R-09-46

Uncertainty aspects of the digital elevation model for the Forsmark area

Mårten Strömgren, Lars Brydsten
Umeå University

October 2009

Svensk Kärnbränslehantering AB
Swedish Nuclear Fuel
and Waste Management Co

Box 250, SE-101 24 Stockholm
Phone +46 8 459 84 00

ISSN 1402-3091
SKB Rapport R-09-46

Uncertainty aspects of the digital elevation model for the Forsmark area

Mårten Strömgren, Lars Brydsten
Umeå University

October 2009

Keywords: Digital elevation model, DEM, Topography, GIS, Forsmark, Surface ecosystem, Biosphere, Error calculation, Quality assesment.

Abstract

A digital elevation model (DEM) describes the terrain relief. A proper DEM is an important data source for many of the different site description models conducted in the Forsmark region. Input data for the Forsmark DEM is elevation data for both land and sea areas of different origin and quality. No statistical analysis of the error in the Forsmark DEM is so far carried out. However, the Forsmark DEM is part of the quality assessment of the regolith depth model for the Forsmark area since it represents the upper surface of the regolith depth model. The aim of this project was to calculate the errors in different areas in the Forsmark DEM and present them in terms of general descriptive statistics.

Measurements have confirmed the knowledge that the 0.25-metre DEM produced from the laser scanning measurements in the Laxemar-Simpevarp area is of very high quality. The 0.25-metre DEM was used to calculate the errors of the 10 and 50-metre DEMs, and the errors for different sea shoreline sources. These error distributions were placed randomly among points for the same data sources in the Forsmark area and used for correction of the original elevation levels. Using the corrected input data for the 10 and 50-metre DEMs and for the sea shoreline, a new DEM was produced. All other input data remained unchanged. The error for the Forsmark DEM was calculated for areas within the data sources corrected from the 0.25-metre DEM. The 0.25-metre DEM from the Laxemar-Simpevarp area was also used for a calculation of how density of input data points used in interpolation affects quality in a 20-metre DEM.

Part of the input data was removed in the sea area, new DEMs were produced and compared to the existing Forsmark DEM within the areas of the removed data, to get a measure of the error in these areas of the DEM.

In areas of input data for the sea shoreline, the quality of the Forsmark DEM is high. The errors within the SKB 10-metre DEM are slightly less than within the extension for the 50-metre DEM. However, the statistical analysis has revealed errors larger than 10 m within the areas of these data sources. In the sea area, the highest quality is found within the extensions of the measurements from shallow bays, the base map, and the detailed area of the measurements from the Geological Survey of Sweden (SGU). In the regional area of the measurements from the Gelogical Survey of Sweden, the uncertainty is much larger and errors of almost 10 m have been calculated. In the extension of the digital nautical chart/nautical chart, the uncertainty is very high with a mean error of almost 5 m, a standard deviation of a little more than 3 m, and a maximum error of more than 35 m.

No error calculation was done in the lakes in the Forsmark area. However, comparing the maximum distance between points in the lakes with measurements from the sea area would suggest a quality at least as good as in the shallow bays. No error calculation was done for the inlet channel to the nuclear power plant and nor for the small areas where false depth values are placed.

Sammanfattning

En digital höjdmodell (DEM) är en modell som beskriver reliefen i terrängen. Den är en viktig del av indatat till olika modeller som tas fram över Forsmarksområdet i samband med platsbeskrivningarna. Indata till den digitala höjdmodellen över Forsmarksområdet är punkter som anger höjdnivåer för både land och hav från ett stort antal olika datakällor av olika ursprung och kvalitet. Ingen statistisk analys av felen i höjdmodellen över Forsmarksområdet är så här långt utförd. Höjdmodellen över Forsmarksområdet ingår dock in den kvalitetsbedömning som har gjorts för jorddjupsmodellen över Forsmarksområdet då den utgör överytan för jorddjupsmodellen. Syftet med det här projektet var beräkna felen i olika delar av den digitala höjdmodellen över Forsmarksområdet och presentera dessa med beskrivande statistik.

