General Assembly of the Nordic Geodetic Commission, Copenhagen, May 29 – June 2 2006, Technical Report No. 1, DTU Space, Denmark, pages 84 – 92.

Ågren J, Svensson R (2007): Postglacial Land Uplift Model and System Definition for the new Swedish Height System RH 2000. Lantmäteriet, Reports in Geodesy and Geographic Information Systems, 2007:4, Gävle, Sweden.

Ågren J (2009): Beskrivning av de nationella geoidmodellerna SWEN08_RH2000 och SWEN08_RH70. Reports in Geodesy and Geographic Information Systems, 2009:1, Gävle, Sweden.

USER CENTRED DESIGN APPROACH TO IMPROVE NAVIGATION MAPS FOR VISUALLY IMPAIRED USERS

Hanna Stigmar[1] and Kirsten Rassmus-Gröhn[2]

[1]GIS Centre, Lund University, Sölvegatan 12, S-223 62 Lund, Sweden
[2]Certec, Department of Design Sciences, Lund University, Box 118, 221 00 Lund, Sweden
hanna.stigmar@lantm.lth.se, kirre@certec.lth.se

Abstract

This paper presents the methods and results of a study aiming at improving visual presentation of navigation information for persons with visual impairment, as well as sighted persons. The overall method used was a sequential, longitudinal combination of user centred techniques. The methods applied were: focus group discussions combined with a diary study, survey questionnaire, and usability test. The combination of the methods revealed a promising combination of visualizations to use for visual presentation.

Keywords: *Usability test, focus group discussion, diary study, survey questionnaire, map, location based services, pedestrian navigation, visual impairment.*

1 Introduction

Location based services (LBS) and their applications are becoming more and more user-friendly. A large number of applications have become available, in line with users' wishes and needs. Nowadays it is also possible for the user to apply personal settings, in order to make the applications even better.

The maps in LBSs are often based on traditional design. This is not optimal for all users. Visually impaired people, for example, obviously have other needs than fully seeing people. For a visually impaired user it can be difficult to match the map to the real world, as many real-world objects cannot be seen. In order for people with vision problems (or people who need their eyes for something else) to fully take advantage of the information it needs to be presented to the user in a proper way.

The European project HaptiMap (Haptic, Audio and Visual Interfaces for Maps and Location Based Services) aims at making maps and location based services more accessible by using several senses like touch, hearing and vision. This paper presents a study that focuses on the visual part of the service, and how to make the presentation of the information as good as possible for sighted persons as well as for persons with visual impairment (including elderly people). The aim is to present a map that will provide the most important visual information in a way that interferes as little as possible with the navigation task.

2 Background

This section presents three aspects that are important to consider when designing maps for location based services. They are: what information to present, what kind of visualization type to use, and how to choose the colours.

2.1 Information types to present

The types of information to display in an LBS map should be determined based on external as well as internal issues and limitations. External issues are the kind of application, the user's preferences and limitations, and alternative signals (sound and haptic signals) that can be used. Internal limitations are the map size and map legibility.

The kind of application has a central role concerning the types of information to display. The information should be useful in its context. In a navigation application the most important information is that which gives the user guidance to his/her destination (for example when to turn and when to walk straight ahead, and what to look for). In order to give the user a sense of the navigation environment, and get valuable "extra" information, surrounding information can be useful (for example the street network, buildings, landmarks, and other points of interest).

The user's abilities and limitations are very important aspects. These set the prerequisites for which information the user will be able to perceive, and how it will be perceived. Because of their limitations, visually impaired people often need other types of information than the information that is presented on a traditional map. The information should give the user an idea about the surroundings that cannot be seen. In order to simplify a navigation task, the user often divides the route into several sub-routes with the help of objects or sounds that are noticed along the way. The application should take this into account and provide information about landmarks and context, as well as the remaining distance (Magnusson et al. 2009b, Loomis et al. 2005, Golledge et al. 2004). For visually impaired users the size of symbols, the density of them, and the symbology by which they are presented are vital aspects. The user's preferences are also important usability issues. Each user is different, and by enabling adaptive application settings much can be won. This can concern zoom levels, kind of egocentricity, symbol size, colour scales, etc.

The information that should be displayed thus depends on a number of different issues. It should be the most important information that cannot be given equally well by other signals. The information must also be presented in an appropriate way. These limitations are rather similar to the ones that exist for tactile maps. Apart from the fact that tactile navigation maps present more, and somewhat different, information than corresponding maps for seeing people, they give a good idea of what information can be seen as the most important. Generally, tactile maps are highly generalized, while keeping typical shapes and features. For tactile city maps, city parts and major streets are enough, while for city part maps also buildings can be included. For very large-scale maps the major streets, some buildings, and footpaths are presented (Eriksson et al. 2003). These large-scale maps can be regarded as corresponding to the navigation maps in our application.

2.2 Visualization types to use

The visualization type to use really has to do with which interface to choose for the presentation of the information towards the real world. The most common visualization type for geographic information is the 2D map. However, the use of 3D and so-called 2.5D maps is also becoming quite widespread.

When it comes to screen based presentations and web pages, one important aspect is clearness of line. Generally, guidelines tell us that to increase the usability, the presentation should be clear and simple, without unnecessary elements. It should be simple to identify each element based on its own geometrical characteristics. Good examples are: two-dimensional graphics, comic-like graphics, images drawn with clean lines, images drawn with thick lines, and simple traits, stylized. Bad examples are: photos, three-dimensional graphics, sketches in a pictorial style, uncertain border lines that are overly detailed, and thin lines (Dini et al. 2006). However, it has also been suggested that since we live in a three-dimensional world it is of interest for us to be able to picture the world in three dimensions as well, to make the image similar to our own view. In addition to providing aesthetically pleasing real-world perspective, these representations also provide a communication support. In this way 3D cartographic products provide a more intuitive acquisition of space than their 2D counterparts, and therefore "fit" into our already existing mental models better (Jobst and Germanchis 2007, MacEachren 1995, Hay 2003).

Evaluations of navigation with 2D maps versus 3D maps have been performed for example by Porathe (2006) and Schilling et al. (2005). The study performed by Porathe showed that egocentric 3D maps provided the fastest decision-making and the least mistakes. Schilling et al. did not get same clear result as Porathe did. For them no significant difference could be shown between 2D and 3D. However, as they point out, because the test was performed outdoors there were some problems with sunlight on the screen used for the 3D map (as opposed to the used 2D paper map) and with the GPS coverage. Also, the test persons in this study clearly stated beforehand that they preferred to use 2D maps. Despite this, the 2D maps did not show a better result.

Some drawbacks can be found when using cartographic 3D representations, however. Like the real world 3D views, the representations contain distortions due to linear perspective, hidden objects, various scales in the same picture. On the other hand, these drawbacks can be somewhat reduced by using multimedia techniques such as orthographic scene rendering or interactive camera movements (Jobst and Germanchis 2007).

As for 2D representations, 3D representations can be visualized either photorealistic or non-photorealistic, or perhaps with elements of both. There has been a tendency in Virtual Reality (VR) that visualizations should be as photorealistic as possible, to give the best possible reference to the real, or alternative, world represented. The representations are often regarded as more impressive and truthful this way. Apart from correct geometries photorealistic representations use true colours and textures to make the presentation as real-looking as possible. This can be done by hand, but it can also be done using photos of the objects in question. Kang et al. (2007) have used photos to create true textures of building facades and roads. The textures have been generated using a method similar to generating ortho-images. This way the textures get an even more real-looking appearance. In some aspects, non-photorealistic visualizations can be more effective. For visualizations of geospatial data one can see the similarities with cartographic generalization in general. With the non-photorealistic visualizations, details such as texture and colour can be simplified. These "generalizations" lead to a simpler and often easier-to-grasp representation. Instead details such as enhanced edges, emphasized illumination and shading, and terrain illustrations can be used to make the objects more "realistic". Also the choice of colours has great impact. Often it is favourable to base the colours on semantics and aesthetics, and to keep the number of colours limited. Furthermore, other features, such as height of buildings, can be exaggerated or reduced in order to make the representation more similar to the user's view (Döllner 2007).

2.3 Colours to use

Colour is one very important aspect when it comes to all kinds of visualizations. Proper use of colour can facilitate and improve the interpretation and reading of the presentation. Visu-

al impairment affects perception, which may reduce also the visual effectiveness of certain colour combinations. In order for more people to be able to use a visual application properly it is important to use the contrast of colours. However, often the ability to discriminate between different colours (hues) is diminished with colour deficit. Instead, lightness is the most effective attribute when creating contrast (Lighthouse 2009). Three guidelines are given by Lighthouse (2009):

- Exaggerate lightness differences between background and foreground, and avoid using colours of similar lightness next to each other.
- Use dark colours of red, purple, violet, or blue, together with light colours of blue-green, green, yellow, or orange.
- Avoid using contrasting hues from adjacent parts of the hue circle (see Lighthouse 2009), but if doing so make sure they differ in lightness.

According to Vision Australia (2009) good colour visibility for foreground and background is achieved if both brightness difference and colour difference are greater than threshold values. The *brightness difference* (Δbr) is defined as the absolute difference in brightness (br_1, br_2) for the two neighbouring objects:

$$\Delta br = |br_1 - br_2| \qquad \text{(eq. 1)}$$

where

$$br_1 = \frac{(red_1 \cdot 299) + (green_1 \cdot 587) + (blue_1 \cdot 114)}{1000}$$

red_1 etc. is the colour component for the first object, and br_2 is defined analogously.

The colour components are given as RGB values (between 0 and 255). The difference in brightness for two adjacent colours should exceed 125.
The colour difference (Δh) is defined as:

$$\Delta h = |red_1 - red_2| + |green_1 - green_2| + |blue_1 - blue_2| \qquad \text{(eq. 2)}$$

where
red_1 etc. is the colour component for the first object, and red_2 etc. is the colour component for the second object.

The colour values are given as RGB values (0 - 255). The difference in colour for two adjacent colours should exceed 500. Thus, the requirement of having a brightness difference larger than 125 and a colour difference larger than 500, gives us the possibility to have three (for colour brightness this is just barely three) adjacent colours and still attain a good colour visibility.

Other guidelines are provided by Dini et al. (2006) and refer to screen based presentations and web pages. Concerning background features, these should not be overly strong as this will make it more difficult to visually separate background from foreground. Preferably, the background colour should be uniform, or in sharp contrast with the foreground elements.

A fairly common visual impairment is colour blindness. This impairment affects about four percent of the population (more men than women). Being colour blind is a rather mild impairment when it is still possible to see lightness difference and quite many hue differences. Red-green colour blindness is the most common, but there are others (Brewer 1999).

An interesting remark, made by Brewer (1999), is not to become too concerned about which colours people will like. People have different taste, and there will always be somebody who does not like the colours that you do. Thus, the organization and matching of colours for presentation is more important.

3 Methods

In this study, we have evaluated different types of map and navigation visualizations in order to find which is preferable for use in pedestrian navigation. The overall method we have applied is a sequential, longitudinal combination of user centred techniques, which is in line with the HaptiMap user study guidelines (Magnusson et al. 2009a). The user activities have also been combined with needs elicitation in the project on other parts of the service beside the purely visual one (see also Magnusson et al. 2009b). In this paper, however, we only report on the parts that have relevance for the visual maps. In short, there are three user activities that have led to a map usability test: focus group discussions, a diary study and a questionnaire. The aim was to start the design process with a wide perspective, and then step by step narrowing it down in order to find the most appropriate final design. This approach has also been taken for example by Nivala et al. (2009) who used questionnaires, empathy probes and focus-group discussions to identify potential users and their tasks during a hike.

3.1 Focus group discussions and diary study

Focus group discussions aim at giving potential users the possibility to freely discuss their opinions about a specific question. The discussions can be structured to some extent, for example concerning the issues and sub-issues to discuss, but should generally pick up the users' opinions and continue from there. A focus group should be composed of potential end users, preferably 3-6 persons per group. Too large groups may make it difficult for each participant to express his or her opinions, while in too small groups it can be difficult to get the discussion started and run freely. The advantages of focus group discussions are the participants' freedom to express their opinions. This has the potential of giving results that the designer might not have foreseen. It also has the potential of providing the designer with the reasons why the users think the way they do, which can be helpful in predicting their behaviour. Some disadvantages can be that it may be difficult to supervise the discussion and keeping it at a desired level while not letting the designer's own opinions colour the discussion. More information on focus groups and how to successfully carry out discussions can be found in Magnusson et al. (2009a) and Morgan (1997). Focus group discussions have, for example, been used by Nivala et al. (2005) in order to evaluate a mobile map service prototype.

Diaries emerged into the tool palette of user-centred designers at the outset of the extensive interest in cultural probes (see e.g. Gaver et al. 1999). Diaries have been employed within ethnography, written by the ethnographers themselves, while in design-related diary studies, the potential users are the authors of the diaries. In a typical diary study, the participants are asked to keep record of activities and thoughts on a specific issue. The benefits of a diary are that the written format may enable participants to express such issues about their lives that might be omitted when asked directly in speech. The diary also gives people more time to think about the subject. As a diary study explicitly encourages people to look at and think about their lives, it renders them sensitive to the issues – in their own context – that may be of relevance for a design project. For that reason, diaries are a good means for preparing users for project-related interviews. More about diary studies can be read in Koskinen et al. (2003). Sohn et al. (2008), for example, used a two-week diary study in order to capture users' mobile information needs.

Case description

In this study we had two focus group discussions and a diary study. The diary and the preceding first focus group discussion served mainly as a preparation for the users. The two focus group discussions each lasted two hours, and were carried out with one week in between. In the week between the focus group meetings, the participants were asked to fill in the diary, which contained questions about their travel, and questions about current travel aids that they used and also travel aids that they might imagine they could use.

As part of the second focus group discussion (which also contained a hands-on workshop, see Magnusson et al. 2009b) 16 different example maps, as well as a pictogram map, had been prepared. The maps were of the kinds that are often used in today's navigation services: 2D maps with streets, buildings and land use. They were of different scales and generalization levels, and with commonly used symbologies. The participants were asked to give their opinions on these maps; whether they were considered as good or bad for use in a navigation service, why so, and what a perfect map would look like.

Our target users were elderly and people with visual impairments. The participants were recruited in two ways. We already had a list of persons with visual impairments, and recruited among them. The elderly were recruited via a "Senior university" (Pensionärsuniversitetet). In total, we had 8 users, who were divided into two groups.

Case results

The discussions lead to the conclusion that the participants wanted a visualization that is more alike the real world, and with more detailed information. This could be enabled by using for example photo-realistic visualizations or 3D maps with relative height of buildings. Also, more detailed information was requested, for example buildings with entrances, open spaces with exits, and logos for shops and restaurants. This information should be presented at a large-scale level. The colours of the symbology should be bright in order to facilitate map reading.

3.2 Survey questionnaire

Questionnaires can be used for experimental or exploratory user studies. A well-designed questionnaire used effectively can gather information on different aspects of design and potentially from a large number of participants. Questionnaires rely on self-reporting. For this reason, they are not an objective source of data and may sometimes express a phenomenon known as "social desirability" (the tendency to answer based on how people think questions should be answered, rather than expressing their own opinion). More details on questionnaires can be found in Courage and Baxter (2005) and Visser et al. (2000). Elias et al. (2009) performed a questionnaire-based user test in order to find the importance of landmarks in navigation, while Sayda (2005) used a web-based questionnaire to validate an LBS prototype for hikers.

Case description

In order to follow up the discussions about the visual maps, a survey questionnaire was developed. The aim of the questionnaire was to get some opinions on different types of visualizations that had been developed based directly on the result of the focus group discussions. Here the target users were not specifically visually impaired people, but rather "any user". In this context, the questionnaire was used to identify the type of visualizations that had the potential to be the most useful in an LBS navigation service visualization. 11 different types of navigation maps or visualizations were presented, and the participants were asked to select the visualization(s) (three or less) they preferred for an LBS navigation service visualization. All 11 visualizations are shown in Appendix A. The questionnaire took approximately 10 minutes to answer.

Our survey questionnaire was distributed at a convention, and was answered by 26 visitors; 12 men and 14 women, with an average age of 46 years. Most of the participants (18) had hardly or never used a GPS navigator for navigation.

Case results

The results of the questionnaire led to the conclusion that the visualization considered to be most useful was a small-scale 2D topographic map with buildings. This was followed by a semi street-level 3D map, a pictogram, a small-scale satellite image, a small-scale 2D topographic map without buildings, and a small-scale photorealistic 3D visualization, in that order. These six visualizations are shown in Figure 1(a-f), together with the number of votes (out of a total of 53).

Figure 1(a)
The small-scale 2D topographic map with streets, buildings, and land use was considered the most useful
(10 votes).

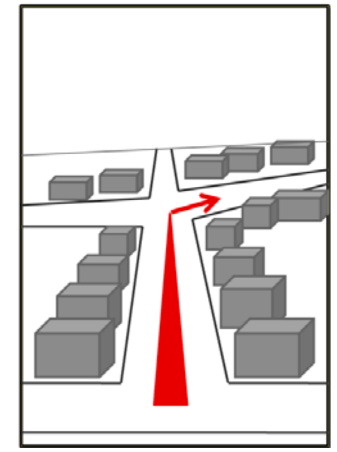


Figure 1(b).
The simplified semi street-level 3D map was considered the second most useful.
(8 votes).

Figure 1(c)
The pictogram was considered the third most useful (7 votes) together with the small-scale satellite image.

Figure 1(d)
The small-scale satellite image was considered the third most useful (7 votes) together with the pictogram.

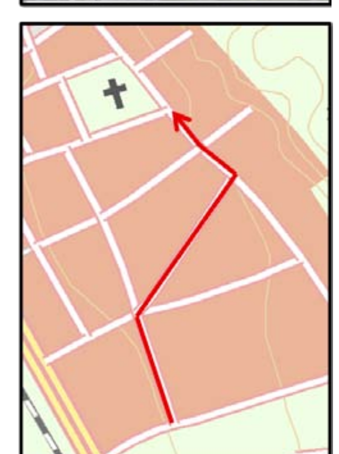


Figure 1(e)
The small-scale 2D topographic map without buildings was considered the fifth most useful (6 votes) together with the small-scale photorealistic 3D visualization.

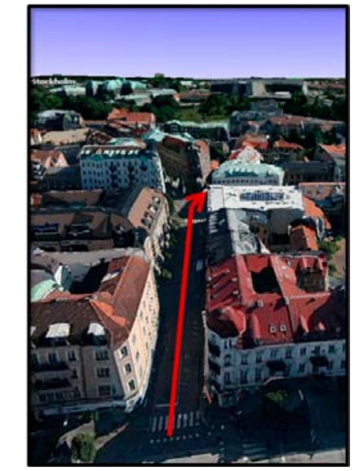


Figure 1(f)
The small-scale photorealistic 3D visualization was considered the fifth most useful (6 votes) together with the small-scale 2D topographic map without buildings.

3.3 Usability test of visual maps

Usability testing that evaluates the use of (especially) technical products is a common activity in the interaction design process. Usability has a number of different definitions, one being the ISO 9241-11 from 1998:

"...the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use..."

Schneiderman (1992) and Nielsen (1993) both divide the usability concept into five different qualities: learnability, efficiency, memorability, error handling and satisfaction. There are also different ways of conducting usability tests, and it is possible to conduct usability tests with different levels of prototypes. A usability test typically is task focused, i.e. a user is asked to perform a certain type of task using the product under evaluation. Different types of measures can be collected during a usability test, e.g. the number and severity of errors the user makes or the time it takes to complete a task.

Traditionally, most usability tests are conducted in a lab environment, among other things because of the possibility to isolate variables and simplify the data collection. Recently, the discussion about the need to do usability tests in context has re-emerged with the use and evaluation of mobile devices (cf. Nielsen et al 2006). When evaluating location based services, like navigation applications based on GPS, lab settings are particularly difficult, as the typical task involves movement (walking, cycling) in a larger area. Furthermore, the natural environment for device aided navigation is outdoors (at least using GPS) that brings with it added diversity to the usage situation like sun glare on the screen, snow fall, traffic noise and numerous other issues, which are hard to simulate indoors. There are also indications (Nielsen et al. 2006) that the use in context will elicit more usability problems with a product. Usability tests are usually concluded by a post-test questionnaire or interview along with a debriefing. Field-based usability tests for mobile map services have for example been performed by Nivala et al. (2003) and Dillemuth (2005).

Case description

Based on the results from the pilot study and the survey questionnaire, we decided to evaluate three types of visual navigation aids in a situated (carried out in a realistic context) mobile usability test. The three types of visualizations were: a 2D map view, a 3D map view, and a photo view. One example of each type of visualization, showing the same part of the route, is given in Figure 2(a-c).

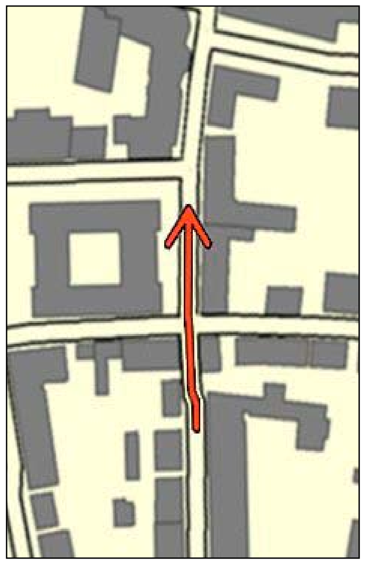


Figure 2(a)
The 2D visualization used in the study.

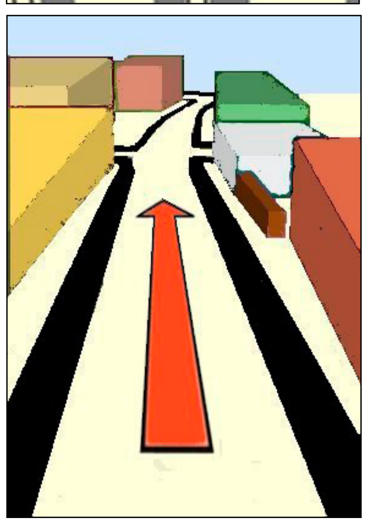


Figure 2(b)
The 3D visualization used in the study.

Figure 2(c)
The photo realistic visualization used in the study.

The 2D map was chosen based on the results of the questionnaire, where a small-scale 2D map was considered the most useful. For the usability test, the 2D map was developed bearing in mind design guidelines for visually impaired people (see Lighthouse 2009, Vision Australia 2009). Only three colours were used in the map in order to keep the contrasts and lightness differences large. The map was of scale 1:3000, and is shown in Figure 2(a).