Mätningar med totalstation har bekräftat att den digitala höjdmodellen med 0,25 m upplösning, framtagen utifrån laserscanningen i Laxemar-Simpevarpsområdet, är av mycket hög kvalitet. Denna höjdmodell har använts för att beräkna felen för höjdmodellerna med 10 och 50 m upplösning och felen för olika datakällor för kustlinjen i Laxemar-Simpevarpsområdet. De beräknade felfördelningarna slumpades ut bland punkter för motsvarande datakällor i Forsmarksområdet och användes för att korrigera de ursprungliga höjdnivåerna. En ny digital höjdmodell tillverkades genom att använda korrigerad indata för höjdmodellerna med 10 och 50 m upplösning och för kustlinjen. Ingen förändring gjordes av övrig indata till höjdmodellen över Forsmarksområdet. Felen i höjdmodellen över Forsmarksområdet beräknades inom områden där datakällorna korrigerats med felfördelningar från höjdmodellen med 0,25 m upplösning. Denna hödmodell användes också för att beräkna hur olika punkttäthet i indata till interpoleringen påverkar kvaliteten på en digital höjmodell med 20 m upplösning.

En del av indata för havsområdet togs bort, nya höjdmodeller togs fram och jämfördes med den nuvarande höjdmodellen över Forsmarksområdet inom områden där indata tagits bort.

I områden med indata för kustlinjen är kvaliteten på höjdmodellen över Forsmarksområdet hög. Felen inom SKB:s 10 m höjdmodell är något mindre än inom höjdmodellen med 50 m upplösning. Den statistiska analysen har dock visat på fel större än 10 m inom utsträckningen för dessa höjdmodeller. I havsområdet är kvaliteten högst där mätningar för de grunda havsvikarna gjorts, inom underlagskartan och inom SGU:s detaljerade område. I SGU:s regionala område är osäkerheten mycket större och fel på nästan 10 m har beräknats. Osäkerheten i den del av höjdmodellen över Forsmarksområdet som ligger inom det digitala sjökortet/sjökortet är mycket stor med ett medelfel på nästan 5 m, en standardavvikelse på drygt 3 m och ett maximalt fel större än 35 m.

Ingen felberäkning utfördes i sjöarna i Forsmarksområdet. En jämförelse mellan de maximala avståndet mellan punkter i sjöarna med mätningar från havet skulle dock antyda att kvaliteten i sjöarna är minst lika hög som i de grunda havsvikarna. Ingen beräkning av felen gjordes för inloppskanalen till kärnkraftverket och de små ytor där falska djupvärden är placerade.

Contents

1 Introduction

For siting of the repository of spent nuclear fuel, SKB has undertaken site characterisation at two different locations, Forsmark and Laxemar-Simpevarp. The surface system part of the site descriptive model includes, e.g. hydrology, Quaternary deposits, chemistry, vegetation, animals, human population and land use. Access to a proper digital elevation model (DEM), describing the terrain relief, is important for many of the different models constructed for the Forsmark region. No statistical analysis of the uncertainty in the Forsmark DEM is so far carried out. However, a quality assessment of the regolith depth model over the Forsmark area, in which the Forsmark DEM represents the upper surface, is presented in /Hedenström et al. 2008/. The aim of this project was to calculate the errors in different areas in the Forsmark DEM.

DEM resolution is the size of the rectangular DEM cells. A DEM is constructed by an interpolation of irregular spaced elevation data. In these models, Kriging interpolation was used. Kriging is a geostatistical interpolation method based on statistical models that include autocorrelation (the statistical relationship among the measured points). Kriging weights the surrounding measured values to predict an unmeasured location. Weights are based on the distance between the measured points, the prediction locations, and the overall spatial arrangement among the measured points.

Input data to the Forsmark DEM have many different sources, such as existing DEMs, elevation lines from digital topographical maps, paper nautical charts, digital nautical charts, and depth soundings in both lakes and the sea /Strömgren and Brydsten 2008/. The quality of these data sources is also different. The density and distribution of input data is different for land and sea areas; most of the data in land areas is distributed in regular grids throughout the landscape. In sea areas, on the other hand, the elevation data is distributed densely along survey lines and there is no elevation data at all in large areas between the survey lines.

These differences are the reason why different methods are used for calculation of the error for land and sea areas in the Forsmark DEM. For land areas, error distributions calculated for similar input data in the Laxemar-Simpevarp area from the high quality 0.25-metre DEM produced from the laser scanning measurements /Nyborg 2005/ were used for calculation of the errors. For sea areas, part of the input data was removed, new DEMs were produced without the removed data, and areas of the removed data were used for the statistical analysis of the errors in the existing Forsmark DEM. ArcGis 9.2 was used for all GIS-calculations.

2 Method

2.1 Calculation of errors in the land area in the Forsmark 20-metre digital elevation model

Six sources were used to collect elevation point data for land areas of the existing 20-metre DEM for the Forsmark area: 1) the existing digital elevation model (DEM) from the Swedish national land survey (LMV) with a resolution of 50 metres. 2) the SKB DEM with a resolution of 10 metres /Wiklund 2002/, 3) measured values from lakes /Brunberg et al. 2004/, 4) fixed points from the property map, 5) brook measurements /Brydsten and Strömgren 2005/, and 6) areas referring to sea shoreline.

In the land area in the Forsmark DEM, the focus of the error calculation is on the areas of the 10 and 50-metre DEMs, and in the area of elevation data for the sea shoreline. The errors for the areas of fixed points and brook measurements are not calculated. However, due to the high quality of input data, these small areas should be considered of higher quality than the rest of the land areas. The lakes will be discussed later in this report.

In October and November 2007 total station measurements of grid points (elevation values converted to points from the DEM cells) from the 0.25-metre DEM /Nyborg 2005/, the SKB 10-metre DEM /Wiklund 2002/, and the 50-metre DEM from the National Land Survey (LMV) were performed in the Laxemar-Simpevarp area /Strömgren and Lindgren 2010/.

The result of these measurements shows that the 0.25 metre DEM is by far of highest quality (Table 2-1). Based on the excellent quality, the 0.25-metre DEM should be very representative of the land surface of the Laxemar-Simpevarp area within its extension of an approximately 17.5 km^2 large area. If we make the assumption that the 0.25-metre DEM is almost equal to the true land surface, the statistics of the errors for the 10 and 50-metre DEMs, and sea shoreline data for the Laxemar-Simpevarp area can be calculated: Grid points from the 10 and 50-metre DEMs were joined to grid points from the 0.25-metre DEM in ArcGis 9.2. This GIS function (point to point join) gives a new attribute with the distance to the closest point in the join to dataset. Points in the actual dataset with a distance of 0 metre were selected, i.e. points having exactly the same coordinate as points from the 0.25-metre DEM. Points from the 10 and 50-metres DEMs (169,263 and 6,608 points, respectively), were selected and exported to new point layers. The difference between elevation levels from these 10 and 50-metre grid points and the grid points from the 0.25-metre DEM were calculated (Table 2-2). Using Zonal Statistics in the ArcGis 9.2 Spatial Analyst extension, the difference between the 0.25-metre DEM and sea shoreline data referring to property map, digitized sea shoreline, and sea shoreline measured using DGPS (differential global positioning system) was calculated. No laser scanning measurements are performed in the Forsmark area. Input data for the land area of the existing Forsmark DEM, is as mentioned earlier, large areas covered with grid points from the 10 and 50-metre DEMs.

Table 2-1. Statistical analysis of the difference between the 0.25, 10, and 50-metre DEMs, and total station measurements. The unit is metre. No normal distribution test is performed. The last three columns show the 5, 25, 75, and 95 percentiles, respectively. This calculation shows that the 0.25-metre DEM is by far of highest quality.