The 3D map was particularly requested by the visually impaired participants in the focus group discussions. They wanted 3D visualizations with simplified geometries and bright colours ("comic like"). This type of visualization was also considered the second best in the survey questionnaire study. It has been suggested that 3D presentations are unfavourable to use for visually impaired people, because of shadows etc. (Dini et al. 2006). Therefore we produced a clean 3D visualization with the buildings as boxes in relative height with realistic but somewhat enhanced colours. The map had a just above-street-level perspective (similar to current car GPS views). The map is shown in Figure 2(b).

The idea of a photorealistic visualization was originally a request from a focus group participant who wanted the visualization to show what the user could not see herself. This type was also considered the sixth most useful in the survey questionnaire study. The visualization was produced by taking photos of the route (in street-level) and adding navigation information (arrows). In order to make the pictures more suitable for presentation on a screen, the colours were somewhat enhanced. The visualization is shown in Figure 2(c).

In our study, the test consisted of a navigation task in a small-town urban environment where the participants were asked to walk a 1.5 km route while guided by a visual navigation aid prototype. The prototype consisted of a smart phone with touch screen and a sequence of adapted guiding visualizations that each showed the next navigation action (i.e. walking straight forward, or taking a turn). To simplify the mock-up, the test participants were to change the navigation images by hand at certain positions indicated by the test supervisor. The route had been divided into three parts so that each participant could try each type of visualization.

The navigation task was performed individually. Each user was given a different sequence of the guiding pictures, to eliminate learning effects and effects that were tied to a certain part of the route. During the navigation task, the test supervisors were walking behind the participant, giving instructions on when to change to a new navigation image, and information about an upcoming change of visualization type. The three sub-routes were timed, and hesitations as well as mistakes were noted. The navigation task was followed by a post-test semi-structured interview about the maps and the task. The questions concerned the participants' opinions about the visualization types, and whether they thought these types would work in a mobile navigation service. The whole test, including the interview, took approximately one and a half hour to perform. The test was performed three times, each time with two participants.

Six persons participated in this map usability test. Two were fully sighted, two had moderate visual impairments and two were elderly.

Case results

The results of the usability test showed that the three visualizations worked well. The recorded times for each sub part of the route did not show any differences between the different visualizations. The few difficulties some of the participants had with the 2D and 3D visualizations also seemed to be caused by them not being used to that particular type of map, and could probably be avoided with some more training.

The post-test discussions gave some more qualitative results concerning the participants' experiences and opinions of the maps. Most participants regarded all three visualizations as fairly good, but the photo visualization as the best. The reason was the ease it provided for orienting oneself. The 2D map was also regarded valuable because of the overview it provided. However, the two visually impaired participants considered the 2D map as somewhat difficult to read, although they did comment on it as having good colour contrasts. The 3D visualization was considered as the least good. Most participants considered it more difficult to "translate" to the real world because of the simplified graphics.

A "perfect" navigation aid should, according to the participants, be designed as a combination of detailed turning instructions (given as photos) and an overview (given as a 2D map or a satellite image). The two fully sighted participants wanted these two types of images to be displayed at the same time on a split screen image, while the visually impaired and elderly participants wanted them displayed one at a time with the possibility to flip between them.

4 Discussion and conclusion

In this study, a combination of four user centred design methods was used to investigate accessibility aspects of visual presentation in a navigation service for pedestrians. The four methods were applied step-by-step and have given us information to the next functional prototype in our project.

The extensive preparation before the usability test was done in part because there were other issues than the visual maps that were discussed and investigated. Otherwise, this may be a too expensive approach. However, we feel that these preparations have been quite valuable. If forced to cut out any of the steps described above, the first candidate would be the first focus group discussion, then the questionnaire, and last the diary study. The second focus group discussion seemed to be the single most important activity to inform the usability test design; however, it is hard to tell if that discussion would have been as valuable without the preceding meeting or the diary.

The questionnaire provided additional, and more general, information as 26 participants answered (to compare with the 8 persons in the focus groups). The questionnaire could also have been sent out to a broader audience, perhaps also using a web questionnaire. And as it was, the informants did not have much GPS experience, and it would be interesting to ask experienced GPS users on their preferences as well. The result of the questionnaire did agree quite well with the result of the focus group discussions. The simplified 3D map requested by the focus group participants was selected as the second most useful by the questionnaire participants, while the very detailed satellite images and photo-realistic visualizations (although small-scale) were selected as the third and fifth most useful.

In the usability test, the use of measurement of the walking time as an indication of performance can be discussed. In a navigation task like ours, the extra time spent when making a mistake seems to disappear in the total time for the navigation task. Equally, the time or effort spent by a user to interpret the map with a less favourable visualization seems not to influence the walking time for the users. The number of mistakes and the answers to the post-test questions have been the largest sources of information, and the users also commented on which maps they found to be most straining to use. A more quantitative approach to collect information on mental demand would be to perform a work load investigation (e.g. NASA TLX, NASA 2010). Also, since our test was based on observation by a test supervisor walking behind a user, it was hard to note the exact number of times a user looked at the screen, a figure which would potentially inform us about the usability of a certain map type.

The usability test showed that the 2D map, the 3D map, and the photorealistic visualization all worked well. However, the participants regarded the photorealistic visualization as the easiest to use as it provided great help in orienting oneself. The participants also valued the overview given by the 2D map. A good solution would therefore be to combine these two. This would provide the possibility to get an overview and very detailed information at the same time. The sighted persons wanted the two visualization types displayed at the same time on a split screen image. The visually impaired and elderly participants, however, instead wanted the two visualizations displayed one at a time with the possibility to flip between them.

This type of sequential, longitudinal study is rather time consuming, which makes it difficult to involve a large number of test persons. However, according to Nielsen (1993), major usability problems can be identified also using quite small user groups. A number of three to five users in a group is considered as an optimal size in order to perform many tests.

None of the participants seemed to have any problems with the colours or colour contrasts. However, as there are many types of visual impairment, it would be preferable if a navigation service could offer, if not free colour palettes, at least a number of different colour schemes for the user to choose from. This way, personalization can move yet another step forward.

Acknowledgements

The authors are grateful to the European Commission which co-funds the IP HaptiMap (FP7-ICT-224675). We would also like to thank all our test persons, as well as our HaptiMap colleagues for valuable comments.

References

Brewer, C. A., 1999. Colour Use Guidelines for Data Representaton. In: Proceedings of the Section on Statistical Graphics, American Statistical Association, Alexandria VA, pp. 55-60.

Courage, C. and Baxter, K., 2005. Understanding your users: a practical guide to user requirements. Elsevier & Morgan Kaufmann Publishers.

Dillemuth, J., 2005. Human-map Interaction for Mobile Cartography. Masters Thesis, University of California, Santa Barbara, USA.

Dini, S., Ferlino, L., Gettani, A., Martinoli, C. and Ott, M., 2006. Educational software and low vision students: evaluating accessibility factors, Universal Access in the Information Society, Vol. 6, No. 1, pp. 15-29.

Döllner, J., 2007. Non-Photorealistic 3D Geovisualization. In: Cartwright, W., Peterson, M. P. and Gartner, G. (eds.), Multimedia Cartography, Springer, pp. 229-240.

Elias, B., Paelke, V. and Chaouali, M., 2009. Evaluation of User Variables in Topographic Feature Recall for the Informed Selection of Personalized landmarks. In: Gartner, G. and Rehrl, K. (eds.), Location Based Services and TeleCartography II, From Sensor Fusion to Context Models, Lecture Notes in Geoinformation and Cartography, Springer-Verlag Berlin Heidelberg, pp. 121-136.

Eriksson, Y., Jansson, G. and Strucel, M., 2003. Taktila kartor – Handledning i kartframställning (in Swedish). Svensk punktskrift, Sweden.

Gaver, B., Dunne, T. and Paceni , E., 1999. Design: Cultural probes. Interactions, Vol. 6, Issue 1, pp. 21–29.

Golledge, R. G., Marston, J. R., Loomis, J. M. and Klatzky, R. L., 2004. Stated Preferences for Components of a Personal Guidance System for Nonvisual Navigation, Journal of Visual Impairment & Blindness, March, pp. 135-147.

Hay, R., 2003. Visualisation and Presentation of Three Dimensional Geoscience Information. In: Proceedings of the 21st International Cartographic Conference (ICC), Cartographic Renaissance, Durban, South Africa, pp. 374-382.

Jobst, M. and Germanchis, T., 2007. The Employment of 3D in Cartography. In: Cartwright, W., Peterson, M. P. and Gartner, G. (eds.), Multimedia Cartography, Springer, pp. 217-228.

Kang, Z., Zhang, Z., Zhang, J. and Zlatanova, S., 2007. Rapidly Relizing 3D Visualisation for Urban Street Based on Multi-Source Data Integration. In: Cartwright, W., Gartner, G., Meng, L. and Peretson, M. P. (eds), Geomatics Solutions for Disaster Management, Springer, pp. 149-163.

Koskinen, I., Battarbee, K. and Mattelmäki, T., 2003. Emphatic Design. User Experience in Product Design. IT Press/Edita, Helsinki.

Lighthouse, 2009. Lighthouse International, available at: http://www.lighthouse.org/.

Loomis, J. M., Marston, J. R., Golledge, R. G. and Klatzky, R. L., 2005. Personal Guidance System for People with Visual Impairment: A Comparison of Spatial Displays for Route Guidance, Journal of Visual Impairment & Blindness, April, pp. 219-232.

MacEachren, A. M., 1995. How Maps Work, Representation, Visualization, and Design. The Guilford Press, New York, USA.

Magnusson, C., Rassmus-Gröhn, K., Tollmar, K., Deaner, E. (eds.), 2009a. User Study Guidelines (for the HaptiMap project), http://haptimap.org/documents.html

Magnusson, C., Rassmus-Gröhn, K., Tollmar, K. and Stigmar, H., 2009b. Exploring User Requirements for Non-Visual Mobile Navigation Systems. In: Gross, T. et al. (eds.), INTERACT 2009, Part I, LNCS 5726, pp. 754–757.

Morgan, D. L.,1997. Focus groups as qualitative research, Sage Publications (CA).

NASA, 2010. NASA TLX, http://humansystems.arc.nasa.gov/groups/TLX/

Nielsen, C. M., Overgaard, M., Bach Pedersen, M., Stage, J., Stenild, S., 2006. It's worth the hassle!: the added value of evaluating the usability of mobile systems in the field. In Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles.

Nielsen J., 1993. Usability Engineering, Morgan-Kaufmann.

Nivala, A.-M., Sarjakoski, L. T, Jakobsson, A. and Kaasinen, E., 2003. Usability Evaluation of Topographic Maps in Mobile Devices. In: Proceedings of the 21st International Cartographic Conference, Cartographic Renaissance, pp. 1903-1913.

Nivala, A.-M., Sarjakoski, L. T. and Sarjakoski, T., 2005. User-centered Design and Development of a Mobile Map Service. In: Hauska, H. and Tveite, H. (eds.), ScanGIS'2005 – Proceedings of the 10th Scandinavian Research Conference on Geographical Information Science, pp. 109-123.

Nivala, A.-M., Sarjakoski, L. T., Laakso, K., Itäranta, J., and Kettunen, P., 2009. User Requirements for Location-Based Services to Support Hiking Activities. In: Gartner, G. and Rehrl, K. (eds.), Location Based Services and TeleCartography II, From Sensor Fusion to Context Models, Lecture Notes in Geoinformation and Cartography, Springer-Verlag Berlin Heidelberg, pp. 167–184.

Porathe, T., 2006. 3-D Nautical Charts and Safe Navigation. Doctoral Dissertation, Mälardalen University.

Sayda, F., 2005. Involving LBS users in data acquisition and update. In: Toppen, F. and Painho, M. (eds), Proceedings of AGILE 2005, 8th Conference on Geographic Information Science, pp. 33-42.

Schilling, A., Coors, V. and Laakso, K., 2005. Dynamic 3D Maps for Mobile Tourism Applications. In: Meng, L., Zipf, A., Reichenbacher, T. (eds.), Map-based Mobile Services, Springer, pp. 227-239.

Schneiderman, B., 1992. Designing the User Interface, (2nd ed.) Addison-Wesley.

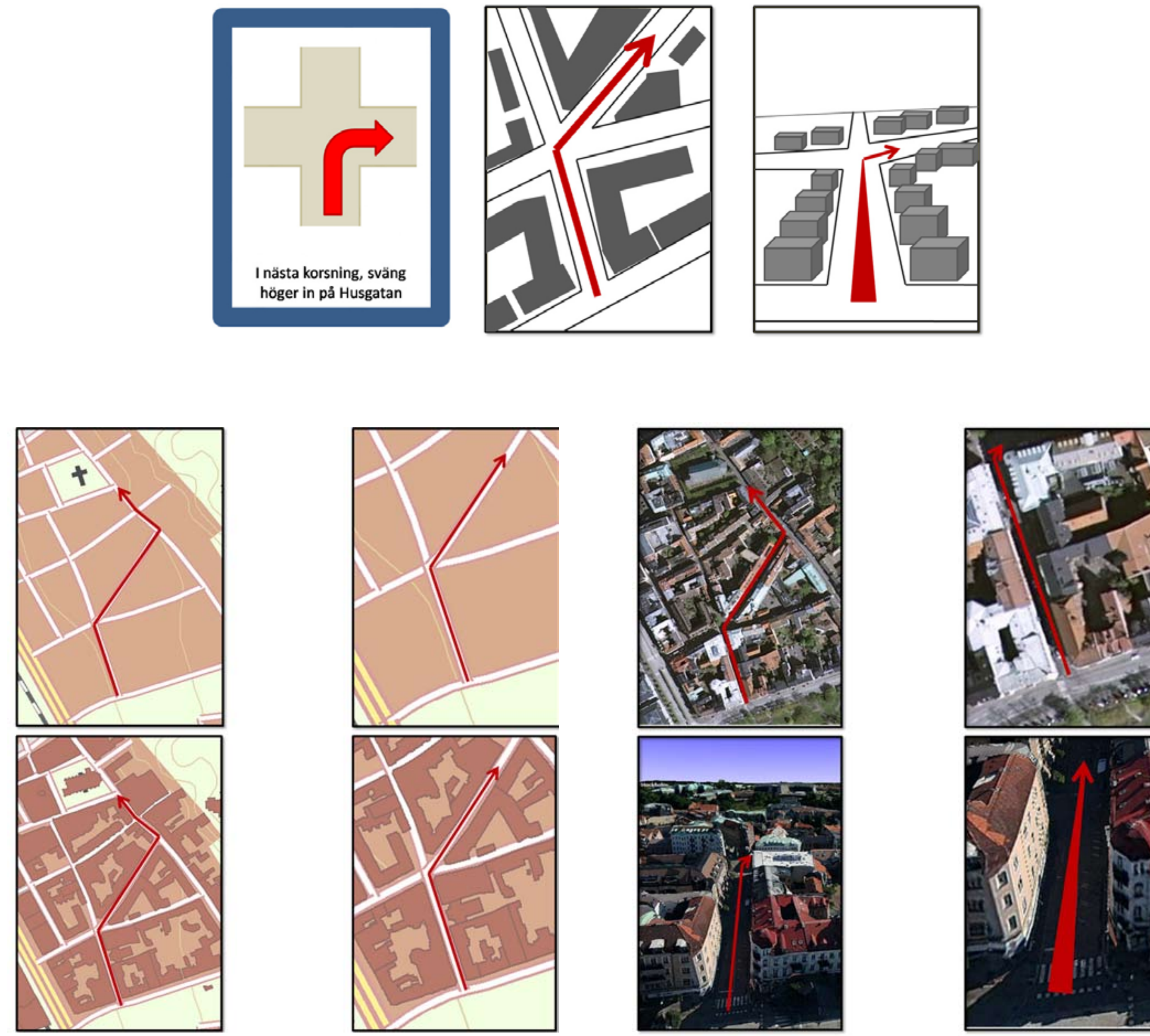
Sohn, T. Li, K. A., Hollan, J. and Griswold, W. G., 2008. A diary study of mobile information needs. In: Proceedings of CHI 2008, Florence, Italy.

Vision Australia, 2009. Vision Australia, blindness and low vision services, available at: http://www.visionaustralia.org.au/.

Visser, P.S., Krosnick, J.A., & Lavrakas, P.J., 2000. Survey research. In: Reis, H.T. and Judd, C.M., Handbook of research methods in social and personality psychology. Cambridge University Press.

Appendix A

The 11 visualizations from the survey questionnaire study.



Identifying areas of a map that are difficult to read

Perola Olsson[1], Karsten Pippig[1,2], Lars Harrie[1] and Hanna Stigmar[1]

[1]GIS Centre, Lund University, Sölvegatan 12, SE-223 62 Lund, Sweden
[2]Institute of Cartography, Technische Universität Dresden, DE-01062 Dresden, Germany
perola_olsson@hotmail.com, karsten.pippig@tu-dresden.de, lars.harrie@nateko.lu.se, hanna.stigmar@lantm.lth.se

Abstract:

The aim of this paper is to develop and evaluate methods to analytically identify areas of a map that are difficult to read. The approach is to compare such areas that were found in user tests with two analytical methods: the threshold method and the cluster method. These methods are implemented and evaluation shows that both methods have the potential of identifying areas that are difficult to read. A main advantage with the cluster method is that it is computationally more efficient than the threshold method, which makes it the best candidate for integration in a real-time service. The cluster method, or similar, could in the future be used as a guidance in the generalization process to improve map legibility in e.g. web services.

Keywords: *cartography, map legibility, web services, map reading, expert user tests, clustering.*

Introduction

The usage of maps has changed profoundly during the last decade. Today, a large portion of the maps are screen maps originating from web services. We could anticipate that this change in map use will continue in the future, especially since most countries are building up national spatial data infrastructures, which are partly based on geoportals. A change that most likely will occur is that more web based map services will distribute geospatial data rather than predefined maps. This development will provide possibilities to integrate geospatial data from several sources. However, creation of maps using geospatial data from several sources in real-time introduces new challenges; one such challenge concerns the legibility of the maps.

In automated cartography legibility is mainly studied in a bottom-up approach. By legibility constraints a program identifies features, or group of features, that are regarded as non-legible (e.g. overlapping or too close features). Then generalization algorithms are applied to resolve these non-legibility problems. The aim of this study is somewhat different. Instead of studying isolated violations of the legibility rules (constraints) we aim at identifying areas in a map that a map reader perceives as difficult to read. The rationality behind this approach is that a map reader may accept isolated violations of the legibility rules, at least in a real-time map, but he/she cannot accept areas of the map that are not readable. Therefore, it would be interesting to establish analytical methods for identifying the areas that are hard to read. The real-time generalization should then mainly focus on resolving conflicts in these areas. This approach inevitably leads to maps where generalization is performed solely on areas that are identified as hard to read; in other words, the type and level of generalization applied is not the same for all parts of the map. This situation may be undesired; however, in real-time generalization of maps viewed in a geoportal time efficiency and readability can be considered as more important than an evenly generalized map.

Related work

Identifying areas that are difficult to read in a map is linked to the presence of clutter. It has been shown that clutter has a negative effect on the performance and likeability of visual presentations (Phillips and Noyes, 1982). In cartography the removal of clutter is performed by generalization. However, one problem in this aspect is to know when a map is too cluttered and in need of generalization. To perform this we need measures of clutter and also an evaluation strategy of the effect of generalisation. One example of the latter is given by Jansen and van Kreveld (1998). A grid is placed on a map and a clutter function is applied to quantify the amount of clutter in each grid cell. The evaluation if performed by comparing the amount of clutter for each cell before and after the generalisation.

There have been several studies for measures of clutter and quantifying the information in maps. The measures can be categorised into the following classes:

- information amount, e.g. the number of objects (Phillips and Noyes, 1982; Wolfe, 1994; Oliva et al., 2004), the number of objects of a particular type (Töpfer and Pillewizer, 1966).
- spatial distribution of information, e.g. distribution of objects (MacEachren, 1982), their symmetry and organization (Oliva et al., 2004), entropy measures for objects and points (Bjørke, 1996; Li and Huang, 2002).
- complexity of information, e.g. sinuosity (João, 1998), total angularity (McMaster, 1987), and line connectivity (Mackaness and Mackechnie, 1999; Fairbairn, 2006).
- symbology, e.g. different aspects of colours (e.g. contrast) of the visualized objects (Eley, 1987; Oliva et al., 2004), legibility of graphics (Robinson et al., 1995; Spiess 1995).

Some researchers have proposed that a single measure never can explain if an area in a map is hard to read (i.e., is cluttered) and that you have to consider synthesis of measures (Rosenholtz et al, 2005, 2007; Schnur et al., 2010; Stigmar and Harrie, 2011; Stigmar et al., 2011). However, to the authors knowledge there have not yet been any studies that are using measures to explicitly find areas in the map that are difficult to read.

Methodology

First we prepared map data (section 3.1). These map data were used to perform user studies to identify which areas in the map that are perceived as hard to read by a user (section 4). Then we developed, implemented, and tested two analytical methods for identifying these areas: the threshold method (section 5) and the cluster method (section 6). Finally, we performed an evaluation of these analytical methods with respect to the outcome of the user studies (section 7).

For the user test the map was prepared using ArcGIS from ESRI. The threshold and cluster methods were implemented in Java. JTS Topology Suite (JTS) (JTS, 2011) was used for geometry operations and OpenJump (OpenJump, 2011) for visualization.

Map data

In the user test a map in the scale of 1:50 000 was used. The map was created from a geographic database from the Swedish mapping, cadastral, and land registration authority (Lantmäteriet) and from the local municipality of Helsingborg (Helsingborgs kommun).

In the case studies of the threshold and cluster methods contour lines and land cover were not included. The reason for this is that these feature types are considered as belonging to the background in the visual hierarch and should be symbolized with less distinct symbols, such as pale colours and/or thin lines. Hence, they will influence map legibility less compared to objects that are placed higher in the visual hierarchy. For example, in the urban part of the map contours are barely visible.

User Tests

Twelve test persons were participating in the test. Seven of the participants were male and five female, with an average age of 40 years. Six of the participants worked on social planning; one on detailed planning, two on regional planning, and three on both. The other six participants worked with GIS at university level. The participants' occupational experience ranged from one year to 30 years, with an average of 9 years. They worked with maps for an average of 50% of their total working time.