Data source	Number of measure-ments	Mean	Median	Standard deviation	Minimum difference	Maximum difference	5%	25%	75%	95%
0.25-metre DEM	484	0.01	0.00	0.19	–1.99[1)]	0.73	–0.26	–0.07	0.11	0.24
10-metre DEM	493	0.34	0.38	1.86	–7.33	4.84	–2.95	–0.43	1.18	3.59
50-metre DEM	69	0.00	0.21	1.72	–5.29	2.86	2.75	–1.15	1.30	2.38

[1)] The minimum error in the 0.25-metre DEM is in very steep terrain, where a small difference in horizontal direction gives a very large error.

Table 2-2. Statistical analysis of the difference between the 10 and 50-metre DEMs and the 0.25-metre DEM from the Laxemar-Simpevarp area. The unit is metre. No normal distribution test is performed. The last three columns show the 5, 25, 75, and 95 percentiles, respectively. The dispersion from the mean error is larger in the 50-metre DEM compared to the 10-metre DEM. However, larger errors are found in both data sources.

Data source	Nr of points	Mean	Median	Standard deviation	Minimum difference	Maximum difference	5%	25%	75%	95%
10-metre DEM	169,263	0.31	0.33	1.11	–11.16	9.29	1.47	–0.21	0.89	2.03
50-metre DEM	6,608	0.36	0.23	2.17	–8.92	12.00	2.72	–0.93	1.35	3.78

If we make the assumption that the error distributions for the 10 and 50-metre DEMs, and the sea shoreline data are the same for the Forsmark area as for the Laxemar-Simpevarp area, it is possibly to calculate the error for the Forsmark DEM in areas covered by input data from these sources. The error distributions from the Laxemar-Simpevarp area were placed randomly among grid points from the 10 and 50-metre DEMs, and sea shoreline data for the Forsmark area. Using the error distributions from the Laxemar-Simpevarp area, the data from the Forsmark area was corrected. All grid points from the 10 and 50-metre DEMs, and sea shoreline data from the property map, digitized sea shoreline, and sea shoreline measured using DGPS, used as input data to the original Forsmark DEM were replaced by corrected points, all other input data remained unchanged. Using the same Kriging parameters, 20-metre resolution, and extension for the interpolation procedure as for the existing Forsmark DEM /Strömgren and Brydsten 2008/, a new DEM was interpolated from this new data base. This new DEM should represent a raster surface produced from true elevation levels in data from the 10 and 50-metre DEMs, and sea shoreline data. The difference between the existing Forsmark DEM and this new DEM reflects the error in the existing Forsmark DEM using incorrect input data for the 10 and 50-metre DEMs, and the sea shoreline (Table 2-3).

However, this only shows the consequence of error in input data used for the interpolation procedure and takes no consideration to the point density of input data. Using input data to the Oskarshamn DEM within the extension of the 0.25-metre DEM is a method to investigate how point density affects quality in the Forsmark DEM. Grid points placed every second metre within the 0.25-metre DEM extension, i.e. 121 points/20-metre cell are input data to the Oskarshamn DEM from the 0.25-metre DEM /Strömgren and Brydsten 2008/ (Figure 2-1a). The point density for the 10 and 50-metre DEMs are only 4 points and 0.36 point/20-metre cell, respectively (Figure 2-1b and c).

Table 2-3. Statistical analysis of the difference between the existing Forsmark DEM and a DEM where input data for the interpolation procedure for the 10 and 50-metre DEMs, and sea shoreline data, was corrected using randomly placed error distributions for the corresponding data sources from the Laxemar-Simpevarp area. Only 20-metre cells within the 10 and 50-metre DEMs, and the sea shoreline extensions were used for the calculation. The unit is metre. No normal distribution test is performed. The last three columns show the 5, 25, 75, and 95 percentiles, respectively. In the area of the sea shoreline data sources, the mean errors and the dispersion from the mean errors are small. The dispersion from the mean error is larger in the area of the 10 and 50-metre DEMs and larger errors are more likely to find there.

Data source	Nr of 20-metre cells	Mean	Median	Standard deviation	Minimum difference	Maximum difference	5%	25%	75%	95%
10-metre DEM	322,943	0.32	0.33	1.10	–11.10	9.24	–1.45	–0.21	0.89	2.02
50-metre DEM	972,088	0.35	0.22	1.45	–8.91	11.99	–1.71	–0.51	1.06	2.77
Sea shoreline from Property map	46,503	0.16	0.19	0.17	–2.02	0.85	–0.07	0.09	0.26	0.35
Digitized sea shoreline	2,685	–0.00	0.00	0.19	–0.82	3.92	–0.21	–0.01	0.00	0.17
Measured sea shoreline	819	–0.00	0.00	0.09	–0.74	0.37	–0.12	–0.00	0.00	0.10

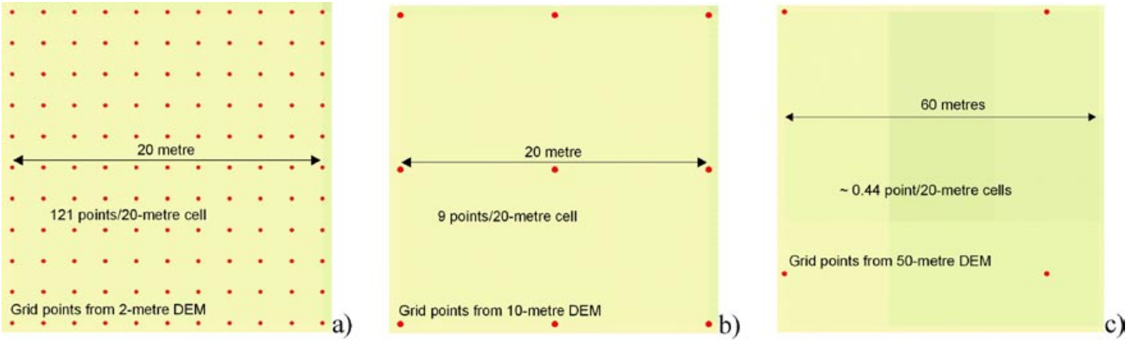

Figure 2-1. *Different point density in 20-metre cells for grid points from a 2-metre DEM a) a 10-metre DEM b), and a 50-metre DEM c).*

The high point density within the 0.25-metre DEM should have produced 20-metre cells of excellent quality within this area of the existing Oskarshamn DEM. Comparing the surface from the Oskarshamn DEM, within the 0.25-metre DEM extension, with surfaces produced using only 0.25-metre DEM grid points every 10th and 50th m (corresponding to grid points from the 10 and 50-metre DEMs), shows how point density affects quality of the 20-metre cells produced in the interpolation procedure. The result of such a comparison is presented in Table 2-4.

2.2 Calculation of errors in the sea area and lakes in the Forsmark 20-metre digital elevation model

Contrary to the land area there are no measurements of input data for the Forsmark DEM that can be used for error calculation in the sea area. Another big difference is the spatial distribution of data. In the land area most input data is distributed in regular grid points throughout the landscape. In the sea area, on the other hand, most input data is gathered along survey lines and large areas are not covered by any elevation data at all. These differences are the reason why another method for calculation of error in the sea area in the Forsmark DEM is used than for the land area.

Elevation data for sea areas has mostly been obtained from measurements, nautical charts, and other maps /Strömgren and Brydsten 2008/.

Some input data to the Forsmark DEM is removed in the sea area and used for calculation of error in other data sources. Measurements from survey lines performed by the Geological Survey of Sweden (SGU) /Elhammer and Sandkvist 2005/ and measurements from shallow bays /Brydsten and Strömgren 2004/ should be of comparable high quality and are therefore used for validation.

Table 2-4. Statistical analysis of different point density in input data for the interpolation procedure. The table shows the difference between the existing Oskarshamn DEM (100 points/20-metre cell in the input data for the interpolation procedure, within the 0.25-metre DEM), and two 20-metre DEMs produced using only points from the 0.25-metre DEM every 10th and 50th m (4 points and 0.36 points/20-metre cell, respectively) in the input data for the interpolation procedure. The calculations are done within the extension of the 0.25-metre DEM. The unit is metre. No normal distribution test is performed. The last three columns show the 5, 25, 75, and 95 percentiles, respectively. The statistics show that a high quality can be expected in a 20-metre DEM using a point density comparable to the 10-metre DEM. Using a 50-metre DEM, on the other hand, gives larger dispersion from the mean error and larger errors can also be expected.