Test Procedure

The expert test was performed individually at the participants' place of work in order to reflect their everyday working situation. The test consisted of several parts, but here we only describe the part that is of interest for this particular study. In this part the tests were performed as user evaluations where the participants were asked to give their opinions on different aspects of the map, and to mark specific problem areas. The participants were asked to do this both in a general perspective as well as with different legibility problems in mind. These legibility problems were based on the legibility measures described in Stigmar and Harrie (2011).

Result

Figure 1 shows the areas in the map that were regarded as difficult to read by at least two test persons.

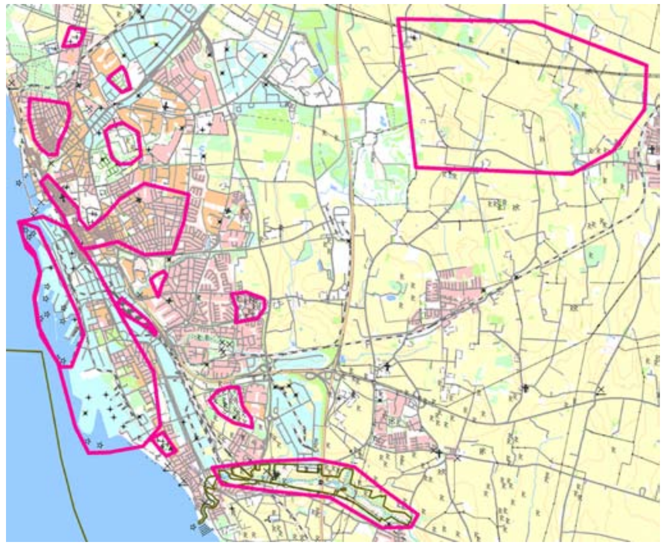


Figure 1. Areas of the map that are hard to read according to the user test. Copyright: Lantmäteriet and Helsingborgs kommun.

Threshold method

Method

The threshold method is an approach to detect areas that are hard to read by using grid cells. For this, a grid is applied over the entire map where each grid cell represents a specific map region and is linked to the map's geometry. This makes it possible to carry out information measurements based on the geometry within each grid cell and to determine the information content quantitatively. A previous study on analytical estimation of map legibility by Stigmar et al. (2011) compares several information measures and provides threshold values for each measurement. The thresholds are derived from empirical tests (different to the test described in section 4) and can be seen as the critical points when features become illegible in relation to a specific information measure. In this study six information measures, taken from Stigmar et al. (2011), are considered and implemented in the method. For each grid cell the value of the measures are computed; if the value for any of the information measures exceeds a respective threshold value that grid cell is regarded as hard to read. The following measures are used (threshold values in bracket):

- Number of objects (<11/cm²)
- Object line length (<17cm)
- Number of vertices (<450/cm²)
- Number of object types (<17)
- Degree of overlap for disjoint objects (<3)
- Angularity (<40/cm)

The *Degree of overlap for disjoint objects* (DO) is defined as the sum of intersections between disjointed features. To ensure good legibility, an outline of 0.3 mm is adopted to the features. For this reason, a buffer is taken into account for each feature. The buffer size is based on the symbology size and a requirement of the minimum separation of 0.3 mm. The Degree of overlap is described as

where n is the number of objects, and δ_{ij} is equal to 1 if object i and j are disjoint and otherwise zero.

The angularity is defined as the sum of all the changes in direction of a line divided by its total length.

Results

As described in section 5.1, most of the measured values are relative to an area unit. Therefore, the choice of cell size is crucial. In this study, a cell size of 1 cm² is used. This means illegible areas are roughly represented (Figures 2 and 3). Smaller cell sizes would state illegible areas with a higher accuracy, but leading to a higher computation time.

The algorithm is designed to determine the individual measures sequentially for each grid cell. To reach high efficiency the method starts by evaluating measures that only requires simple geometry computations (e.g. number of objects). Measures that require complex topological computations and, hence, a high running time, such as degree of overlap or angularity, are tested later. This approach is efficient in areas with high density; once a grid cell is identified as difficult to read no further measures need to be evaluated for that cell. In areas with low density the algorithm is efficient since the computations are faster due to less detail. However, for areas with medium density efficiency decreases since a high number of measures must be evaluated and these cells are richer in detail than the low density areas. To minimize object density, only data layers that present features in the fore- or middle ground were used. Figure 2 provides the result from the threshold method using all data layers except land cover and contour lines. To shorten the computations a test was performed where also road data were left out. The result of the latter test is presented in Figure 3.

Figure 2. Areas of the map that are hard to read as identified by the threshold method - all data types except land cover and contour lines are included.

Figure 3. Areas of the map that are hard to read as identified by the threshold method - all data types except roads, land cover, and contour lines are included.

Discussion

It seems to be possible to imitate the results of the performed user tests by an automated determination of legibility based on the threshold method. The results are affected by the grid cell size, the location of each grid cell, the specific threshold values, and the data layers.

In the tests different object types have been used in the computations. When the contour lines and land cover data are included large areas of the map are identified as hard to read. The reason is that contour lines and land cover polygons are made up of a large number of points that increase the number of vertices within a grid cell. However, these points do not make the map hard to read, but they are necessary to make the lines smooth. They just affect the map reading slightly and both object types are located in the background in the printed map. It seems appropriate to exclude these data layers from an automatic analysis. At least, the measure values should be weighted according to if a feature belongs to fore-, middle, or background.

Currently the grid tiles are joined together seamlessly. Depending on the location of the grid tiles, information can be lost when analyzing the legibility since the calculations refer just to each grid cell. If an area with high density of information is distributed over several tiles, thresholds may not be exceeded. A solution could be an overlap of tiles, for example 50%. Thus, every area of the map would be distributed over several tiles and areas with poor legibility would be detected. This would, however, be associated with a higher computational time.

As stated in section 5.2, the use of smaller cell sizes would identify the shapes of the illegible areas with a higher accuracy. However, due to the high computation time, we have worked with a lower grid resolution. This leads to a rough representation of the areas that are hard to read. However, since the purpose is to find the illegible areas rather than to identify their shape accurately this is acceptable.

Compared to the result of the user test much larger areas are detected as illegible by the threshold method. Currently, each measure does classify a grid cell as hard to read on its own, if this measure exceeds the threshold value. Most probably the measures affect and imply each other. In a future development of the method, constraints can be established to identify cells in the context of different measures. Some measures may classify a cell as illegible alone, while other measures only make a grid cell illegible if the threshold is exceeded also by other measure(s).

Cluster method

Method

The basic idea of the cluster method is to find dense clusters of map points. Map points are defined as the midpoint of point objects and as vertices on line and polygon objects. Clusters are then identified by applying the density based algorithm DBSCAN (Ester et al., 1996) on the map points. DBSCAN is able to find clusters of all shapes, and points that do not belong to any cluster are regarded as noise - that means that all points are not included in a cluster. In our application we are interested in finding areas where the map points are dense and form clusters there. Then the convex hulls of these clusters are regarded as areas that are difficult to read.

DBSCAN identifies an area as dense if there is a minimum number of points minPts within the distance ε from a point. If there are more than minPts points within the distance ε from a point, that point is regarded as a core point. Another point that is within the distance ε from a core point, but is not a core point itself, is a border point. Points that are not within the distance ε of any core point are regarded as noise (see Figure 4). For details about the algorithm see Ester et al. (1996) or Han et al. (2001).

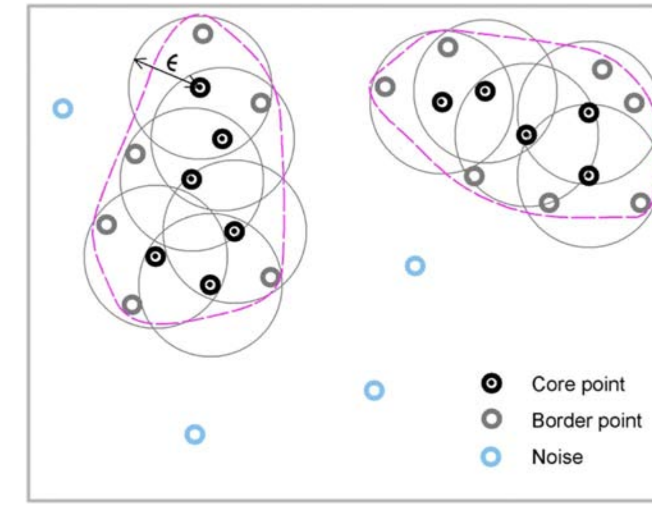


Figure 4. Two clusters identified by DBSCAN. minPts = 3.

To reach a higher correlation between the areas that are identified as difficult to read by DBSCAN and the areas perceived as hard to read in the user test, modifications of the basic idea were made.

- When a single map point is used to represent a point object, the impact of point objects on legibility is underestimated compared to the user test (Figure 1). To improve this situation a number of map points were added for each point object.
- The influence of point objects can also be increased indirectly by decreasing the impact of line and polygon objects. Especially smooth curves on lines and polygons tend to be overestimated in terms of illegibility. These gentle curves are formed by several short line segments that result in a large number of map points being created. These map points may result in clusters that are not due to poor legibility. To decrease the impact of line and polygon features the number of map points was reduced with the simplification algorithm Douglas Peucker (Douglas and Peucker, 1973). A threshold was defined for the ratio between line length and the number of vertices of that line to decide on which lines and polygons to apply Douglas Peucker.
- Areas with parallel straight lines, such as railway yards, are not identified as difficult to read by DBSCAN when map points are created at vertices only. Such areas are potentially difficult to read; hence, a maximum distance for line segments maxDist was defined. For line segments with a length exceeding maxDist, map points were added.
- A threshold for minimum area of the convex hulls representing the clusters was applied to omit small clusters.

Result

Figures 5–6 show the areas identified by the cluster method. As shown, which areas that are identified as difficult to read is strongly affected by the parameters *minPts* and ε. In the figures the maximum length for line segments is 200 m (if longer, additional map points are added).

Large ε results in less, but larger clusters. The nature of convex hull prevents the shapes of the clusters from being clearly shown. This results in areas that are not part of a cluster, being covered by the convex hull; hence, large clusters are occasionally overlapping other clusters. With decreasing ε as in Figure 6 the number of clusters is increasing and their size decreasing, and the shapes of the convex hulls conform more to areas with poor legibility.

minPts=7 and ε=115m

Figure 5. Clusters created with minPts=7 and ε=115m. Max length of line segments is 200 m

minPts=8 and ε=75m

Figure 6. Clusters created with minPts=8 and ε=75m. Max length of line segments is 200 m.

Figure 7 and 8 shows areas that were identified as difficult to read by DBSCAN with settings as in Table 1. In Figure 7 marginally more areas that were perceived as difficult to read in the user test are identified by DBSCAN compared to Figure 8. However, Figure 7 also includes more areas that were

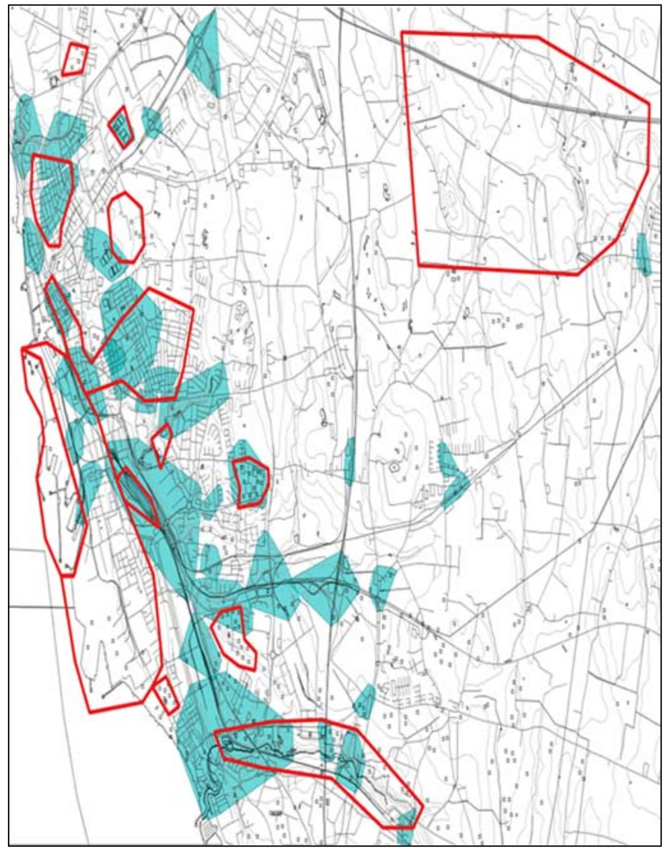


Figure 7. Clusters created with parameter values as in Table 1.

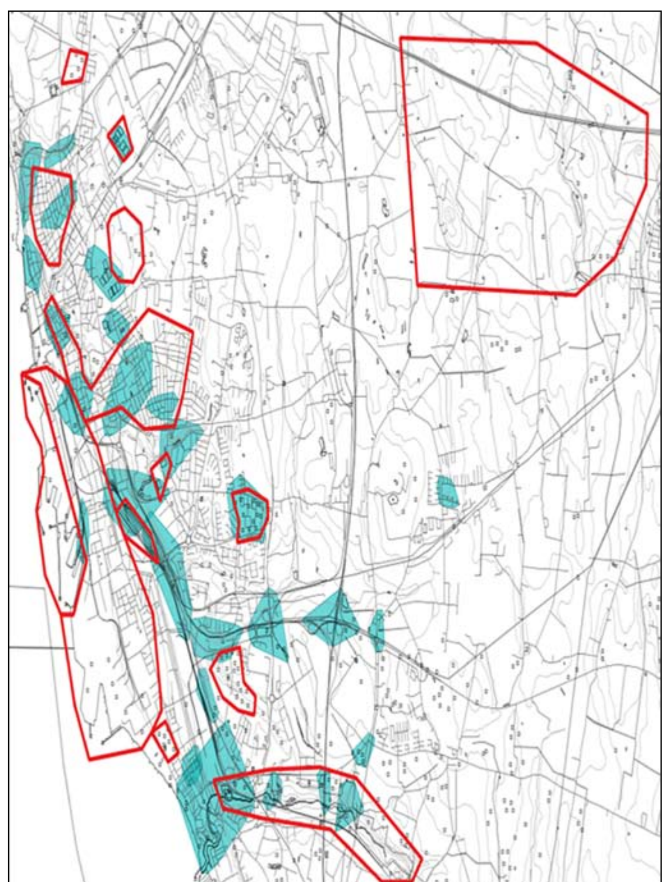


Figure 8. Clusters created with parameter values as in Table 1.

not perceived as difficult to read.

To reduce the number of clusters a threshold was applied to remove small areas. 250,000 m^2 was tested as it corresponds to 1 cm^2 in scale 1:50,000. This is the unit for the thresholds defined in Stigmar and Harrie (2011) and utilized in the threshold method. However, this size of the minimum area excluded some of the areas that were identified as hard to read in the user study. Hence, a minimum area of 100,000 m^2 was applied.

	Figure 7	Figure 8
minPts (DBSCAN)	**15**	**12**
ε (DBSCAN)	**75 m**	**75 m**
Additional map points per point object	**8**	**16**
Ratio line length/no. of vertices (if feature should be simplified)	**50 m**	**50 m**
tolerance (Douglas Peucker)	**10 m**	**12 m**
maxDist (line segments)	**50 m**	**100 m**
Minimum area (convex hull of cluster)	**100,000 m^2**	**100,000 m^2**

Table 1. Parameter values for the cluster method to identify the clusters in Figure 7 and 8.

Discussion

The main concern in Figure 7 and 8 is that line objects, mainly gently bending roads, are overestimated in terms of poor legibility; several areas that were not perceived as hard to read in the user test are identified as difficult to read by the cluster method. This situation can be improved by adjusting the parameters listed in Table 1. It is, however, so that different parameters give similar results when modified leading to a complex study.

Two important parameters are the parameters required by the DBSCAN algorithm, namely minPts and ε. Figure 5 and 6 clearly shows how these parameters affect the resulting clusters. However, the tests performed showed that there is no major difference between a low value for minPts and a high value for ε, versus a high value for minPts and a low value for ε.

Another concern is that the influence that point objects have on legibility is underestimated compared to the user test. The approach to add additional map points for each point object does improve the result. However, the additional computations decrease performance. An approach where the number of map points that are added for a point object is related to the symbology would be interesting to test; for simple symbols only a few map points could be added, and for complex symbols several map points. That would, however, increase the complexity of the method.

The parameters that decide when the Douglas Peucker algorithm is applied and how much the algorithm should simplify an object were also adjusted to improve the result. If a line or polygon is to be simplified is decided by the ratio between line length and the number of vertices. The tests showed that the ratio should be low to have a major effect on the result. The parameter tolerance (max perpendicular distance) required by the Douglas Peucker algorithm also has a major effect on the result. Naturally a high tolerance reduces the number of map points resulting in less clusters being created. However, the level of simplification must be considered so that line and polygon features are not simplified too much.

The last parameter affecting the clusters being created by cluster method is the maximum length of line segments, maxDist. For this parameter there is a conflicting aspect between a low value and a high value. A low value results in a large amount of map points being added and large clusters are created. A high value on the other hand leads to areas that are difficult to read, according to the user test, not being identified.

Finally a threshold was applied on the clusters that were created by the cluster method. Minimum area for clusters naturally removes the smallest clusters; these are likely not a major problem from a legibility perspective since they are small. However, the test showed that this threshold should be kept low, otherwise the smallest areas identified as hard to read in the user test will be omitted.

In an extended future study it is likely that the result can be improved by:

- increasing the maxDist between map points along line segments. This would both decrease the impact of line and polygon objects and reduce the number of computations.
- increasing the tolerance for Douglas Peucker to decrease the impact of lines and polygons.
- adding more additional map points for each point object; this would, however, increase the number of computations.

This should be done in combination with different values for the parameters minPts and/or ε as required by DBSCAN

Another concern is that the convex hull is not representing the shape of large clusters (Figure 5 and 6) very well. This situation could be improved by implementing a better aggregation method of points, such as the method proposed by Joubran and Gabay (2000)

Evaluation

When the threshold and cluster methods are evaluated it should be noted that all the data used in the user test (Figure 1) are not included when the methods are tested. This means that all areas perceived as difficult to read in the user test cannot be identified by the threshold and cluster methods; these areas, which are described below, are not considered in the evaluation.

The upper right part – this large area was perceived as difficult to read due to the contours, which are not considered in the tests of the methods.

The lowermost area (elongated polygon in East-West direction) – this area is likely to be perceived as difficult to read due to a nature protection area (distinct green line with complex shape) and contours along the gully. Since the green line is a symbology problem, and contours are not included in the test of the methods, this area cannot be identified by either the threshold or the cluster method.

The harbour area (West) – in the dataset the coastline is represented by water in the land cover data. Since land cover is not considered in this case study the complex shape of the wharf is not found. Hence, this area cannot be found by either the threshold or cluster method.

In Figure 9 and 10 the areas identified as hard to read in the user test are symbolized with a thick line (red), areas identified by the threshold method are bright (green), and areas identified by the cluster method are dark (green). In Figure 9 all data layers except contours and land cover are considered and settings for the cluster method are as in Table 1, Figure 7. The threshold method identifies all areas from the user test except a small part of one area in the North-East part. However, large areas that were not perceived as hard to read in the user test are included. The cluster method fails to identify larger portions of the areas that were perceived as difficult to read in the user test, and three small are totally missed. However, substantially fewer areas that were not perceived as difficult to read in the user test are identified as difficult to read by the cluster method.

Figure 9. Comparison of areas of the map that are hard to read identified by (light green) the threshold method and by (dark green) the cluster method (settings as in Table 1 – Figure 7). All data except contour lines and land cover are used.

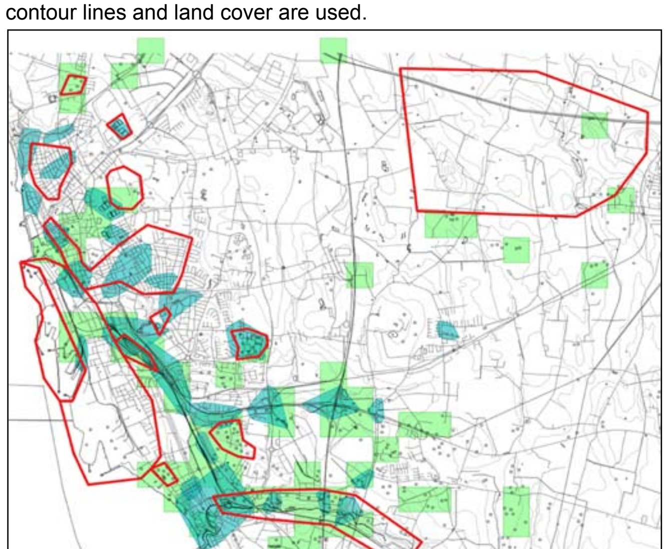


Figure 10. Comparison of areas of the map that are hard to read identified by (light green) the threshold method and by (dark green) the cluster method (settings as in Table 1 – Figure 8). All data except contour lines and land cover are used – for the threshold method also roads are excluded.

In Figure 10 contour lines and land cover are excluded from the test. For the threshold method also roads are excluded. For the cluster method settings are as in Table 1, Figure 8. In this test the threshold method identifies fewer areas, but it also fails to identify several areas that are hard to read according to the user test. The cluster method also identifies fewer areas compared to Figure 9; however, nearly the same result as for the test in Figure 9 is reached for areas that are hard to read according to the user test.

In Figure 10 the threshold method mainly fails to identify areas with several roads which is natural since roads are not considered in the test. The major concern for the cluster method is the point objects as discussed earlier. This situation is likely possible to improve by adding a higher number of map points for point objects. By adjusting the parameters minPts and ε for DBSCAN more areas would be identified as difficult to read. However, identifying more areas will inevitably include more areas that are not perceived as hard to read. These conflicting aspects must be considered if the method is further developed.

Discussion

The main aim of the threshold and cluster methods is to identify areas in a map that are difficult to read as part of the generalization process. The outcome of the methods should guide the generalization process by identifying which areas that most likely requires generalization. Figure 11 shows an example of how the threshold and/or cluster methods can be used in a legibility service. The service is utilized to improve legibility when data from several web services are viewed in a geoportal. A user connects to a geoportal via the Internet and requests a map that consists of geographic data from several external services. Since both the threshold and cluster methods require the geographic data, these external services must be download services, such as Web Feature Service (WFS) that enables a user to download the data. The geoportal retrieves the data requested and sends them to the legibility service. The legibility service identifies areas that are difficult to read, and applies generalization operations to enhance legibility. Finally a map is created from the generalized data and returned to the user.