Point density calculations	Nr of 20-metre cells	Mean	Median	Standard deviation	Minimum difference	Maximum difference	5%	25%	75%	95%
10-metre DEM	43,549	–0.00	0.00	0.05	–2.16	3.08	–0.00	–0.00	0.00	0.00
50-metre DEM	43,549	–0.07	–0.00	1.12	–7.53	7.72	–2.04	–0.65	0.56	1.68

However, it is actually not the elevation data that is used for the error calculation, but instead the 20-metre cells from the Forsmark DEM in areas of removed data. These cells are to a high extent produced from these removed data points, since they have had large weight in the Kriging interpolation procedure that is used for producing the DEM.

The measurements from the Geological Survey of Sweden are divided in a regional and detailed area. The distance between survey lines are approximately 1,000 m in the regional area. The distance between surveys lines in the detailed area and in the shallow bays are, on the other hand, only approximately 100 m. However, in some parts of the detailed area and shallow bays, survey lines cross other survey lines (Figure 2-2).

Some of the measurements from the crossing survey lines were removed from the original data base used for interpolation of the existing Forsmark DEM, and 20-metre cells within the areas of these removed data were used for error calculation of areas between survey lines. The distance between elevation data points in the survey lines from the Geological Survey of Sweden and the measurements from shallow bays is mostly only 2–3 m (i.e. 7–8 points/20-metre cells). This would suggest a high quality of the 20-metre cells used for error calculation of the sea area in the Forsmark DEM (see the calculation on point density described on the previous side).

The Geological Survey of Sweden has also interpreted elevation data /Elhammer and Sandkvist 2005/ in the detailed area, which is used in the interpolation of the Forsmark DEM. These interpreted elevation data was also removed from the original data base used for the interpolation procedure. Using the same Kriging parameters, 20-metre resolution, and extension as for the existing DEM /Strömgren and Brydsten 2008/, a new Forsmark DEM was produced from this modified data base.

No crossing survey lines exist in the regional area of the measurements from the Geological Survey of Sweden. Survey lines from the detailed area (Figure 2-3a) were removed (Figure 2-3b) from the data base used for the interpolation of the existing Forsmark DEM to imitate the regional area. All measurement points from the detailed area within 300-metre from these test zones were also removed to give an even better imitation of the condition in the regional area. From this modified data base, a second Forsmark DEM was produced.

Elevation data from the regional measurements performed by the Geological Survey of Sweden was used to calculate the error of base map data, and data from the digital nautical chart (the Swedish Maritime Administration) /Strömgren and Brydsten 2008/. Survey lines were removed from the input data to the existing Forsmark DEM within the areas of base map data, and digital nautical chart/nautical chart. Using the same Kriging parameters, 20-metre resolution, and extension as for the existing Forsmark DEM, a third DEM was produced from this modified data base.

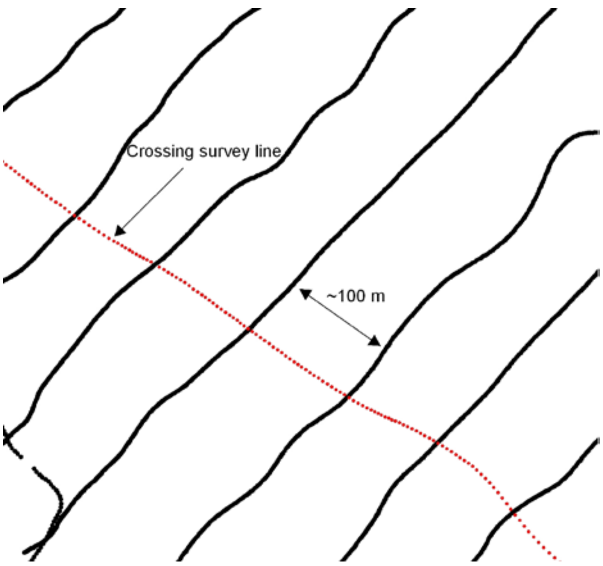

Figure 2-2. *Area where survey line crosses other survey lines.*

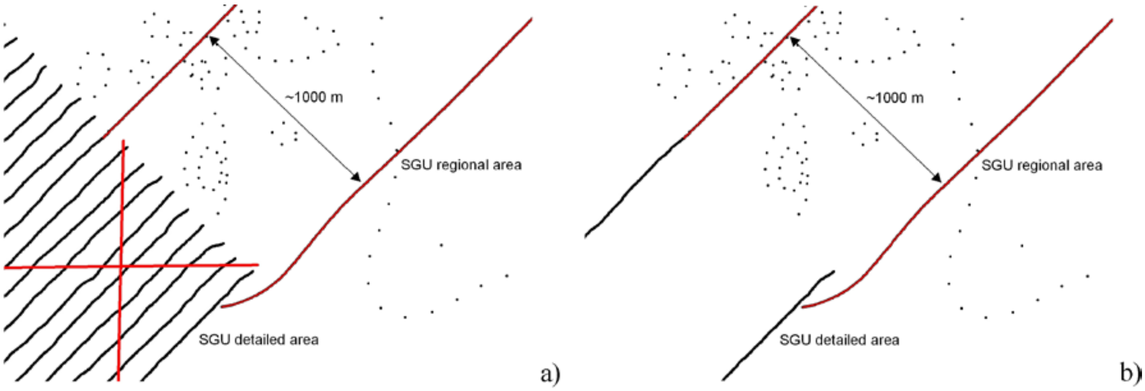

Figure 2-3. *Area where survey lines from the detailed area of the measurements performed by the Geological Survey of Sweden (SGU) a) were removed b) to imitate the condition in the regional area.*

The three new DEMs produced from modified input data bases were compared to the existing Formark DEM within the 20-metre cells covered by the removed data (Figure 2-4).

Descriptive statistics for the error calculation of these areas are shown in Table 2-5. It was not possible to calculate the error of the lakes in the Forsmark DEM in a similar way as for the data sources in the sea area since no proper data for validation exists. No error calculation was done for the inlet channel to the nuclear power plant and areas where false depth values are placed either.