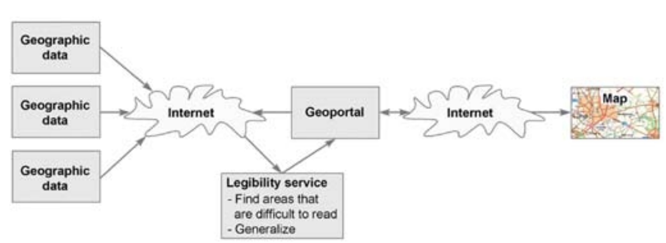


Figure 11. A possible future implementation of the threshold and/or cluster methods as a legibility service to improve legibility when a map is viewed in a geoportal.

How a legibility service identifies areas that are difficult to read in an efficient manner can be discussed. A major difference between the threshold method and the cluster method is that the threshold method considers several measures that are based on empirical studies. This implies that it is likely to identify areas that are hard to read with a higher accuracy. However, the cluster method by far outperforms the threshold method in time efficiency. A possible approach would be to apply the cluster method on the original data to find a candidate set of areas with poor legibility. The threshold method could then be applied to perform a refined search on the candidate set. Finally, generalization operations are applied on the areas identified as difficult to read. However, performance must be considered. It might be so that it is more efficient to apply generalization directly on the areas identified by the cluster method, than first refining the search with the threshold method.

As a future extension it might also be possible to extend the legibility service to play the role of a virtual cartographer. A possible workflow would be:

1. Find areas with poor legibility.
2. Apply appropriate generalization operations on these areas (if any).
3. Apply symbolization methods to improve cartography.
4. Apply text placement methods.

Conclusions

Based on the studies in this paper we can draw the following conclusions:

* Both the threshold method and the cluster method are able to identify most of the areas in the map that are difficult to read (as stated in the user test) where the difficulties stem from cluttering. There are also other reasons for an area to be perceived as difficult to read, such as the symbology used. Areas that are perceived as hard to read due to these reasons are not identified by the methods tested.
* The threshold method is advantageous in the sense that it is built on empirical studies of map reading. This in contrast to the cluster method, which is a method selected by intuition that it should provide a good result. On the other hand, the cluster method is much more efficient and could be utilised in a real-time process. For that reason it is interesting that the cluster method almost provides the same result as the threshold method (cf. Figure 9 and 10). We believe that with some effort in tailoring the cluster method it could be a proper alternative to the more empirically solid threshold method.

Acknowledgements

We would like to thank all the participants in the user tests. Lantmäteriet and the city of Helsingborg are acknowledged for providing geographic data.

References

Bjørke, J. T. (1996) 'Framework for Entropy-based Map Evaluation', **Cartography and Geographic Information Systems**, 23, pp. 78-95.

Douglas, D. H. and Peucker, T. K., 1973. Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature. **The Canadian Cartographer**, Vo. 10, No. 2, p 112-122.

Eley, M.G. (1987). 'Color-layering and the performance of the topographic map user', **Ergonomics**, 30, pp. 655-63.

Ester, M., Kriegel, H.-P., Sander, J. and Xu, X., (1996). A density-based algorithm for discovering clusters in large spatial databases. Proceedings of the 1996 International **Conference on Knowledge Discovery and Data Mining** (KDD –96), August, 226-31.

Fairbairn, D. (2006) 'Measuring Map Complexity', **The Cartographic Journal**, 43, pp. 224-238.

Hangouet, J. F. (1998) 'Voronoi Diagrams on Segments – Properties and Tractability for Generalization Purposes', Technical report for AGENT, Cogit, IGN.

Han, J., Kamber, M. and Tung, A.K.H., (2001). Spatial clustering methods in data mining - A survey. In Miller H.J. and Han, J., 2001. Geographic Data Mining and Knowledge Discovery. Taylor and Francis. ISBN 0-415-23369-0.

Jansen, M. and van Kreveld, M. (1998). Evaluating the consistency of cartographic generalization. In: T.K. ¨ Poiker and N. Chrisman (eds.), **Proc. 8th Int. Symp. on Spatial Data Handling**, pp. 668--678.

João, E. M. (1998) **Causes and Consequences of Map Generalisation**, London: Taylor & Francis.

Joubran, J., and Y. Gabay (2000). A Method for Construction of 2D Hull for Generalized Cartographic Representation. **International Archives of Photogrammetry and Remote Sensing**, Vol. XXX-III, Part B4, Amsterdam, pp. 417-424.

JTS, (2011). JTS Topology Suite, Vivid Solutions. http://www.vividsolutions.com/jts/jtshome.htm.

Li, Z. and Huang, P. (2002) 'Quantitative measures for spatial information of maps', **International Journal of Geographical Information Science**, 16, pp. 699-709.

MacEachren, A. M. (1982) 'Map Complexity: Comparison and Measurement', **The American Cartographer**, 9, pp. 31-46.

Mackaness, W. A. and Mackechnie, G. A. (1999) 'Automating the Detection and Simplification of Junctions in Road Networks'. **GeoInformatica**, 3, pp. 185-200.

McMaster, R. B. (1987) 'Automated Line Generalisation', **Cartographica**. 24, pp. 74-111.

Oliva, A., Mack, M. L., Shrestha, M. and Peeper, A. (2004) 'Identifying the Perceptual Dimensions of Visual Complexity of Scenes', in **Proceedings of the 26th Annual Meeting of the Cognitive Science Society**.

OpenJump, 2011. http://www.openjump.org/.

Phillips, R. J. and Noyes, L. (1982) 'An Investigation of Visual Clutter in the Topographic Base of a Geological Map', **The Cartographic Journal**, 19, pp. 122-132.

Robinson, A. H., Morrison, J. L., Muehrcke, P. C., and Guptil, S. C., (1995). **Elements of cartography – sixth edition**. John Wiley and sons, Inc., New York, USA.

Rosenholtz, R., Li, Y., Mansfield, J. and Jin, Z. (2005) 'Feature Congestion: A Measure of Display Clutter', in **Proceedings of SIGCHI Conference on Human Factors in Computing Systems**, pp. 761-770.

Rosenholtz, R., Li, Y. and Nakano, L. (2007) 'Measuring visual clutter', **Journal of Vision**, 7, Art. 17, pp. 1-22.

Schnur, S., Bektaş, K., Salahi, M., Çöltekin, A., 2010. 'A Comparison of Measured and Perceived Visual Complexity for Dynamic Web Maps', in **Proceedings of GIScience 2010**, Zurich, 14-17 September, 2010.

Speiss, E. (1995) 'The need for generalization in a GIS environment', in **GIS and Generalization – Methodology and Practice**, ed. Müller, J.-C., Lagrange, J.-P. and Weibel, R., Taylor and Francis

Stigmar, H., and L. Harrie, (2011). Evaluation of Analytical Methods to Describe Map Readability. **The Cartographic Journal**, in print.

Stigmar, H., Harrie, L. & Djordjevic, M., (2011). **Analytical estimation of map legibility**, submitted.

Töpfer, F. and Pillewizer, W. (1966) 'The Principles of Selection', **The Cartographic Journal**, 3, pp. 10-16.

Wolfe, J. M. (1994) 'Guided Search 2.0: A revised model of visual Search', **Psychonomic Bulletin & Review**, 1, pp. 202-238.

MAKING SWEDISH ENVIRONMENTAL GEODATA INSPIRE COMPLIANT: A HARMONIZATION CASE STUDY

Finn Hedefalk and Anders Östman

Gävle GIS Institute, University of Gävle, Nobelvägen 2, 80167 Gävle,
finnhedefalk@gmail.com; Anders.Ostman@hig.se

Abstract

The European project Nature-SDIplus has developed data and metadata specifications for three INSPIRE Annex III themes: Habitat and Biotopes, Bio-geographical regions and Species distributions. These serves as a foundation for the thematic groups developing the corresponding INSPIRE specifications. The aim of this study is to test a data harmonization approach to make Swedish environmental geodata and metadata compliant with these specifications. In the harmonization process, we use offline transformations that are split into one spatial and one non-spatial part, and standardized formats to allow vendor neutrality. Moreover, we extend the compliance tests to the data and metadata specifications by validating against both eXtensible Markup Language (XML)-schema and Schematron. Finally, we identify harmonization processes that may be costly or have negative impacts on data quality. The harmonized data and metadata are thereafter published as network services compliant with OGC Web Service specifications. The output from our method is data and metadata that are valid to the Nature-SDIplus data specifications and metadata profiles. Although the usage of standardized formats facilitates vendor neutrality, the non-spatial transformation procedures expressed in interoperable languages seem to be insufficient to execute all the mapping rules. Therefore, some of these transformations cannot be executed in a vendor neutral environment without modifications. Furthermore, by splitting the harmonization into two manageable parts, we avoid some limitations about XML schema translations in existing spatial transformation tools. Additional findings are: (1) by extending the validation with Schematron tests, we find non-compliances that have been missed during the XML schema tests; (2) costly processes are identified, which are caused by missing elements and by unstructured information given as comments; and (3) degradation of the positional and thematic accuracy occur during the harmonization.

Keywords: *INSPIRE, Geodata harmonization, ETL, OGC Web services, Nature-SDIplus, Validation, Data quality.*

INTRODUCTION

The Infrastructure for Spatial Information in the European Community (INSPIRE) is a directive that aims at establishing a Spatial Data Infrastructure (SDI) in Europe (EC, 2007). The main goal of an SDI is to provide easy access to harmonized geodata and metadata. These should be available by network services that allow users to discover, view and download the harmonized data. The INSPIRE Directive entered into force 2007 and is currently under implementation in the member states. Network services are being set up and the monitoring and reporting of the process has started. The harmonized datasets provided by the member states shall, according to the directive, comply with specific data specifications, and should be available in the period of 2012 – 2019 (INSPIRE, 2011).

The Swedish legal text of the directive got transited into national law January 1th, 2011 (SFS, 2010:1767). Accordingly, the information responsible authorities shall now ensure that their harmonized data and metadata will be available by network services. A problem that many of these authorities may face relates to data harmonization; that is, to make the data and metadata compliant with the data specifications. Should the harmonization procedures be executed offline or on-the-fly? What kind of operations should be executed and what tools should be used? Will there be costly transformations? What impacts on data quality will the transformation have? What kind of architecture should the network services be built on? And how to know if the data and metadata are compliant with the specifications? This paper aims to study these questions through a case study. The aim of the case study is to describe the data harmonization procedures and experiences, and to identify operations that are costly or reduce the data quality considerably.

A few recent studies exist about data harmonization for making data INSPIRE compliant. Within the ESDIN project, Hemmatnia, Broecke and Rammsdonk (2010) use a combined approach, which includes both preprocessing transformations through Extract, Transform and Load (ETL) and on-the-fly transformations. That is, they first transform the datasets and then store them in an intermediate relational database with schemas close to the INSPIRE models. When a request arrives to the download service, the final transformations are made on-the-fly and data are provided as INSPIRE compliant datasets in Geography Markup Language (GML) 3.2.1. By using this approach, the performance should be improved compared to using only on-the-fly transformations. Moreover, by using database schemas close to the INSPIRE models, the data themes are not treated as isolated datasets and links between the themes can be established and maintained. For example, features in the theme Geographical Names can be linked to features in both Addresses and Administrative Units. In a later study, Broecke et al. (2011) fully transformed all datasets and stored them as eXensible Markup Language (XML) data types in an XML-enabled database (one that accepts XML as input and provides some XML functions (Bourret, 2005)). By doing so, no demanding on-the-fly transformations or complex database schemas were needed, and the performance when providing data should be further enhanced. Moreover, by using XML-enabled databases, links between different themes could still be maintained by the use of XML Linking Language (XLink).

In another study, Gedrange, Neubert and Röhnert (2011) harmonized cross-border spatial data between Germany and the Czech Republic. Since the harmonization procedures for spatial and attribute data are often different, they split them into two separate parts. In addition, they created bilingual catalogues from the feature catalogues of the source datasets. The latter approach may be crucial when different datasets should be harmonized without having any data specification to comply with. As regards to Transformation Services that should be able to make data INSPIRE compliant (EC, 2009), Waters et al. (2010) show with their prototype that such transformations are possible in a vendor neutral environment. They use standardized formats such as GML and Rule Interchange Format (RIF) as integrators, which allow them to combine both proprietary and non-proprietary tools. Furthermore, the HUMBOLDT project has developed an open-source software framework for data harmonization and services integration. This framework aims to cover the whole harmonization process. They have also developed several scenario specific case studies. In some of these, data harmonization is done by integrating several of the developed tools and services in a workflow editor (Fichtinger et al., 2010). Other projects that have dealt with harmonization for INSPIRE are for instance EUROADIN, GIS4EU and MEDISOLAE-3D (Mendive, Cardoso and Cabello, 2010)

Few studies explain the validation phase of the harmonization, although Hemmatnia et al. (2010) and Beare, Payne and Sunderland (2010a) validate their data against the INSPIREs GML 3.2.1 Application schemas. Nevertheless, it may still be needed to also validate both the metadata and data according to the INSPIRE Schematron rules. Schematron is a rule-based validation language and ISO standard (ISO, 2006), which can define semantic rules that GML application schemas (i.e. XML schemas) cannot. Furthermore, it may be easy to comply to a data specification if degradation in the data quality is allowed. However, such degradation is not in line with the user needs, and it may produce useless data. This may cause additional costs, and therefore it is important to address these problems caused by the harmonization process.

The research in this paper was initially carried out in the context of the Nature-SDIplus project, which addresses among other things the data harmonization and validation issues. Nature-SDIplus aims at gathering good practices on data harmonization of European environmental datasets. With this, the project hopes to facilitate the member states' harmonization tasks. Moreover, Nature-SDIplus has developed data models for three of the INSPIRE Annex III themes, namely Bio-geographical regions (BGR), Species distributions (SD) and Habitat and Biotopes (HB). These will later serve as a foundation for the thematic groups developing the corresponding INSPIRE specifications. One of our objectives in Nature-SDIplus was to develop a method for harmonizing Swedish data and metadata to be compliant with the data specifications specified by the project.

More specifically, the primary objective of this paper is to test the combination of the following approaches as applied to Swedish data

- To harmonize all data offline and use a data warehouse approach (Yeung and Hall, 2007; Hemmatnia et al., 2011).
- To allow vendor neutrality as in Waters et al. (2010), Hemmatnia et al. (2011) and Fichtinger et al., (2010), by using standardized formats such as XML and GML.
- To split the process in one spatial and one non-spatial part (Gedrange et al., 2010).

Additional objectives are

- To test the compliance to the data and metadata specifications by extending the XML-Schema compliance tests with Schematron tests.
- To identify harmonization processes that may be costly or have negative impacts on data quality.

MATERIALS AND METHODS

The overall approach is described in Figure 1. Here, source data were extracted, transformed in one spatial and one non-spatial part, validated, loaded into new databases and finally published by standardized network services. In this chapter, we mainly describe the extract, transform and validation parts. The establishments of network services are shown in the results.

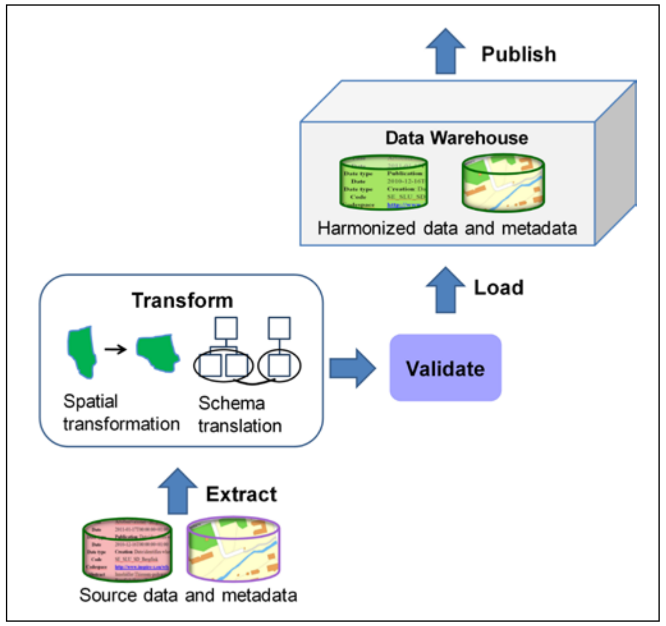


Figure 1. Overall harmonization approach

Materials

Our source data and metadata data were delivered from the Swedish Environmental Protection Agency (SEPA), and the Swedish University of Agricultural Sciences (SLU). As regards to the data from SLU, a web application was developed to extract them from the database. The test data contained information about valuable waters, wetlands, bio-geographical regions and species observations (Table 1). Their spatial extent covered the area of Gävleborg's county in Sweden.

Data provider	Source dataset	Source data format	Metadata	Target specification
SEPA	Valuable waters based on fishing	ESRI Shape + MsAccess	Yes	Habitat and Biotopes (HB)
SEPA	Special valuable waters based on fishing	ESRI Shape + MsAccess	Yes	Habitat and Biotopes (HB)
SEPA	Valuable waters based on nature	ESRI Shape + MsAccess	Yes	Habitat and Biotopes (HB)
SEPA	Special valuable waters based on nature	ESRI Shape + MsAccess	Yes	Habitat and Biotopes (HB)
SEPA	Special valuable waters based on culture	ESRI Shape + MsAccess	Yes	Habitat and Biotopes (HB)
SEPA	Wet lands	ESRI Shape + Excel	Yes	Habitat and Biotopes (HB)
SLU	Artportalen database (6 species observations)	RDBMS	No	Species distribution (SD)
SEPA	Biogeographical regions	ESRI Shape	Yes	Bio-geographical regions (BGR)

Table 1. Extracted source datasets. The target specification column describes the theme that the source data were transformed into. Six different species were extracted from the Artportalen database.

Data and metadata harmonization

The aim of the data harmonization is to solve the different heterogeneities that exist in data, such as syntactical (differences in formats, data types etc.), structural (differences in the data models) and semantic (differences in the meanings) (Yeung and Hall, 2007). To avoid a mix-up of concepts, we define harmonize as the process with the aim of eliminating conflicts among data so they can be combined in a meaningful way, and transform as a general operation for changing data. We split the data harmonization in two parts: One spatial transformation and one schema translation. One reason was the lack of spatial Extract, Transform and Load (ETL) tools that could successfully handle both the spatial and the XML schema translation parts. For example, Safe Software's FME Desktop 2010 had a lack of GML 3.2.1 support and Snowflake's GO Publisher could only handle databases as import format. Therefore we tested the spatial transformations with open-source GIS softwares and the schema translations with data mapping and conversion tools that had strong XML support. Note that we tested the harmonization tools created in the Humboldt project, for the reason of being open source and for their geospatial schema matching and mapping tool HALE (version 2.0.1). However, due to the early stages of development for many of the Humboldt tools, we failed to successfully use them in our work.

Spatial transformation

The spatial transformations (Figure 2) were carried out with the open source software GRASS GIS, which can provide functionalities similar to well-known proprietary GIS products (Steiniger and Bocher, 2009), and batch processing was used to carry out the transformation chains. The main spatial transformations for all datasets were coordinate transformations to the European Terrestrial Reference System 1989 (ETRS89), and format conversions to GML 2.1.2 simple features. The reason for the conversions was to facilitate the schema translation and to allow vendor neutrality.

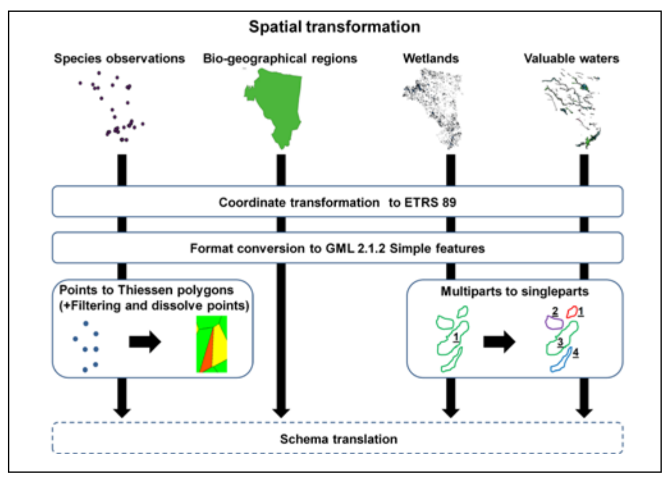


Figure 2. Spatial transformation

Schema translation

The goal of the schema translation was to make the data and metadata compliant with the Nature-SDIplus data specifications and metadata profiles. The schema-translation consisted of three parts

- Schema matching: Finding elements or attributes that have similar meaning.
- Schema mapping: Defining the schema translation operations.
- Schema transformation: Executing the schema translation operations.

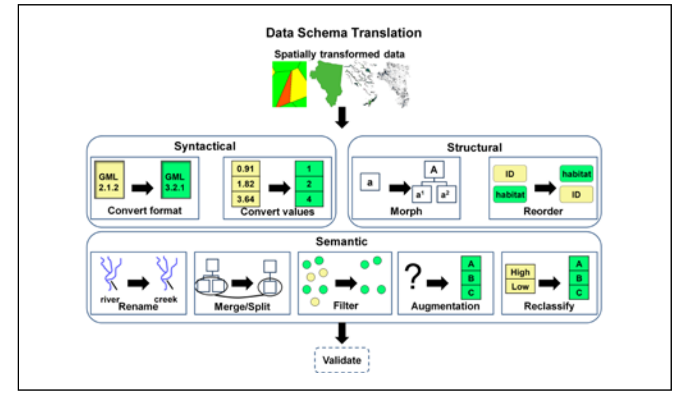


Figure 3. Schema translation operations of data

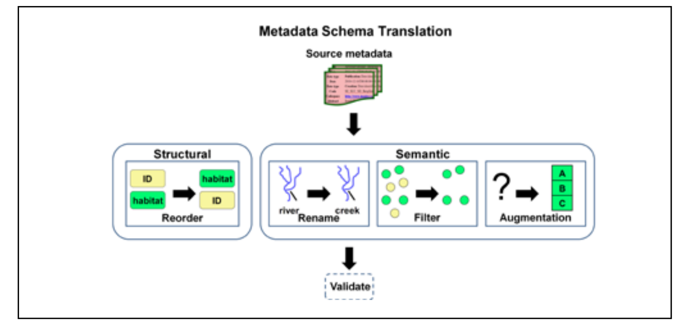


Figure 4. Schema translation operations of metadata

During the matching process, the elements of the source and target schemas that had a shared, or approximately shared, meaning were matched against each other. Before this could be done, an extensive study of the data specifications, metadata profiles, source data and metadata had to be carried out. In the schema mapping process, we defined common operations such as filtering, renaming, reclassification, merging/splitting, reordering, value conversions, morphing and augmentation (Lehto, 2007). During this process, we tried to re-use as many translation operations as possible (as in Hemmatnia et al., 2010). An overview of the operations required is given in Figure 3 and 4.

Validation

Validation languages are usually categorized as grammar-based, such as XML Schema and Document Type Definition (DTD), and rule-based such as Schematron. Grammar-based languages are able to specify rules regarding the structure, form and syntax of documents, whereas rule-based languages may specify the relationships between elements or attributes (Costello and Simmons, 2008). For example, an XML schema may specify certain allowed values in elements <A> and <B>, and the Schematron may thereafter extend this rule by specifying that the allowed values of <B> also depend on the values of <A>. For the data and metadata specifications we used in our tests, both validation languages were needed to cover all the constraints. We carried out our validation tests according to Ford et al. (2011) where both XML schemas and Schematron rules were used for validation (Figure 5).

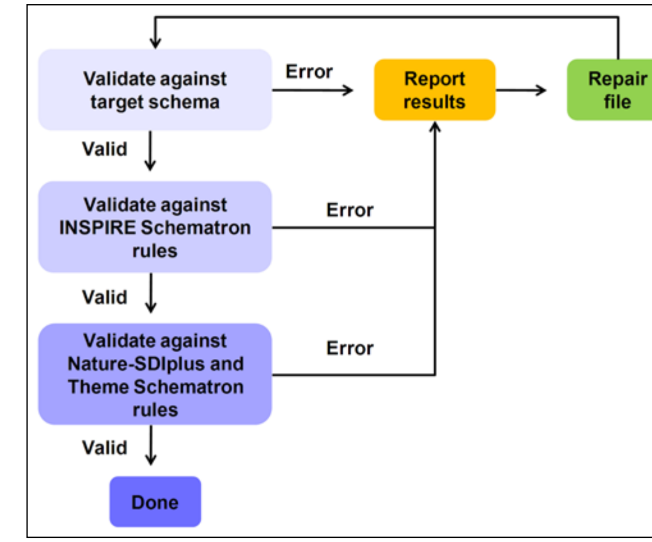


Figure 5. Validation procedure according to Ford et al. (2011).

RESULTS AND DISCUSSIONS

One goal of the Nature-SDIplus project is to harmonize Swedish environmental geodata and metadata so they become compliant with the Nature-SDIplus data and metadata specifications for Bio-geographical regions (BGR), Habitats and Biotopes (HB) and Species distribution (SD). As overall harmonization approach, we used a data warehouse solution and ETL. Moreover, the transformation was split into one spatial and one schema translation part. In the following sections, the results from the establishment of network services, the data harmonization and the validation are presented and discussed.

Establishing of network services

The harmonized data and metadata were published with network services compliant with OGC Web Feature Services (WFS), Web Map Services (WMS) and Web Catalogue Services (CSW). We used mostly open source tools for the establishment, except for the WMS (Figure 6). For download services compliant with WFS, GML 3.2.1 is required as output by the INSPIRE data specifications. Several tools for network services and ETL have had problems handling this format. Nevertheless, Deegree was one of the few tools that were able to support this output. When loading the harmonized data that were to be provided by the WFS, they were stored as the XML data type in an XML-enabled database (in our case PostgreSQL with the spatial extension PostGIS). The reasons for this process flow were mainly to improve the performance by eliminating on-the-fly transformations. Additionally, it facilitated the load and export of GML by not using the relational model since it is fundamentally different from the XML data model (Bourret, 2005) which GML uses. As regards to the WMS, we used the proprietary product Supermap IS .NET, which is one of the leading GIS software products in China. The reason for selecting this product was mainly curiosity and the product fulfilled our expectations. Since WMS is a standard for view services delivering map images, we did not need to fully harmonize the data into GML 3.2.1 for this service.

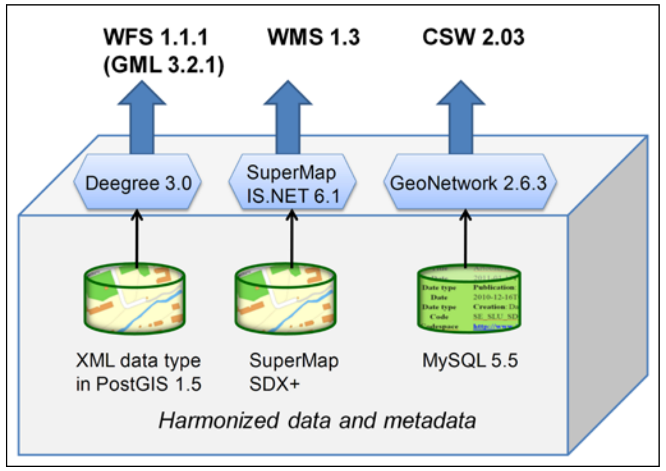


Figure 6. Tools used for the establishment of standardized services

Experiences from the spatial transformations

GRASS GIS was used for the spatial transformations and most operations in Figure 2 were automated with batch processing. However, one unusual transformation was done when harmonizing Swedish species point observations into the NatureSDIplus SD data model. In this model, species distribution should be represented by aggregated fields or polygons. Therefore, we converted the source point data to Thiessen polygons, which reduces the spatial resolution. In addition, the positional accuracy and currentness may be reduced. In the species observation database, multiple observations of the same species in the same location can be reported at different times. The positional accuracy of each observation may also be reported. When creating the Thiessen polygons, observations with same coordinates and year were dissolved into the most recent observation (Figure 7). By doing so, we ensured the highest possible currentness, but with the risk of achieving reduced positional accuracy. This example illustrates the problems of processing spatio-temporal observations.

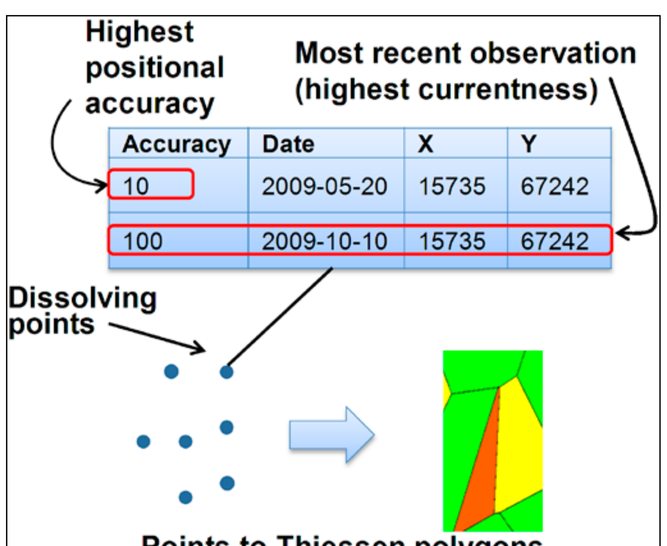


Figure 7. Creation of Thiessen polygons.

Experiences from the schema translations

We used Altova Mapforce to transform the data so they became compliant with the specifications. In this software, the transformation code can be generated in several languages such as Java, C#, C++, Extensible Stylesheet Language Transformations (XSLT), XSLT 2 and XQuery, but it is only the latter three that can be exported and executed in other software environments without modifications. We were, however, unable to efficiently execute all the defined mapping rules with these languages. We do not know whether this is because of limitations in the schema translation tool's functions library or in the mapping languages themselves. XSLT is for instance Turing complete and thus should be able to carry out most calculations (Beare et al., 2010b). Nevertheless, some functions such as database connection and text file import was not possible since XSLT is designed for transforming XML documents (Kay, 2007). Irrespective of the reasons, this is a drawback when the transformations need to be exported and later executed on-the-fly in a vendor neutral environment.

The aim of the schema translation was to fill all mandatory elements according to the target schema and also all optional elements that could be derived from the source data without too much effort. During the schema matching and mapping part, we identified costly transformations and the number of elements in the NatureSDI-plus schemas that had correspondences in the source data and metadata (Table 2). A costly transformation is defined as an operation that required much manual work or that may had negative impacts on the data. Some of these are described in the next two sections. Most of the mandatory elements with no direct correspondences in the source data were easily derived from other elements (in the source data) or from external sources. However, some were more costly; for instance, the metadata lineage element which required knowledge about the datasets process history. Also the conversion of the geometry elements in GML 2.1.2 to GML 3.2.1 was considered somewhat costly for all data themes since it required more time-consuming operations than in other spatial ETL tools. The reason may be that the schema translation tool we used was not specially designed to handle spatial elements.

INSPIRE Theme	Mandatory elements			Voidable elements			Costly transformations	
	Nature-SDIplus	*Provided*	*%*	*Nature-SDIplus*	*Provided*	*%*	*Mandatory elements*	*Voidable elements*
BGR Data	5	3	60%	23	1	4%	1 (20%)	
BGR Metadata	21	18	86%	6	5	83%	1 (5%)	
HB Data	12	3	25%	22	4	18%	3 (25%)	4 (18%)
HB Metadata	22	15	68%	6	1	17%	1 (5%)	
SD Data	8	4	50%	16	6	38%	4 (50%)	5 (31%)
SD Metadata	22	0	0%	7	0	0%	1 (5%)	

Table 2. Costly transformations and number of provided elements. The column "Nature-SDIplus" tells about the total number of elements in the target schemas, whereas the "Provided" columns show how many target elements that had direct correspondences in the source data. Metadata for the Species Distribution (SD) were not available, and therefore we had no direct correspondences for this theme.

Manual operations (costly)

In some cases, manual operations had to be done during the schema translations. The reason was mostly because the source data was quite unstructured. When transforming the Species observations (Table 1) into the SD model, the element distributionInfo in the target model required such operations. This element consists of several child elements which describe the characteristics of the reported species. The following information was sometimes given as unstructured text in comment columns (Figure 8).

- The number of breeding males.
- The number of observed species as range values (as well as integers in the QuantityOr column).
- The residency status (instead of the ActivityOr column).

For these reasons, transformations could not be fully automated.

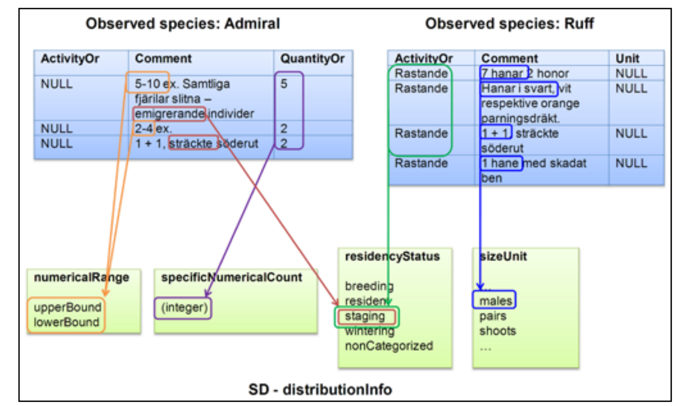


Figure 8. Schema matching between elements in Species observations and the SD model. ActvityOr = Activity, QuanityOr = Number of observed species, Unit = sex of the species.

Another example concerns the Valuable waters based on fish values (Table 1), which was transformed to the HB model. Here, a column containing red listed species (RodHotArt) was mapped to the element habitatSpecies (Figure 9). One solution was then to extract the strings separated by commas and thereafter match them against a list containing Swedish species names and their corresponding scientific names. However, some abbreviations in the source data such as "havs-" and "flod-", and comments such as "sydsvensk." and "(skyddsområde)" required manual renaming operations.

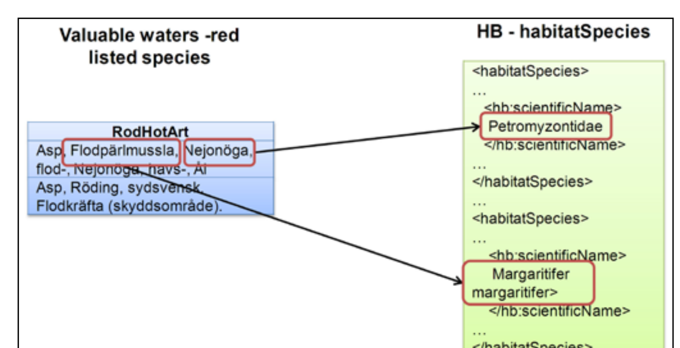


Figure 9. Schema matching between elements in Valuable waters and the HB model

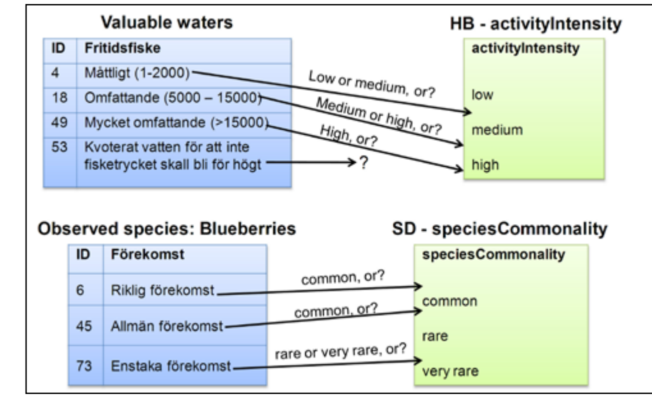


Figure 10. Schema matching between elements in: (1) Valuable waters and the HB model; and (2) Species observations and the SD model.

Approximate reclassifications (costly)

Approximate reclassifications, which may have negative impacts on the data quality, occurred for the source data sets Valuable waters based on fish values and Species observations. The former was transformed to the HB model whereas the latter to the SD model. In both cases, source values that described quantity intervals were matched against enumerations (Figure 10). Both reclassifications may have reduced the thematic accuracy since we were unsure of what quantity intervals the source and target values may have shared. The given quantities may also depend on cultural factors. For instance, what quantity of a certain species do Swedish observers generally define as an isolated occurrence? Furthermore, information may also be lost when values are reclassified back and forth. For instance, when the values (in Figure 10) "Allmän förekomst" and "Riklig förekomst" are reclassified to "common", and thereafter from "common" to perhaps "Allmän förekomst".

Validation

The data and metadata were first validated against their XML schemas and thereafter against the Schematron rules. During the Schematron test we identified non-compliances that were not identified by the schema validation. One such example is that the HB data specification requires that a habitat area should be defined by either a Natura 2000 or EUNIS habitat type. These types of validations are not possible by using ordinary grammar-based validation languages.

CONCLUSIONS AND FUTURE PROSPECTS

The research questions of this paper concerned the harmonization method, tools, costly transformations, quality impacts, validation and architecture of network services. Therefore, the first objective was to test a combined harmonization approach, applied to Swedish data, were the harmonization is made offline, standardized formats such as XML are used and the process is split in spatial and non-spatial part. From this, the following conclusions are drawn

- We successfully harmonized Swedish environmental geodata and metadata so they became compliant with the Nature-SDIplus data specifications and metadata profiles. This shows that the method is feasible.

- Although the usage of XML-based data facilitated the vendor neutrality, the non-spatial transformation procedures expressed in XSLT, XSLT2 and XQuery were insufficient to execute all the mapping rules. One reason is that XSLT is designed to handle only XML transformations. We do not know, however, whether all limitations are because of the mapping languages or because of limitations in the schema translation tool's functions library (XSLT is supposed to be Turin complete). Regardless the reason, manual modifications of the code may be required when executing some of the transformations in a vendor neutral environment.

- By splitting the harmonization into two manageable parts, we avoided some limitations concerning XML schema translations in existing spatial ETL tools. On the other hand,

According to Yeung and Hall (2007) and Broecke et al. (2011), the performance should be improved when using preprocessed data and XML storage in a Data Warehouse. We have made no attempts to validate this statement for our own network services.

The second objective was to extend the validating phase with Schematron tests. When testing against the Schematron rules, non-compliances were found that had been missed during the schema compliance tests. One example is the requirement from the Habitat and Biotopes (HB) data specification which states that a habitat area should be defined by either a Natura 2000 or EUNIS habitat type. This indicates the importance of supplementary semantic rules to ensure compliance against a data specification.

The third objective was to identify processes that were costly as regards to time and quality. Costly (manual) processes were found, which were mainly caused by missing elements and by unstructured information given as comments. For the latter, it was infeasible to automatically extract all relevant information. This may be a problem when many data need to be harmonized. In addition, the positional accuracy got reduced when converting species observation points to Thiessen polygons. Degradations of the thematic accuracy also occurred when approximate reclassifications were done.

Since quality degradations may reduce the usability of the data, further studies are needed for modeling data quality changes in harmonization process. For instance:

- What type of quality degradations that may occur and in what situations?

- What measures should be used? For example, spatial semantic similarity measures (Schwering, 2007) may be developed to measure the possible degradation of thematic accuracy.

- To assure that the harmonized data set adheres to quality requirements, quality assurance procedures (similar to Brennecke et al. (2009)) and optimization methods may be developed.

Another issue is the data volumes when the usage of GML is rising. Storing GML as XML data types in XML-enabled databases may work as a cache and improve performance. Nonetheless, this may complicate the spatial queries and indexes (Lehto, 2009 cited in Hemmatnia et al., 2010), and GML requires large storage space (Lu et al., 2007). An alternative solution is to use Native XML databases with spatial extensions. Since these kinds of databases are using the XML data model, they may be better suited for GML storage (Bourret, 2005). Finally, to streamline the harmonization procedure, transformation tools should be provided as standardized web services. In such way, they will be easily integrated in workflow editors (as in Fichtinger et al., 2010).

Acknowledgements
This work is financially supported by the EU eContentPlus program through the Nature-SDIplus project, project number ECP-2007-317007-Nature SDI Plus, and by the city of Gävle, dnr 06KS23. We also thank Dr Julia Åhlén, for developing the application to extract species observation data from the Artdatabanken database, and the anonymous reviewers for giving valuable feedback.

REFERENCES

Beare, M., Payne, S. and Sunderland, S., 2010a. Prototype Report for the INSPIRE Schema Transformation Network Service. *EC JRC Contract Notice 2009/S 107-153973*, 1 September 2010. RSW Geomatics, 1Spatial and Rob Walker Consultancy.

Beare, M., Howard, M., Payne, S. and Watson, P., 2010b. Development of Technical Guidance for the INSPIRE Transformation Network Service: State Of The Art Analysis. *EC JRC Contract Notice 2009/S 107-153973*, 1 June 2010. RSW Geomatics, 1Spatial and Rob Walker Consultancy. Rob Walker (ed.).

Bourret, R., 2005. XML and Databases. [online] Available at: <http://ece.ut.ac.ir/dbrg/seminars/AdvancedDB/2006/Sanamrad-Hoseininasab/Refrences/6.pdf> [Accessed 29 January 2011].

Brennecke, J., Jakobsson, A., Marttinen, J., Meulenger, L.D., Persson, I., Tsoulos, L. and William, F., 2009. A Common Data Quality Model for Reference Information in Large and Small Scales. *ESDIN Deliverable D8.1, ECP-2007-GEO-317008*, 17 September 2009. Jaana Mäkelä (ed.).

Broecke, J., Kleiker, K., Rammsdonk, K. and Hemmatnia, E., 2011. *INSPIRE Prototypes (Phase 2) – Dutch Kadaster, Release 2.0*. [online] GEORZ-lab & Research. Available at: <http://inspire.kademo.nl/doc/pdf/inspire-prototypes-phase2.pdf> [Accessed 28 January 2010].

Costello R.L. and Simmons R.A., 2008. *Tutorials on Schematron: Two Types of XML Schema Language*. [online] Available at: <http://www.xfront.com/schematron/Two-types-of-XML-Schema-Language.html> [Accessed 24 March 2011].

European Commission (EC), 2007. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), Official Journal of the European Union, Vol. 50.

European Commission (EC), 2009. Commission Regulation (EC) No 976/2009 of 19 October 2009 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards the Network Services, Official Journal of the European Union, Vol. 52.

Fichtinger, A., Rix, J., Schäffler, U., Michi, I., Gone, M. and Reitz, T., 2010 (in press). Data Harmonisation Put into Practice by the HUMBOLDT Project. *International Journal of Spatial Data Infrastructures Research*. (submitted 22 June 2010).

Ford, M., Martirano, G., Schleidt F. and Vinci, F., 2011. Validation Briefcase. *Nature-SDIplus Deliverable D5.1, ECP-2007-GEO-317007*, 17 March 2011.

Gedrange, C., Neubert, M. and Röhnert, S., 2011. Cross-Border Harmonisation of Spatial Base Data between Germany and the Czech Republic. *International Journal of Spatial Data Infrastructures Research*, Vol. 6, pp. xx-yy.

Hemmatnia, E., Broecke, J. and Rammsdonk, K., 2010. Harmonizing Dutch National Geodata Conformant to INSPIRE Using Combined Transformation. *International Journal of Spatial Data Infrastructures Research*, Vol. 5, pp. 365-381.

International Organization for Standardization (ISO), 2006. ISO/IEC 19757-3:2006 *Information Technology - Document Schema Definition Languages (DSDL) - Part 3: Rule-based Validation - Schematron*. Switzerland: ISO.

INSPIRE, 2011. *INSPIRE Workprogramme: Data Specifications.* [online] Available at: <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/2/list/5> [Accessed 20 January 2011].

Kay, M., (ed.) 2007. *XSL Transformations (XSLT) Version 2.0*. W3C recommendation, W3C, January 2007.

Lehto, L., 2007. Schema Translations in a Web Service Based SDI. In: *Proceedings 10th AGILE International Conference on Geographic Information Science*. Aalborg University, Denmark 8-11 May, 2007.

Lu, C.-T., Dos Santos Jr, R.F., Sripada, L.N. and Kou, Y., 2007. Advances in GML for Geospatial Applications. *Geoinformatica*, Vol. 11, pp. 131-157.

Mendive, P., Cardoso, L.J. and Cabello, M., 2010. Procedures for Metadata Profile and data model implementation. *Nature-SDIplus Deliverable D3.5, ECP-2007-GEO-317007*, 17 May 2010.

Schwering, A., (2008). Approaches to Semantic Similarity Measurement for Geo-Spatial Data: A Survey, *Transactions in GIS*, Vol. 12, pp. 5–29.

Steiniger, S. and Bocher, E., 2009. An overview on current free and open source desktop GIS developments. *International Journal of Geographical Information Science*, Vol. 23, pp. 1345-1370.

SFS, 2010:1767. *Lag om geografisk miljöinformation*. Stock-

From Ortelius to OpenStreetMap

Transformation of the map into a multifunctional signpost

Ferjan Ormeling

Utrecht University, The Netherlands
f.ormeling@geog.uu.nl

Introduction

In his story De kaartenmakers [1] ("The map makers"), the Dutch novelist J.M.A. Biesheuvel describes two Norwegians who get their hands on a highly accurate globe, so true-to-life that at one point the border between model and reality fades, and when they study the globe with a microscope they see what is at that very moment actually taking place on the surface of the earth. That story puts into words a secret wish of cartographers: for the landscapes they depict to actually come to life. Actually, this is already possible today in the virtual world; for example, in the computer game SimCity[2] that came out 20 years ago where, almost immediately after you draw a road on a map, you see cars driving on that road. We are coming closer and closer to such a situation in the real world, as well: we see images of the earth's surface – including those of ourselves - coming from satellites in what is extremely close to real-time[3] .

The Antwerp cartographer Ortelius needed a bit more time to collect the information for what we now consider to be the first modern atlas, the Theatrum Orbis Terrarum. In order to collect reliable map material with which to create this atlas, he corresponded with Europe's very best cartographers for ten years before publishing his atlas in 1570. Ortelius marks not only the dawn of modern cartography but also of my career at Utrecht University. The first project, with which I was involved, in 1970, was the preparation for an exposition[4] to mark the four hundredth birthday of Ortelius' atlas. If I compare the practice of Ortelius with the general practice in the 1970s and the situation now, it would appear that more things changed in the world of cartography after this exposition than before it. One exponent of those changes is the recent OpenStreetMap project, in which volunteers collect topographical information on their own. This is an exponent of the current goal of achieving a "well-mapped society", whereby everyone has access to the spatial information that he needs, anytime and anywhere. It is such changes, and their consequences with respect to the future of cartography, that will be discussed here.

Paradigm changes in cartography

The definition of the term *cartography* has gone through quite a few changes during the period that the term has been in use. In approximately 1820, when the term was first introduced in Germany, it encompassed the *production of maps*. When I started my university studies in Groningen, my first instructor in the field of cartography, W.F. Hermans, said the word could simply be defined as *projection theory*, a multiple of ways in which one could depict the earth on a flat surface. He was in good company, too, because the Utrecht professor of cartography Vening Meinesz defined the term similarly[5]. They should have known better, however; as far back as in 1898, the geography teacher Zondervan from Groningen published a cartography textbook[6] in which the field was already referred to as the *visualization of spatial information*. This is a process subject to clear-cut rules, as was demonstrated in the lectures that I attended in 1967 by Koeman, my predecessor in Utrecht. That very year, the grammar of the language of graphics had just been elaborated by Jacques Bertin[7], and we followed his rules as we designed maps so that we could be assured of the proper presentation of geographic information. That term, *geographic information*, had just been introduced in a model by the Czech cartographer Kolačny[8], and that was the impetus for a scientific approach to the transfer of information. This was based upon empirical research: by comparing what map readers read off a map (A', see figure 1) with what cartographers placed on it, (A), one could measure the effectiveness of a map design. The model of this information transfer is no longer used, but it did once play a key role in the development of cartography, because it opened the door for psycho-physical research (that is the comparison of such physical stimuli as map symbols with the perceptual-psychological reactions to those stimuli). It also led to a new definition of the term *cartography*: that definition had then become, in the 1980s, *the production and use of maps*[9].

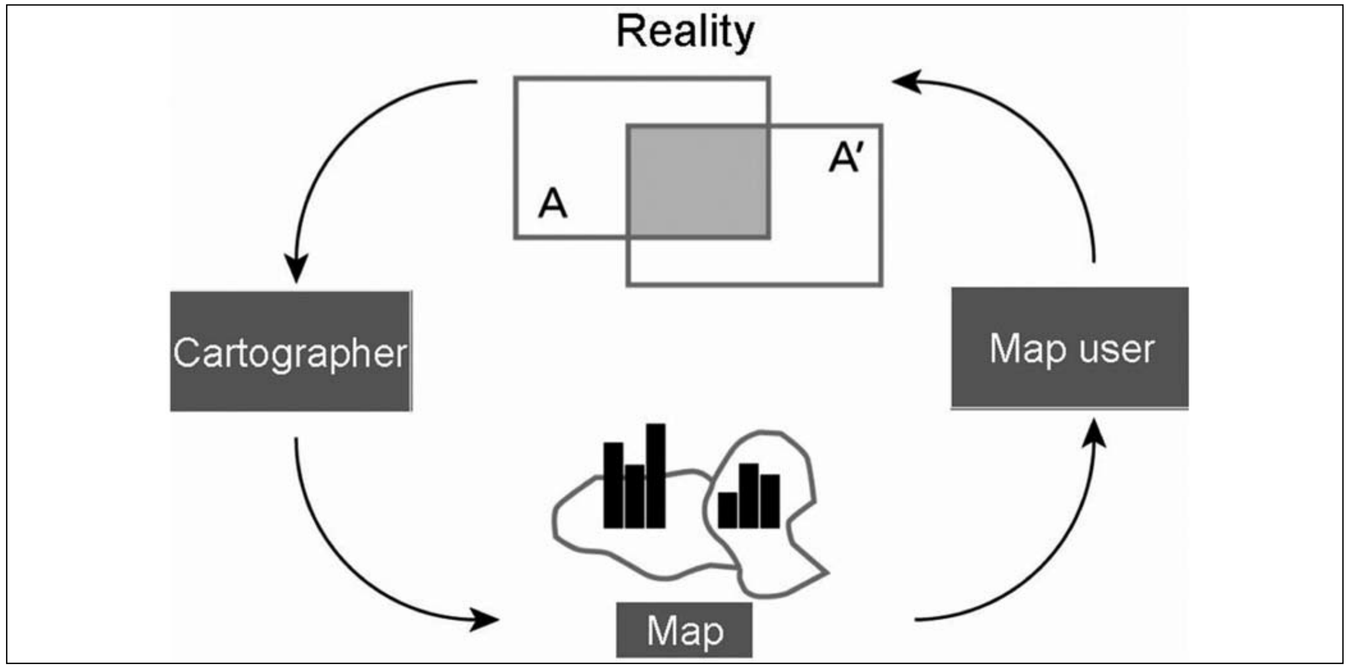


Figure 1 – The cartographic communication model by Kolačny[10]. "INT 1452,© BSH"

Looking back, we did not have enough time to elaborate the research possibilities sufficiently, the process being followed rather too closely on its heels by the development of automation. And Utrecht was a significant part of that development: with the aid of a digitalization unit (Haromat), we carried out research into automatic line generalisation. We experimented with the production of line printer maps; these are maps on which we simulated the various shades of grey by printing letters in various combinations over one another. We see those maps today as incunabula of digital cartography. We learned to work with plotters that could draw borders of areas and also shade those areas. After 20 years of automation, then, we had reached a point where we could use the computer to make maps that were almost as good as the ones that used to be made by hand.

Simultaneously, however, it became clear that the computer could do more than only produce maps: once you had stored the spatial information needed to draw maps in the computer, you could also begin doing some calculations: determining surface area, measuring distances, and carrying out visibility analyses. The part of the field that encompassed this work was called *analytical cartography* or, in more modern words, geo-visualisation. With the new methods of analysis, we had opened the door to geographic information systems. In digital form, then, the information from maps became *a geographic file* which one could then have the computer draw in a form suitable for a specific application.

The arrival of digital geographic files led to a revolution in map production – not so much because we were able to work faster (because if you include all the preparation time, it was certainly not faster!), but because now the map image could be flexibly adapted for various purposes. Once the information was digitally stored in a file, you could easily visualise that which was needed for a certain purpose from that file. In the past, we produced nautical charts that contained all the information that helmsmen might need somewhere at any given time: now all we need on a monitor is the information we require for our own ship: we only need to see the depth contours that are crucial to the draught of our own vessel, making the image considerably clearer (see figure 2).

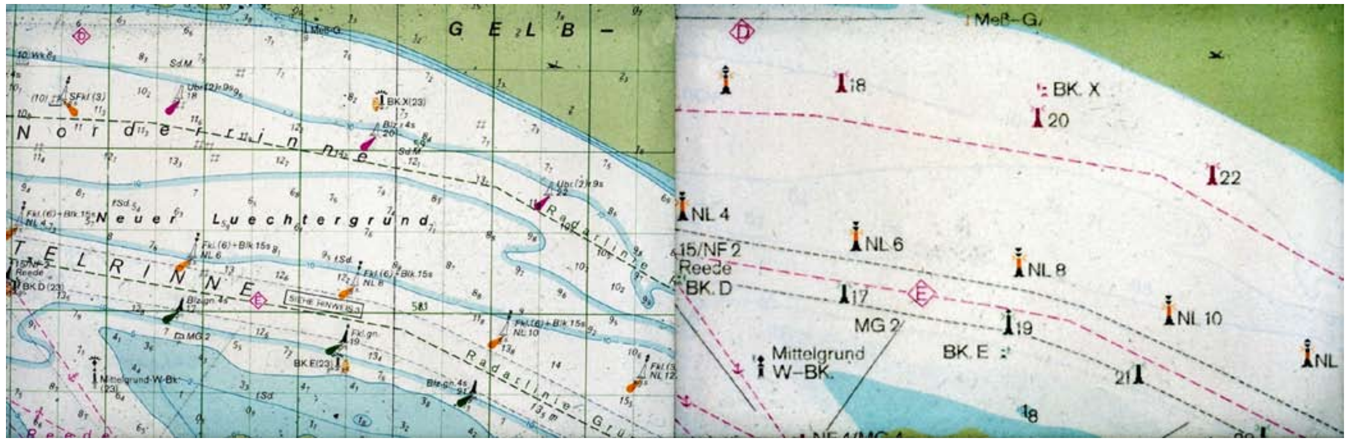


Figure 2: Complete nautical chart and digital version for individual use.

The fact that we were able to use the computer meant that we were able to separate the *storage function* of the map or, in digital form, the geographic file (which describes all the measured aspects of the surface of the earth), from the *communication function* (with which the only objective is to pass on whatever knowledge is required). This breakthrough changes the content of the term cartography once again: now cartography stands for *passing on spatial information to support decision making*.[11] Sometimes this involves maps indirectly, such as in a navigation system in which one listens only to oral instructions, but usually this still takes place based on maps, where we use their unique quality of being able to predict spatial reality as it applies at any given time.

Maps as predictive tools

We view maps as models of reality. The map of *Treasure Island*[12] is a model of a renowned but possibly fictitious Caribbean location. While searching for the treasure, we replace reality by the model, imagining ourselves in that mo-

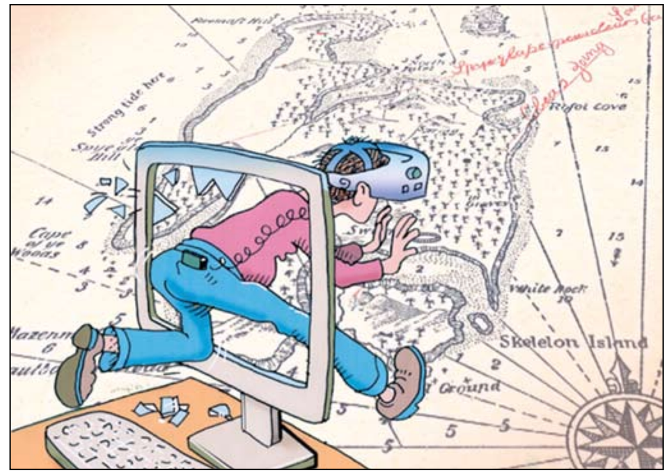


Figure 3: Immersion in the map (drawing A. Lurvink)

del as in a sort of immersion (figure 3). However, when the story ends and the treasure has been found that process nonetheless confronts us with the most important characteristic

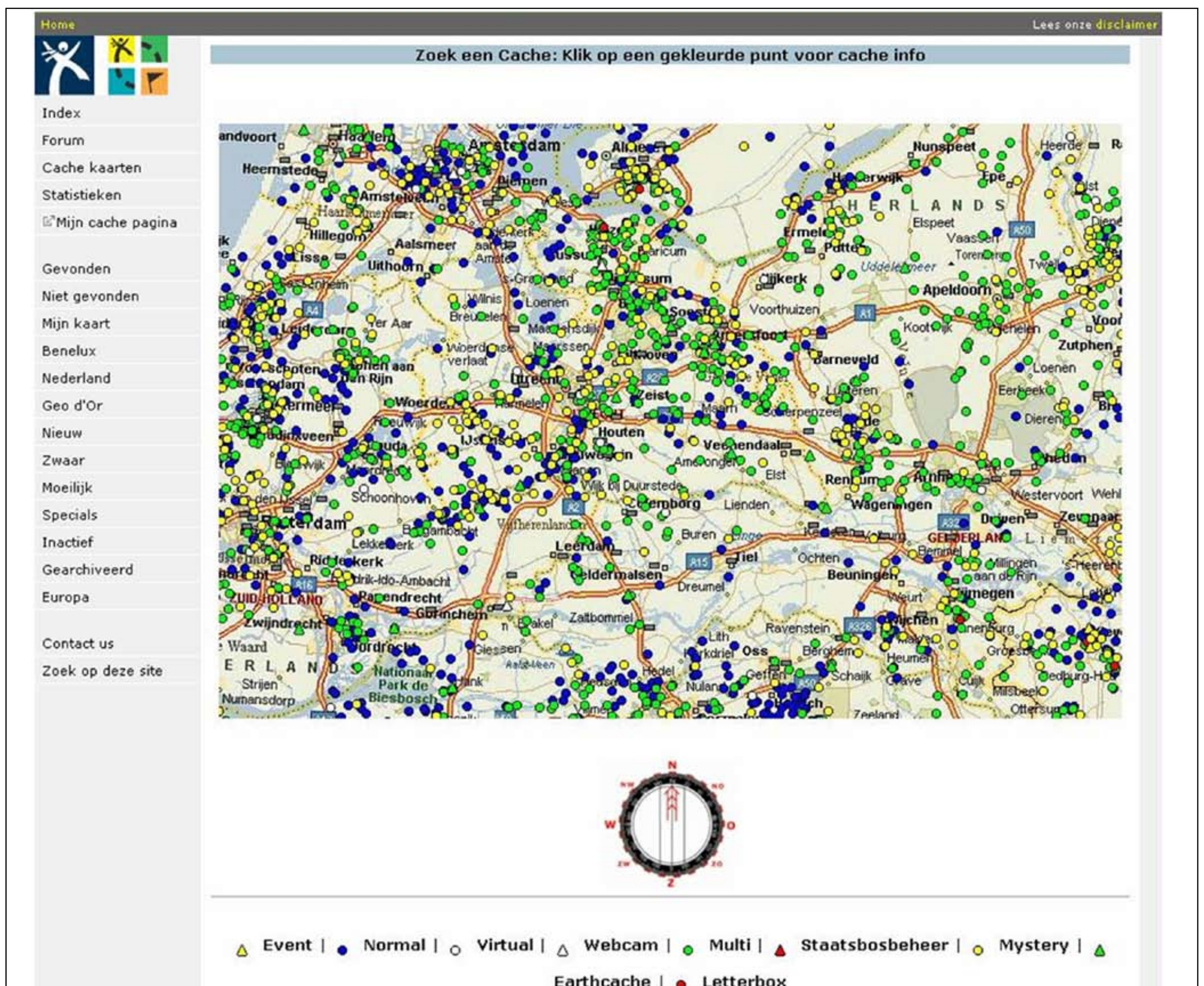


Figure 4: Modern treasure hunting: geo-caches in part of the Netherlands (http://www.geocaching.nl/maps/DisplayCachemaps.php?action=nederland)

of the map, namely that it displays what is in store for us in a spatial sense. If we identify our position in reality, our orientation and destination, on the map, we can determine how we get from one point to the other, and what we will encounter on the way. This is true, in any case, assuming that the map is an accurate model of reality and we obtain the correct impression of reality from that map. Because, after all, this is what it is all about, not that the map is correct but that what we expect in terms of reality from that map is correct. Then we can take relevant and correct decisions. In order for us to be able to form a correct picture of the spatial future, a certain geographic realism and the possibility of animating geographical reality are vital. This is stimulated by displaying the third dimension, for example ground cover images combined with terrain models. What is then still lacking is for us to see how things look at a certain time - during the day or night, in sunshine or snow, at the moment click on a location, as though we were able to place a webcam anywhere we wished. For a few applications, we can already combine the image of weather radar with the map image on our monitor. In this way we obtain living maps - a bit like those described by the novelist Biesheuvel.

These days, searching for a treasure has become a sport: it is called *geocaching*. With their GPS units set at specified coordinates, the innumerable aficionados of this sport (figure 4) search for a treasure in a box hidden somewhere that contains a logbook or camera with which they can confirm that they have found the *cache*. This can be rather difficult using the tiny screen on a GPS-unit: if this is able to show reality in so much detail that they can find their location, that still does not mean they can also recognise the environment – because this requires zooming out, so part of the image must be generated again and there is a risk that one might lose his link with the previous version and thus also with the location. In order to be understood well, information must be visualised correctly on a screen.

It is not only cartographers who make such a GPS application possible – the task also requires specialists in information technology, photogrammetry, remote sensing and geodetics. But it is indeed cartographers that ensure the transfer of spatial information. Cartographers know how to draw information and generalise it correctly if it is to appear without distortion, and how to adapt images to a limited bandwidth and our small mobile screens. We call this *context-specific design* of spatial information.

So we use maps in order to predict a situation at a certain place and at a certain time. Or we use them to determine by which route we can best reach a faraway place. Of course, maps also have other applications, such as analysis, the storage of information, education and advertising. However, being able to make a statement concerning expectations of a situation in some other place is, after all, the most important application. The success of every prediction depends upon the quality of maps – their suitability for an envisaged use – answering the questions of whether they are indeed up to date and complete, whether they contain the right amount of detail, whether their area has been effectively measured, and whether reality has been modelled and categorized in a relevant manner. Together with Menno-Jan Kraak[13.] I recently studied which trends can limit the role of the map as a predictor of space.

The most important trend is the democratization of cartography: more and more map users are generating their own maps from statistical files that are at their disposal, using software packages. They often do this without sufficient cartographic knowledge, so that while the results indeed appear technically attractive they can also give readers an entirely incorrect impression of spatial reality. After all, if you are not aware of the characteristics and possibilities of data to be shown and of mapping techniques, you cannot adjust maps for those areas.

A second limiting factor for the predictive capacity of maps is the increasingly larger gap that is growing between theory and practice. It is the easiest thing in the world for us to combine a wide range of data sets, to carry out overlay operations, or manipulate with buffers, but we do not know how accurate map images resulting from that work will be, even though we are indeed aware of the degree of accuracy of the original maps and files.

We do not even know to what extent, if at all, we may combine various types of sets of data with one another. In addition, we are still too unfamiliar with the degree of accuracy that map images must have in order to guarantee sufficient support of spatial policy, nor are we aware of the likelihood of map users interpreting maps correctly. We do not know whether users read relationships between mapped objects as they are meant to be read, and this is made even more difficult because the range of digital analysis techniques continues to expand daily. Therefore, the area requires research in terms of those questions; however, before I address that subject we must determine whether potential users actually have access, in practice, to the theoretically available spatial information.

Ubiquitous cartography

Just imagine a world in which up-to-the-minute spatial information is always available to everyone who needs it, anytime and anywhere – this is referred to as *ubiquitous cartography*. Imagine you can request and receive information about your environment anywhere you want it, using a mobile computer. Where is the nearest hospital, or theatre, and what is the telephone number you can ring to find out whether there are still tickets available for tonight's show? What type of soil am I standing on, and how do people vote around here? Through the evolution of GPS systems, mobile computers and wireless networks, the realisation of this fantasy is coming more and more within our reach (see figure 5). We have already realized partial aspects of such ubiquitous cartography, such as route specifications for automobile navigation systems.

Other matters demand even further development, such as receiving radiation data perceived by satellites sent to our mobile computers, even though military applications of this probably already exist. The supply of spatial information is also a question of money; in Europe, governments often still charge high prices for such information.

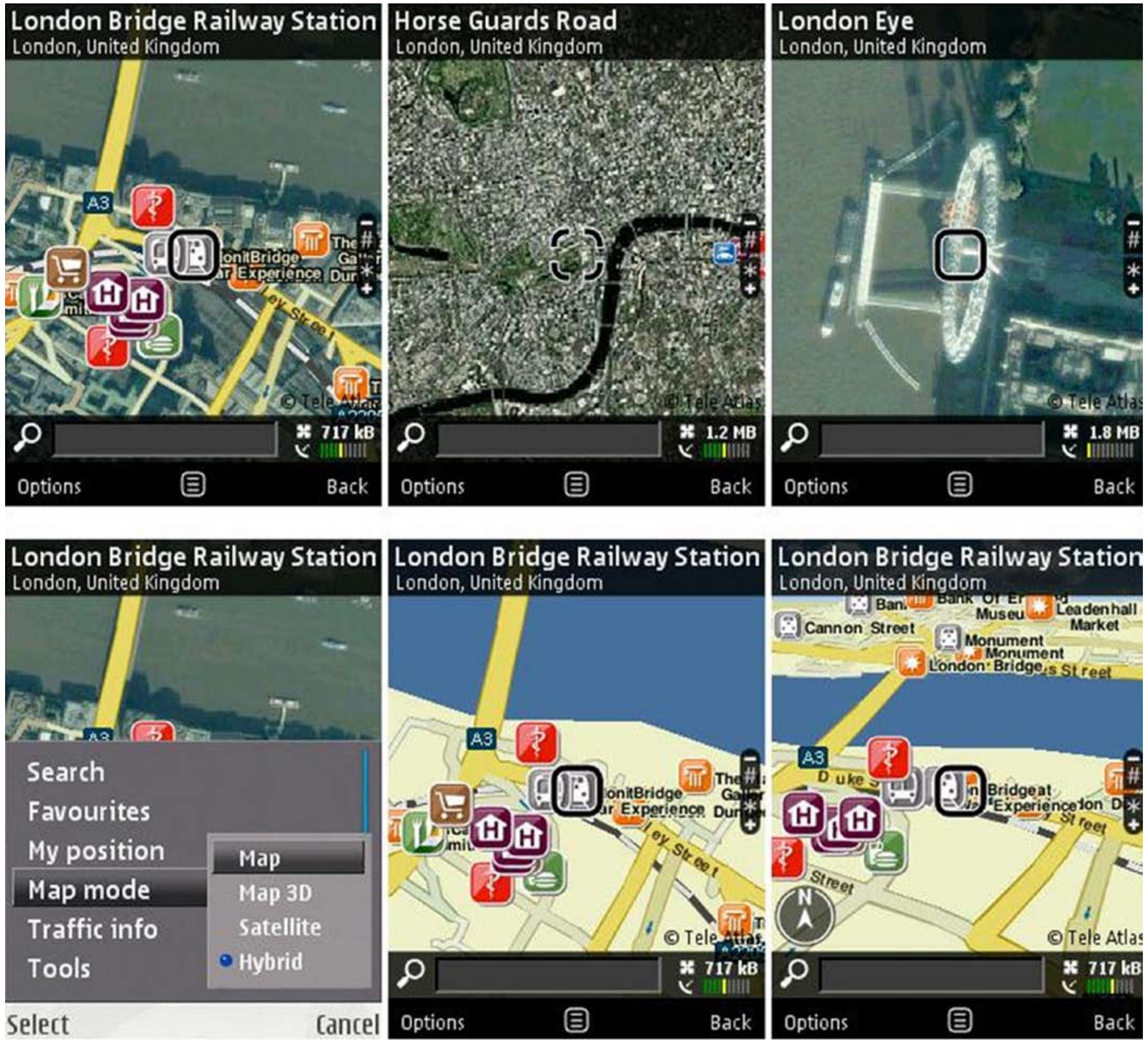


Figure 5 – Various use modes of the Nokia with the Maps 2.0 system (http://www.allaboutsymbian.com/news/item/6704_Nokia_Maps_20_hits_beta.php)

However, over the course of time, we will get to a point where we can combine dynamic information from satellite images with static digital topographic files, so that we can indeed see on our TomTom whether we will run directly into a rain shower if we turn left, or on Google whether the car belonging to the person we want to visit is standing in front of his house or not – so far we still receive the image that was recorded a year ago. How far are we in the Netherlands from those maps that will be available everywhere and at any time?

Changing access to spatial information in the Netherlands

When I started working at Wolters Publishing Company in 1961, all the mapped spatial information of, for example, the Directorate of Public Works and Water Management and the Dutch Forestry Commission, or indeed of the National Geological Service, was actually already available to third parties, that is to say it was available to the general public. The map sheets on which information was recorded concerning coastal defence, nature management or infrastructural matters were for sale, so that non-professionals had a realistic option of viewing this data and thus also having a voice with regard to its future development.

With the exception of the maps of the Topographical Survey, all of the above-mentioned services have now for some time been switched over to the establishment and maintenance of information systems from which their employees can obtain the information they need for their own use. However, if these files have indeed been made accessible to people who are not employees of specific agencies, their use now does require a high level of technical knowledge and a well-filled wallet: it is usually only engineering firms that can afford to acquire the files they need in order to carry out their projects. There are certainly not many Dutch people who know how to work with the important digital geographic files that have been established in order to keep our intensively used delta running smoothly, such as the hydrological information system or the large-scale basic map on a scale of 1:1000. And thus has the provision of spatial information to the greater public been compromised considerably. The same has taken place outside our country, and there it has led to the development of Public Participation GIS, or Participatory GIS[14], an attempt to make GIS techniques and government data files clear and accessible to a broader public, which is to result in the actual establishment of realistic possibilities to share in decision-making as the transparency of government decisions based on GIS activities becomes greater. Here lies one of the tasks of the cartographer of the future.

On the other hand, some government departments at a national, provincial and municipal level have developed websites where the non-professional can obtain information about a number of environmentally-related topics free of charge. Some excellent examples are the cultural-historic files of provinces, or their risk maps. In addition, the commercial sector has taken over a number of tasks from the government. Our city maps used to be based on cadastral maps, but now they are more frequently based on information obtained by commercial map production companies themselves. Such companies as TeleAtlas and NavTeq, who specialize in car navigation systems, as well as Google and Microsoft, have recording vehicles driving around that collect geographical information and convert that information to files from which they can also make road maps. Google (Earth), Microsoft (Virtual Earth), and Terravision produce files based on satellite or aerial photographic recordings with which we can zoom in on Internet, at the expense of advertisers, on any area down to such a level that we can even see our own houses . We can navigate through a city on our own and see the city in three dimensions at any point, in the direction of our choice.

25 years ago, in my inaugural address[15] in 1985, I attacked the deficient cartographic visualisation by the government. I showed examples of the poor visualization of spatial information meant for the Parliament. Since then, ministries have set up visualization departments and not only has the official provision of information since then been greatly improved, but this also applies to the visualisation of spatial data in general: the quantity of maps in the media has increased tremendously, as has the quality of that visualisation. Government departments produce map images or even atlases for the web. An absolute high point is the National Atlas of Public Health of the National Institute for Public Health and the Environment (RIVM) (figure 6). In addition to the increased range of maps that are available, the interest in these maps has increased as well. No one would have dared predict that a national atlas for 100 Euros, the Bosatlas van Nederland, with over 50,000 copies sold would ever top the Dutch list of best sold books[16].

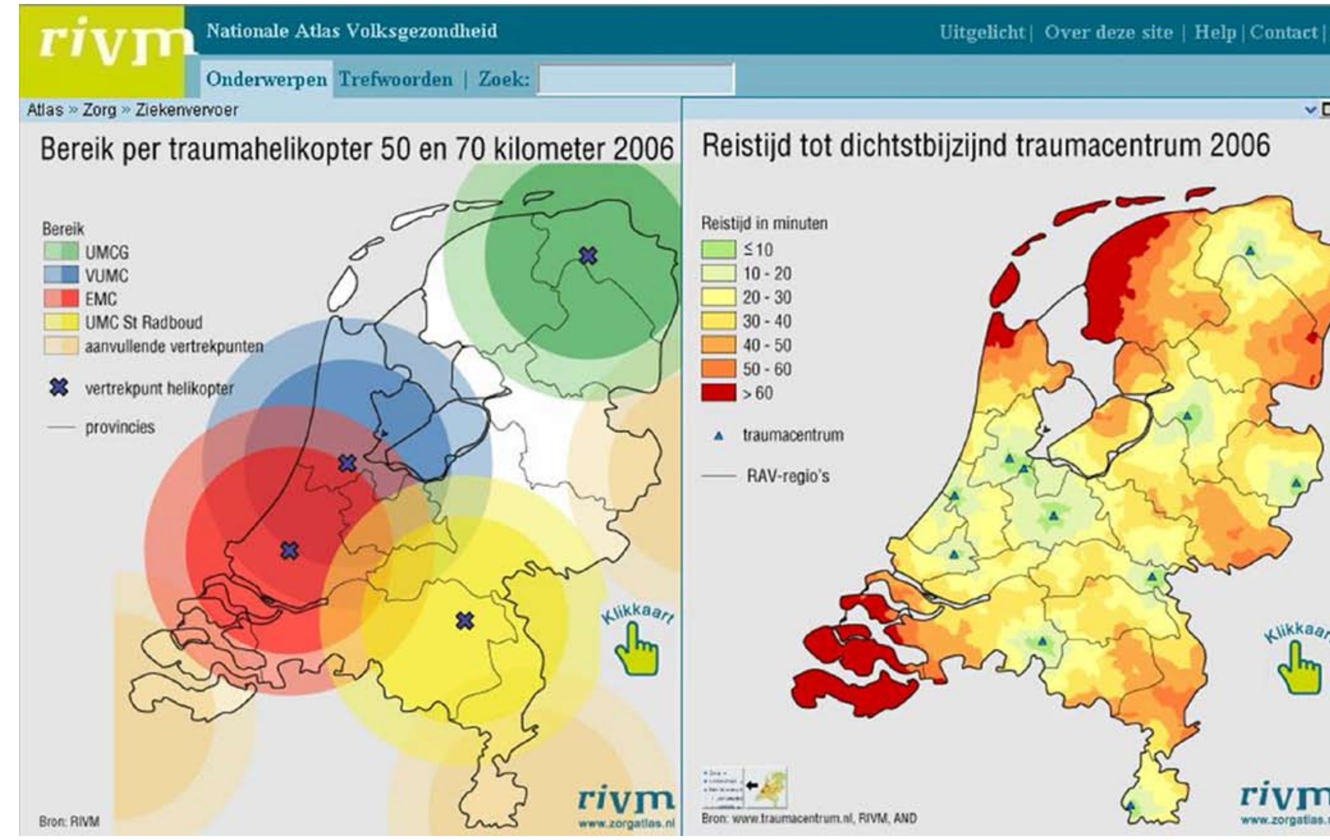


Figure 6. Two maps from the National Atlas of Public Health of the RIVM (http://www.zorgatlas.nl/algemeen/menu/english/).

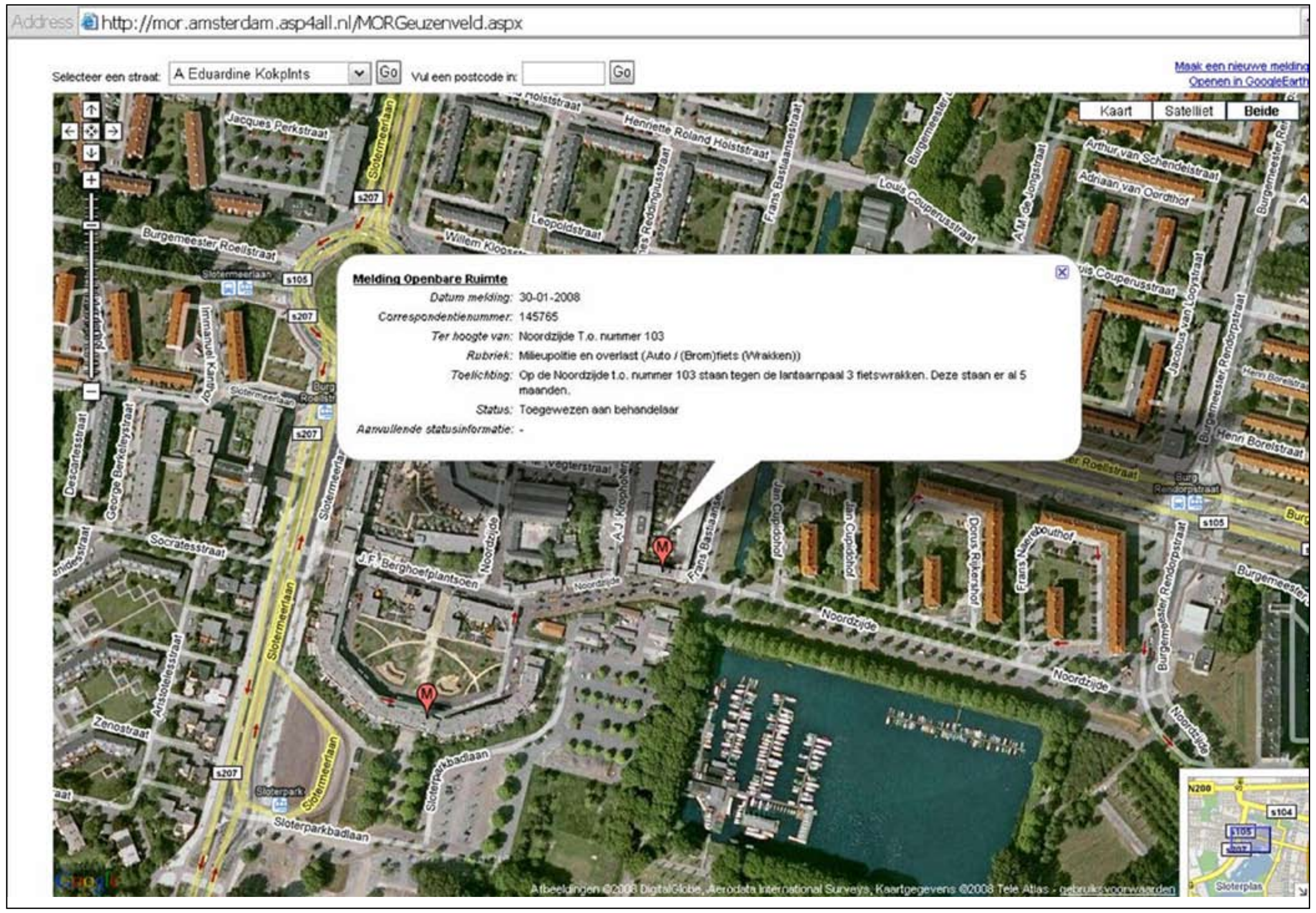


Figure 7 – Report based on the mash-up principle of the district of Geuzenveld-Slotermeer in Amsterdam

Therefore, in addition to there being a wider range of spatial information from the commercial sector, in the Netherlands it has become in some ways more difficult and in other ways easier to acquire specific spatial information from the government. How is the public reacting to this?

Geotagging, OpenStreetMap and Web 2.0

Map production has been democratized. Maps can be adapted more and more to the interests of the user, not only with a bit of reality chosen by that user, given form with the symbols of his preference, in the desired perspective, and also with his own notations, but also - more generally - equipped with information that he adds. At the time of the hurricane Katrina in 2005, the widespread use of *mash-ups* was made clear on a broad scale for the first time: Due to the suboptimal provision of information by the American government, victims themselves were forced to seek ways to find their lost relatives: by referring to a map of the city on Internet, and linking their houses on that map with information about, for example, their temporary address or information about the condition of their house. Residents of the Geuzenveld district in Amsterdam can do the same thing if they wish to report vandalised street furniture for repair to the local authorities (see figure 7). Of course, in order to do this they must be able read a map and have some geographical knowledge. And, as has been proven in some popular TV shows[17], not everyone has that knowledge.

Geotagging, or assigning geographic coordinates, for example to holiday photos, is the latest manifestation of our desire to pinpoint our position. Flickr.com is a website where you can upload your photos so that you can show them to everyone. Last month, some 2 million geotagged photos were added to the site. This has led to a gigantic reservoir of photos of which it is known precisely where they were taken, and if you search at a certain location you will find many photos of that location. On Funda, too, a popular real estate website in the Netherlands frequented by the partners of retired cartography professors, one can find pictures of houses based on the map. By means of geotagging with photos, it is possible to determine more than an exact location; you can also find out about the altitude, time, date and compass direction of the view, so that the photo is almost reproducible, given the right weather conditions and camera. Therefore, because it is now possible to carry out searches in terms of coordinates, both Funda and geotagging use the organising characteristic of mapping. So mapping something also means organizing spatial information! A comparable initiative is *Wikimapia*, where a person can attach sub-maps or remarks to objects on a map.

Another application of the 'mapping urge' is the already mentioned *OpenStreetMap* initiative of making maps oneself, and not leaving that task to government organisations that charge a great deal of money for the files they generate using tax money.

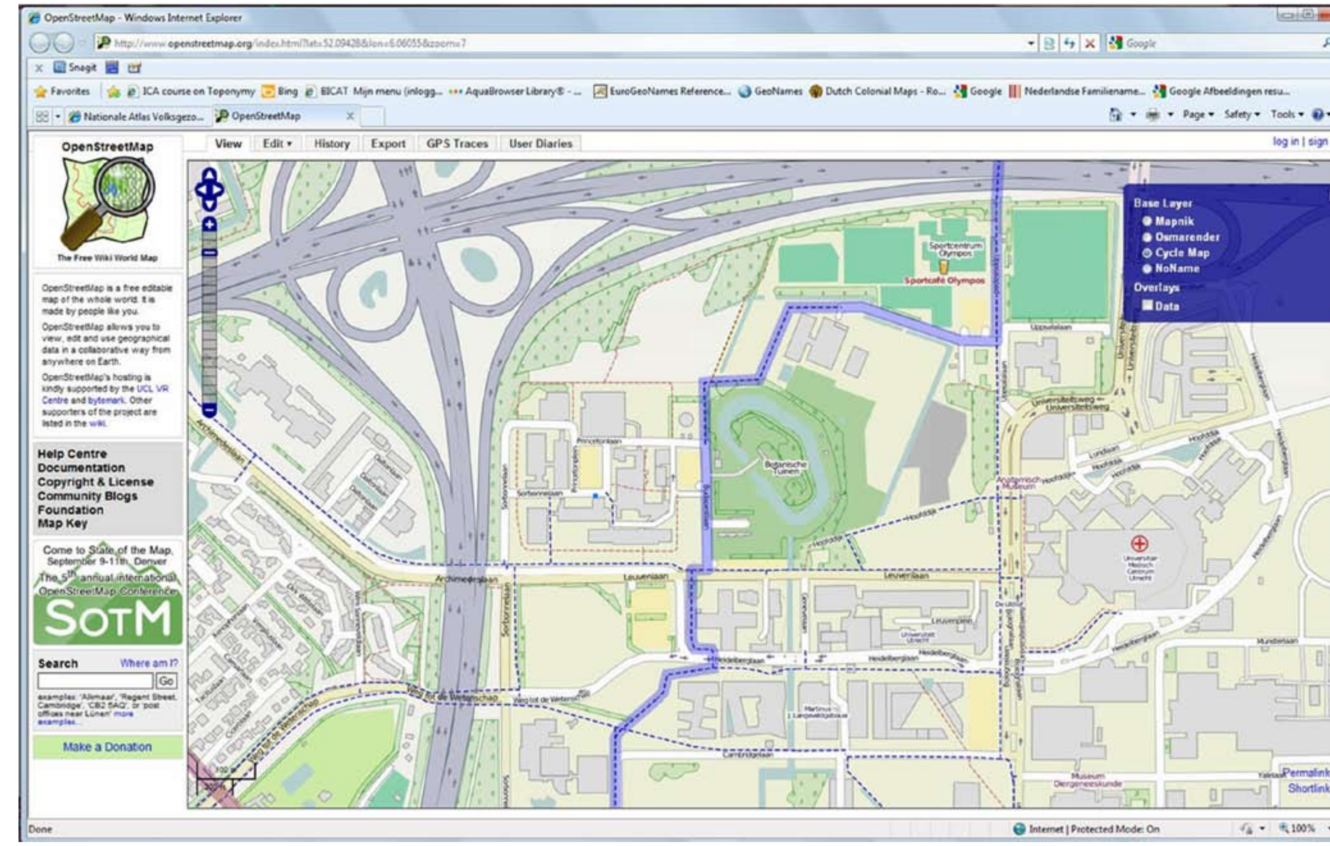


Figure 8 - Detail OpenStreetMap of the Uithof area in Utrecht (http://www.openstreetmap.org/index.html?lat=52.09428&lon=6.06055&zoom=7)

OpenStreetMap is a project focused on the generation of freely available geographic data (figure 8), such as for road maps and city maps, for anyone who wishes to do this. The project was started in England, where the topographic service charges high copyright fees for the use of its information, so that many creative uses of geographic information are not possible. The Netherlands is also mapped according to this project; people often hold 'openstreet mapping parties' at which the coordinates of all roads of a certain area are established using GPS systems. These are all cartographic applications of Web 2.0, the platform on which people join forces to create their own information, of which Wikipedia is also a manifestation. It is participants who determine and control their own data; not data suppliers.

Therefore, a great deal of spatial information has become available, from the government, from companies and from private parties, because people are making maps more than ever before. But to what extent does this information reach its users? It will require research to answer this question

Cartographic research/ Research agenda

From recent publications by the National Research Council (NRC) in the United States, one notes that the interest in spatial data among the public continues to increase[18] . In order to respond to this and steer it, the NRC has prepared a list of research priorities: this means that (1) there must be better inventories of all the spatial data that is now already available, to support science and decision-making, that (2) that spatial data must be made accessible and usable for everyone, that (3) that accessibility must be guaranteed everywhere and all the time, and that (4) it must be made possible for people to work jointly with spatial data from various locations, with the aid of spatial information technology, in order to solve spatial problems. The latter is called *geocollaboration*[19], and people are developing techniques with which they can share their spatial files and digital maps remotely, view them simultaneously, process them together, and make joint decisions on the basis of these facts. In addition, all participants in a project must be able to show each other their location and their spatial information on the same map.

In addition to these NRC plans, we can view the research agenda of the ICA, the International Cartographic Association[20] . This is a programme that we have developed over the last eight years, the goal of which is to steer research efforts in the commissions of the ICA. This focuses partially on analysing large files in order to be able to identify changes based on *data mining* or *change detection* techniques. The development of spatial analysis techniques, the establishment of the quality of our geographic files, and the assessment of the uncertainty inherent to analyses of combinations of files are points that are high on the agenda.

In my opinion, the most fertile topic for cartographic research, in addition to data quality and generalisation, continue to be the psycho-physical studies that I already mentioned at the beginning, which now plays a role in the *usability* studies[22] (here, that is research regarding the effectiveness and efficiency with which certain map users reach a specified goal in specific circumstances). We still know too little about how to use the information on maps and insert this into our current knowledge. During the last five PhD studies that I monitored, the thinking-aloud laboratory (figure 9) was used during this type of study, a laboratory set up for usability studies at the International Institute for Geo-information Science and Earth Observation (ITC).

This research listing of the ICA also contains the history of cartography, because I consider it necessary for those practicing the profession to know how spatial information was collected, visualised and used in the past. If you would look at an old map (figure 10), you would usually just see an attractive graphic image. Historians of cartography, however, would see more: they would see a landscape as perceived through the eyes of a cartographer, containing the information that the client of that time considered important and which the cartographer gave form in his own specific way. So a map is both a source of knowledge of the landscape of a given time, of the society that had it mapped and a representation of the ideas and expertise of a cartographer.

Standardisation is a factor that is also part of the applied research that is vital to the many plans we have with regard to the exchange of information in the future. One aspect that we have actually neglected in the exercise of both cartography and geography (this is certainly true in the Netherlands, which, despite the years of recommendations of the United Nations, is the only country in Western Europe where there is still no geographic names bureau), is the standardization of geographical names. These names are essential to maps; they form the most important interface for the lay user who wishes to know more about his environment.

Figure 9 – Laboratory setup for thinking-aloud research on the usability of maps at the ITC[21].

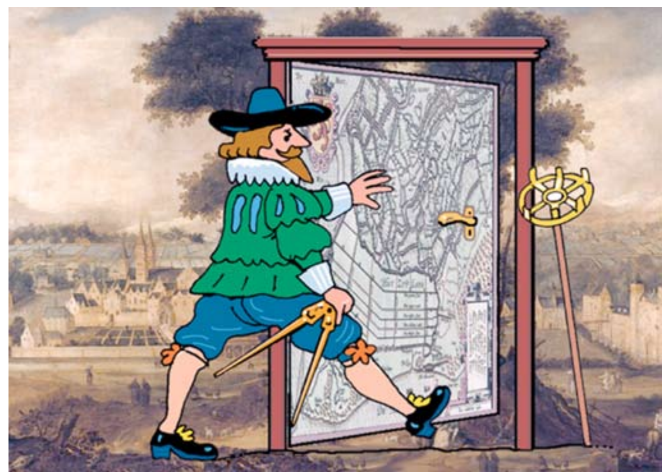


Figure 10 – Old map as a source of knowledge of the early landscape, but also of the cartographer and his patron (drawing A. Lurvink)

Standardisation of geographic names

To me, geographic names have become one of the most fascinating aspects of cartography. The standardisation of geographic names is essential to spatial communication, because one must be certain that one is speaking about the same location when it comes to ambulances, food assistance or travel in general. In the context of the United Nations Group of Experts on Geographical Names[23]. I am permitted to participate everywhere in courses in the collection and processing of geographic names on behalf of maps, and while doing this see how differently different people deal with names: in Lesotho the name of a village is associated with the name of the head of that village, so it will change when he dies. In New Guinea, during shifting cultivation, when people move to another location and thus take their village to that new location, they take the name of the village with them. A bit further south, among the aboriginals of Arnhem Land, men and women each have their own set names for objects in their environment. If we remain in that area and look at the Javanese, they use various versions of names of places in High and Low Javanese, which they use depending upon whether they are speaking to a socially higher or lower level of person[24]. There are differences in the names used by young people and old people, too, because the young are more likely to use slang such as saying LA instead of Los Angeles. In areas where there are linguistic minorities, different versions of names appear for topographic objects. The climate can also lead to different names. Because summer and winter landscapes in some areas, such as in Northern Canada, vary so much from each other, entirely different objects must be given names in the different seasons. Just try to standardise while keeping all these factors in mind!

The international standardisation of geographic names (an important aspect of international communication) is based upon national standardisation, and such a thing has never existed in the Netherlands. According to the recommendations of the Spelling Act of 1947, a proposal was made in 1960 for the way to write geographic names in the Netherlands, but that proposal disappeared into the recesses of the desk of the minister of the Interior. Since that time an incorrect interpretation of municipal law has been a laissez-passer for municipal governments to establish names of topographic elements in their territory in regional dialects. This represents a threat to

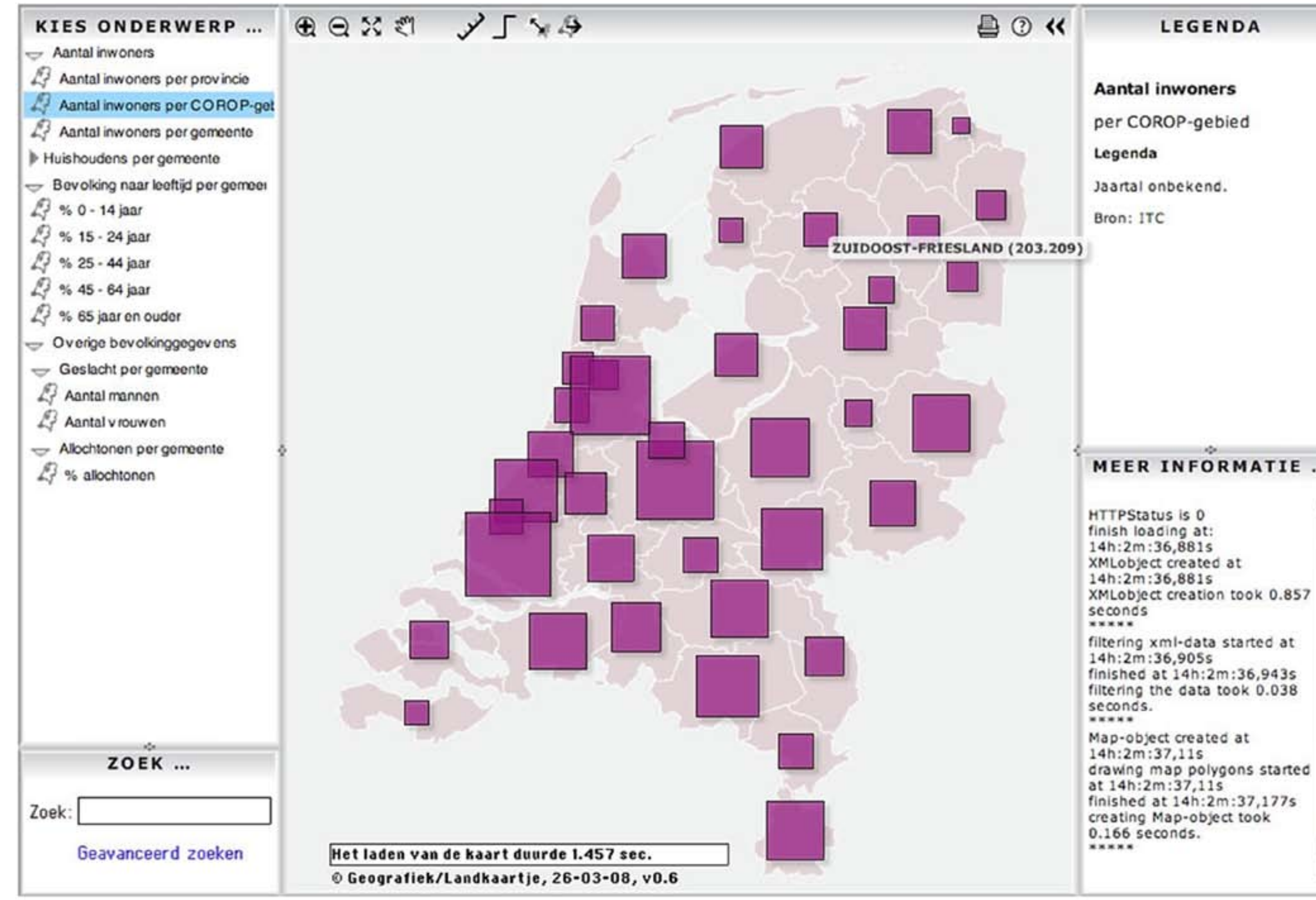


Figure 11 – Screen layout of the new Atlas of the Netherlands. Design by Geografiek and Landkaartje.

the de facto agreement on the spelling of geographic names[25] that was reached, as a result of long and tedious consultation, by mapping services in the 1950s and 1960s.

The standardisation of geographic names is also important when it comes to searching for information based on those names. In that context, cartographers in Utrecht participated in the eContent-project *EuroGeoNames*[26] financed by the European Commission. This was a predecessor of the Inspire project, and its objective was to create a virtual European database of geographic names by combining various different existing databases. The use of unequivocal geographic names optimizes the search function of maps…and this brings me to the portal function of maps and atlases.

The map as data portal

In the field of cartography, atlases are viewed as the ultimate challenge, because the information they contain must be coordinated not only within one map but also between various different maps. The best atlas is then a national one, the most detailed presentation of the spatial knowledge about a country. By way of Internet we can give the national atlas an extra dimension by *geocollaboration*, whereby different institutions collaborate in order to supply spatial information, via a central point or data portal. A website has been developed for this purpose in the United States, called (http://www.geodata.gov), "your one stop for federal, state and local geographic data" (that is to say, the geo-spatial one-stop)[27]. An atlas structures our view of the earth, familiarises us with geographic concepts, and is therefore eminently suited to function as an interface with the GDI, the spatial data infrastructure. This makes an atlas more than just a costly cage in which one captures the earth. We are working on creating such an interface in collaboration with the ITC and commercial partners[28] in the context of a Geo-Information Drive research project[29], based on the national atlas of the Netherlands[30]. With that atlas as a metaphor, we are developing an alternative, sustainable map-oriented access to the geodata infrastructure, via user-friendly solutions, in order to make geo-information accessible to the greater public, something that is also an objective of GeoNovum, the new trailblazer of the national geo-information infrastructure. This requires responsible, systematic visualization, because it is essential that the maps can also be compared with one another. See also figure 11. Improving access to spatial information is also consistent with the objectives on a European level[31] .

In order to keep up with the latest developments, atlases must also keep up with the possibility of including data from their users. The Canadian cartographer Fraser Taylor speaks, in this context, of *Cyber cartographic atlases*[32] , atlases that form contexts within which user-generated data as well as such social digital networks as Web 2.0 and Wiki can be easily integrated. This sounds fantastic, because in this way we enable people to provide information that they consider to be relevant. But is this consistent with the concept of the atlas? Ortelius collected information from the world's best cartographers, whose approved map material (the best available at that time) he used, and this is what made his atlas such a success. So are we going in the right direction with these cyber-atlases? Is active civilian participation enough? The American movement of Critical Cartography[33] thinks that it is. But in my opinion we are running the risk, with such atlases to which anyone and everyone can contribute his own information, that – without exercising professional control over the contents to be added - we are replacing quality by consensus, so that in the long run no one will any longer be able to truly depend on the data. The cartographer of the future will have to monitor the processes of the collection, design and use of information (such as in the OpenStreetMap), in order to inform the public about what is already available in terms of spatial information, so as to ensure that the relevant information is collected and properly given form for each specific application, and he will also have to contribute to the professional use of visualised spatial information. There certainly are challenges enough in our profession!

In closing

The research and educational tasks that I mentioned above will also be carried out in the future, if not in Utrecht, then somewhere else in the world. Of course, I hope that cartographic research and education will continue to remain anchored at Utrecht University and that new cartography lecturing staff will once again be appointed, before all the investments that have been made in the past lose their value. Over the years, Utrecht has been a significant centre of cartographic education and research. It is there where, in the middle of the 19th century, the professor of National Economics and Statistics J. Ackersdijk already was an expert in the field of statistical cartography. It was from here that the astronomer Prof. J.A.C. Oudemans, around 1880, supplied the topographic maps of Indonesia with their correct mathematical principles. In 2008 we are celebrating not only the 50th anniversary of the Netherlands Cartographic Society[34] but also the fact that cartographic lecturing staff have been associated with this faculty for a hundred years, a faculty that can point with pride to its status as the sole university-level graduate cartography programme in the Netherlands from 1970 to 2000. For specialized areas with which technical developments have caught up, tradition need not be a guarantee for continued existence, but I believe that I have demonstrated that cartographic experts will also be needed in the future, as long as we use our eyes to obtain the spatial information we need from maps. The producers of the Bosatlas, the major Dutch school atlas, recently appointed the first cartographer educated in Germany. This person is an outstanding specialist in his field, but that appointment does not detract from the fact that there must continue to be the possibility at Dutch universities for the further development of cartography as an important form of communicating spatial information.

To me cartography, now associated with GIS, has been an inseparable part of geography since the beginning of the university study of that field. I thank the geographers who shared that opinion in the past for having given us the opportunity to practice, teach and further develop our speciality for so many years in Utrecht.

Thank you for your attention.
Ferjan Ormeling

(abbreviated, translated version of valedictory address at the occasion of his retirement as professor of cartography at Utrecht University, April 23, 2008)

Notes

1) Biesheuvel, J.M.A. (1979) De verpletterende werkelijkheid, pp 84-96. Amsterdam: Meulenhoff .
2) Maxis (Emeryville, California) - Sim City. 1989
3) See also YouTube: http://nl.youtube.com/watch?v=fPgV6-gnQaE
4) Bergeijk, V. van, W. van den Dries & E. Micheels (1970) 400 jaar atlas. Utrecht: National University, Institute of Geography
5)Vening Meinesz, F.A. (1941) Kort overzicht der kartografie. Groningen-Batavia: P. Noordhoff.
6) Zondervan, H.(1898) Proeve eener algemeene kartografie, Leiden: Kapteijn.
7) Bertin, J. (1967) Sémiologie Graphique. Paris/The Hague: Mouton.
8) Kolačny, Antonin (1969) Cartographic information: a fundamental concept and term in modern cartography. The Cartographic Journal 6: 47-49, 1969.
9) Ormeling, F.J. and M-J.Kraak (1987) Kartografie. Ontwerp, productie en gebruik van kaarten. Delft: Delft University Press p. viii.
10) Kolačny, A. (1970) Kartographische Informationen – ein Grundbegriff und Grundterminus der modernen Kartographie. Internationales Jahrbuch für Kartographie vol X, 1970, pp 186-193.
11) In the mission of the Netherlands Cartographic Society (NVK) it was stated in 1996: "Cartography is making accessible and transferring spatial information with a view of solving spatial issues, emphasizing visualisation and interaction" (Kartografie is het toegankelijk en hanteerbaar maken en overdragen van ruimtelijke informatie met nadruk op de visualisatie en interactie, afgestemd op het oplossen van ruimtelijke problemen." (NVK adresboek 1996/97, The Hague: Netherlands Cartographic Society, p.6). In 2003 cartography was described in the Strategic Plan of the International Cartographic Association as: "the unique facility for the creation and manipulation of visual or virtual representations of geospace – maps – to permit the exploration, analysis, understanding and communication of information about that space." http://cartography.tuwien.ac.at/ica/en/ICA_Strategic_Plan_2003-08-16.pdf
12) Robert Louis Stevenson (1883) Treasure Island.
13) Ormeling, F.J. and M-J.Kraak (2007) Maps as predictive tools. Mind the gap. Keynote held at the 23rd International Cartographic conference of the International Cartographic Association, Moscow 2007. pp 18-21, Abstracts of Papers, 23rd International Cartographic Conference . Roskartografiya: Moscow 2007.
14) Sieber, R. (2006) Public participation geographic information systems: a literature review and framework. Annals of the Association of American Geographers 96 (3), 2006, pp.491-507.
15) Ormeling, F.J. (1986) - Imaging (Beeldvorming) (Inaugural address). Kartografisch Tijdschrift 12, 2 (1986), 16 pp.
16) The Bosatlas van Nederland (Wolters-Noordhoff Atlasproducties. Groningen 2007) was among the Top Ten books in the week of 27/10/2007 (NRC Handelsblad 27-10-2007).
17) "De vakantieman". Travel programme broadcast by RTL4 1990-1999, presented by Frits Bom in collaboration with Peter Mekenkamp (Department of Cartography, University of Utrecht).
18) Committee on Intersections between geospatial information and information technology, National Research Council (2003) IT Roadmap to a geospatial future. Washington: National Academies Press.
19) Parker, Chris (2007) Shaping a future for geographic information. Geo Informatics, March 2007, pp 12-16.
20) Virrantaus, Kirsi and David Fairbairn (2007) ICA research agenda on cartography and GI Science. Keynote held at the 23rd International Cartography Conference of the International Cartographic Association, Moscow 2007.
21) From: Elzakker, C.P.J.M. van (2004) The use of maps in the exploration of geographic data. Utrecht / Enschede: KNAG / Geoscience Department, University of Utrecht / ITC. Netherlands Geographical Studies 326.
22) "To date, virtually nothing is known about the usability of geospatial technologies. Even less is understood about the extent to which those technologies can be matched to human conceptualizations of geographic phenomena or about the use to which the information will be put. It will be necessary to develop new tools to track how individuals and groups work with geospatial technologies, to assess which approaches are most fruitful, and to identify the usability impediments imposed by the technologies. Such understanding will be vital for tailoring user-centered design and other usability engineering methods to the needs of general audiences working with geoinformation.
In particular, it will be important to establish which techniques can measurably improve how effectively and productively geo-information is used by the general public, students, and other non-specialist audiences. As noted previously, current HCI research methodologies look at people's interaction with technology rather than at how technology is applied to support people's interaction with information. Cognitive and usability assessment techniques do not address visually enabled technologies or ones intended for application to ill-structured problems." p 93, National Research Council (2003) IT Roadmap to a geospatial future. Washington: National Academies Press.
23) See for the activities of the UNGEGN: http://unstats.un.org/unsd/geoinfo/UNGEGN/bureau.html.
24) I thank Tjeerd Tichelaar for the example of the kromo and ngoko-version of the name of the garrison city between the mountains Arjuno and Semeru: Mambeng resp. Malang. In addition there are examples for the pairs of names Pasedahan/ Pasuruhan and Kedinten /Kediri .
25) Ormeling, Ferjan (2005) - Verzameling en vaststelling van de spelling van aardrijkskundige namen in Nederland. See http://www.meertens.nl/books/veldnamen/ormeling.pdf
26) See for a description of the project http://www.eurogeographics.org/eurogeonames.
27) Goodchild, M.F., Pinde Fu and Paul Rich (2007) Sharing geographic information: an assessment of the geospatial one-stop. Annals Association of American Geographers 97 (2), 2007, 250-266.
28) Geografiek (http://www.geografiek.nl/) and Landkaartje (http://www.landkaartje.nl/).
29) Ruimte voor Geo-Informatie(RGI)-project no 111, National atlas as portal to the Geodata-infrastructure.
30) Kraak, M-J., F.J. Ormeling, W. Broeder, E. MacGillavry, W. van den Goorbergh - The Dutch National Atlas in a GII environment: the application of design templates. Abstracts, 23e International Cartographic Conference of the International Cartographic Association, Moscow 2007.
31) European Umbrella Organization for Geographic Information (2000) Towards a strategy for geographic information in Europe – A consultation paper. Eurogi: Apeldoorn.
32) Challenge for the industry is brainware. Interview with Fraser Taylor, GIS Development, December 2007
33) See Crampton, Jeremy W. and John Krygier (2006) An introduction to Critical Cartography. ACME: An International E-Journal for Critical Geographies, 4 (1), 11-33.
34) In 2002 the Netherlands Cartographic Society became part of the association Geo-Informatie Nederland

PRESENTATION OF AUTHORS

Jonas Ågren, Lantmäteriet Gävle

Jonas Ågren has a doctor of technology degree in geodesy from the Royal Institute of Technology (KTH) in Stockholm. He presently works at the Geodetic Research Division of Lantmäteriet, mainly with research and development in geoid determination, gravimetry and geodynamics. He is chairman of the Working Group for Geoid and Height Systems within the Nordic Geodetic Commission (NKG).

Runar Svensson, Lantmäteriet Östersund

Runar Svensson has a master of science degree in geodesy and photogrammetry from the Royal Institute of Technology (KTH) in Stockholm. He has long experience of geodesy and surveying, both from Lantmäteriet's production and development divisions. He now works as a cadastral surveyor in Östersund, Sweden.

Perola Olsson, GIS Centre, Lund University, Lund, Sweden

Perola Olsson received a Master of Science in Geomatics at Lund University, 2011. The title of the master thesis is "Cartography in Internet-based view services - methods to improve cartography when geographic data from several sources are combined".

Karsten Pippig, Institute of Cartography, Technische Universität Dresden, Germany

Karsten Pippig studied Cartography at Dresden's University of Technology and finished his studies in 2010. Within the scope of his diploma thesis, he concentrated on automated generalisation in map production. Since 2011 he is member of the academic staff and PhD student at the Institute of Cartography, Technische Universität Dresden. His current research is related to mobile map applications based on user generated content.

Lars Harrie, GIS Centre, Lund University, Lund, Sweden

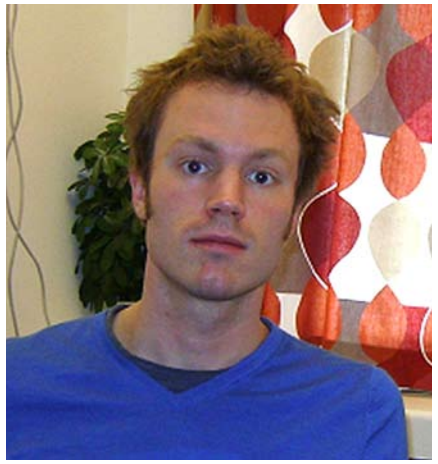
Lars Harrie received a Master of Science in Geodesy from KTH, Stockholm, Sweden in 1993 and a PhD in Geographic information science from Lund University in 2001. Since 2006 he is senior lecture at the GIS Centre, Lund University. His main research interests are algorithms for processing and presenting geographic data, methods for improving the national spatial data infrastructure, and map services on the Internet.

Hanna Stigmar, GIS Centre, Lund University, Lund, Sweden

Hanna Stigmar received her Master of Science in Surveying at Lund Faculty of Engineering in 2001. She received her PhD degree in 2010 at Lund Faculty of Engineering, Sweden. Her thesis focused on methods to improve the legibility and usability of maps in real-time map services. She is currently working at the GIS centre at Lund University.

Kirsten Rassmus-Gröhn

Kirsten Rassmus-Gröhn is a senior lecturer at the Department for Design Sciences at Lund University. Her research focuses on non-visual and less-visual aspects of cross-modal interaction design in mobile and stationary contexts. The goal of the design in the mobile context is to limit the users' need for visual attention to a small screen in route



Finn Hedefalk, Gävle GIS Institute, University of Gävle

Finn Hedefalk has a BSc in Geographic Information Systems and an MSc in Geomatics. The last three years he has worked in various European research projects in Italy and at the University of Gävle. In the projects, he has mainly addressed geodata harmonization issues in the EU-member states. Within the project GeoTest, he is now testing the implementation feasibility of the INSPIRE Annex II and III data specifications. His main interests are GIS and related sciences, and his research interest lies within geodata harmonization and quality assurance methods.

Anders Östman, Gävle GIS Institute, University of Gävle

Anders Östman is a Land Surveyor (KTH 1978) and he has also a PhD in photogrammetry (KTH, 1986). Östman worked with GIS development at Intergraph Scandinavia AB and as a professor in Geographic Information Technology at the Luleå Technical University. Since 2004, he holds a professorship in geomatics at the University of Gävle. Östman is one of the founders of AGILE, a European research organization. He is also a Swedish academic representative of the European Spatial Data Research (EuroSDR).

Since the early 1990s, Östman works on issues related to Spatial Data Infrastructures. During recent years research focus has been on questions related to schema transformations, test of SDI components and 3D modeling.

Ferjan Ormeling, Utrecht University

Ferjan Ormeling (1942) has been teaching cartography from 1969 onwards, since 1985 as professor, and has been engaged in atlas cartography since he started to work as a part-time atlas editorial assistant in 1961 with Wolters-Noordhoff Atlas Publications. His other main interest is toponymy; he did his PhD on the representation of minority names on official topographic map series in Europe. This has paid off, as he is now vice-chair of the United Nations Group of Experts on Geographical Names (UNGEGN), and he was one of the initiators of the EuroGeoNames project. He is engaged in organizing toponymy courses all over the world for the UNGEGN. As he has problems in keeping up with new digital developments in cartography he now focuses on the history of cartography in the former Dutch colonies and is expected to write a monograph on the development of thematic cartography in the Netherlands. From 1987-1999 he chaired the ICA Commission on Education and Training, and from 1999 to 2007 he served as the ICA Secretary-General.

Avsändare:
Kartografiska Sällskapet
c/o Lantmäteriverket
801 82 Gävle

Kart & Bildteknik

Mapping and image Science

Scientific edition