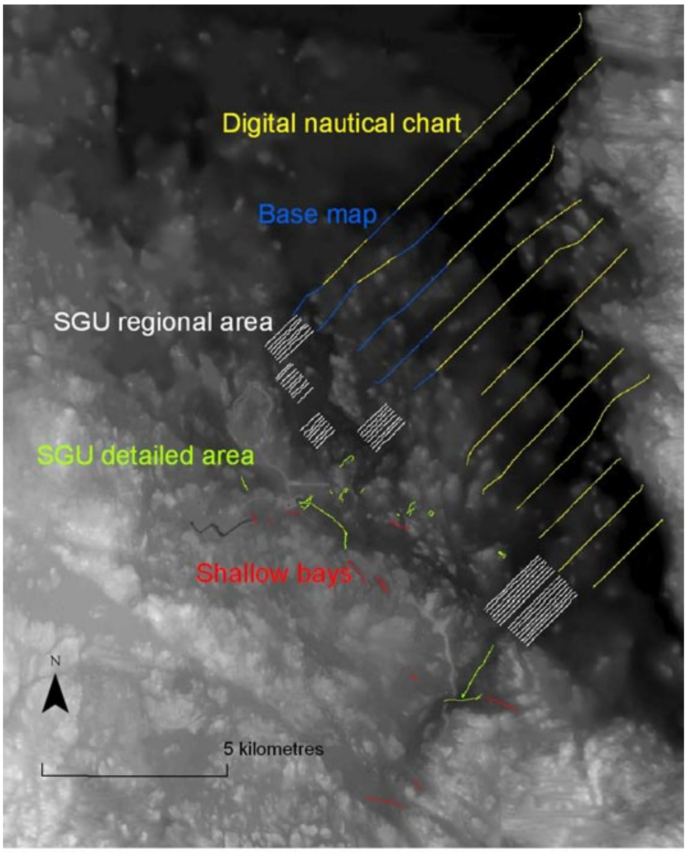

Figure 2-4. *20-metre cells used for calculation of errors for the data sources shown in the figure, placed on top of part of the existing Forsmark 20-metre digital elevation model (DEM).The errors for the regional area of the measurements from the Geological Survey of Sweden (SGU regional area) are not calculated from overlapping cells within this area, but instead from the from the detailed area of the measurements from the Geological Survey of Sweden.*

Table 2-5. Statistical analysis of the error in the existing Forsmark DEM in 20-metre cells covered by different input data sources for the existing Forsmark DEM (see also Figure 2-4). The unit is metre. No normal distribution test is performed. The last three columns show the 5, 25, 75, and 95 percentiles, respectively. In the shallow bays and in the detailed area of the measurements from the Geological Survey of Sweden (SGU), the mean error and dispersion from the mean error are small. In the area of base map data, the mean error and dispersion from the mean error are slightly larger. In the regional area of the measurements from the Geological Survey of Sweden (SGU), the dispersion from the mean error is even larger, and it is also likely to find larger errors there. In areas with only data from the digital nautical chart/nautical chart, the mean error and dispersion from the mean error is very large, and a maximum error larger than 35 m has been calculated there.

Data source extension	Nr of 20-metre cells	Mean	Median	Standard deviation	Minimum error	Maximum error	5%	25%	75%	95%
SGU regional area	3,899	0.25	0.46	2.19	–9.89	7.31	–4.03	–0.72	1.48	3.36
SGU detailed area	502	–0.28	–0.16	0.70	–4.03	1.93	–1.74	–0.47	0.02	0.69
Shallow bays	265	0.06	0.07	0.47	–3.00	1.45	–0.63	–0.12	0.27	0.81
Digital nautical chart/ nautical chart	4,184	4.84	3.09	5.52	–5.44	35.55	–0.87	1.44	6.54	16.45
Base map	824	–0.69	–0.57	1.02	–4.48	3.13	–2.55	–1.25	–0.09	0.82

3 Results and conclusion

3.1 Errors in the Forsmark digital elevation model

Data sources for the 20-metre digital elevation model (DEM) for the Forsmark area are shown in Figure 3-1, and the calculated errors in the Forsmark DEM for areas within these data sources are shown in Table 3-1.

The data sources for the sea shoreline have low mean errors, between 0.00 and 0.16 m a standard deviation ranging from 0.09–0.19 m, and maximum and minimum errors not larger than a few metres. The mean error in the SKB 10-metre DEM is 0.32 m with a standard deviation of 1.10 m, and the uncertainty in the 50-metre DEM is only slightly larger with a mean error of 0.35 m and a standard deviation of 1.45 m. However, larger error are likely to be found in both the areas of the 10 and 50-metre DEMs since the maximum and minimum errors for these areas are more than 10 m. The error calculated in the point density calculation presented in Table 2-4 would only have a minor effect on the mean error and the spread from the mean error in the area of the 10-metre DEM (a point density of 4 points/20-metre cell). However, the result from this test suggest an increasing spread from the mean error within the 50-metre DEM (0.36 points/20-metre cell) if the effect of point density is considered. The distance between points for the sea shoreline data from the property map, and the digitized sea shoreline are 5 m (4 points/20-metre cell), and approximately 2 m for areas measured using DGPS (approximately 10 points/20-metre cell). This indicates that error due to the density of input data points for the interpolation should be small for areas within these data sources.

The statistics for the errors in the Forsmark DEM in the sea area should be interpreted different than for land areas. For land areas, the errors are calculated for the whole area of the input data sources. For the sea area, on the other hand, the errors are only calculated for areas between the survey lines or in areas where no input data exists.

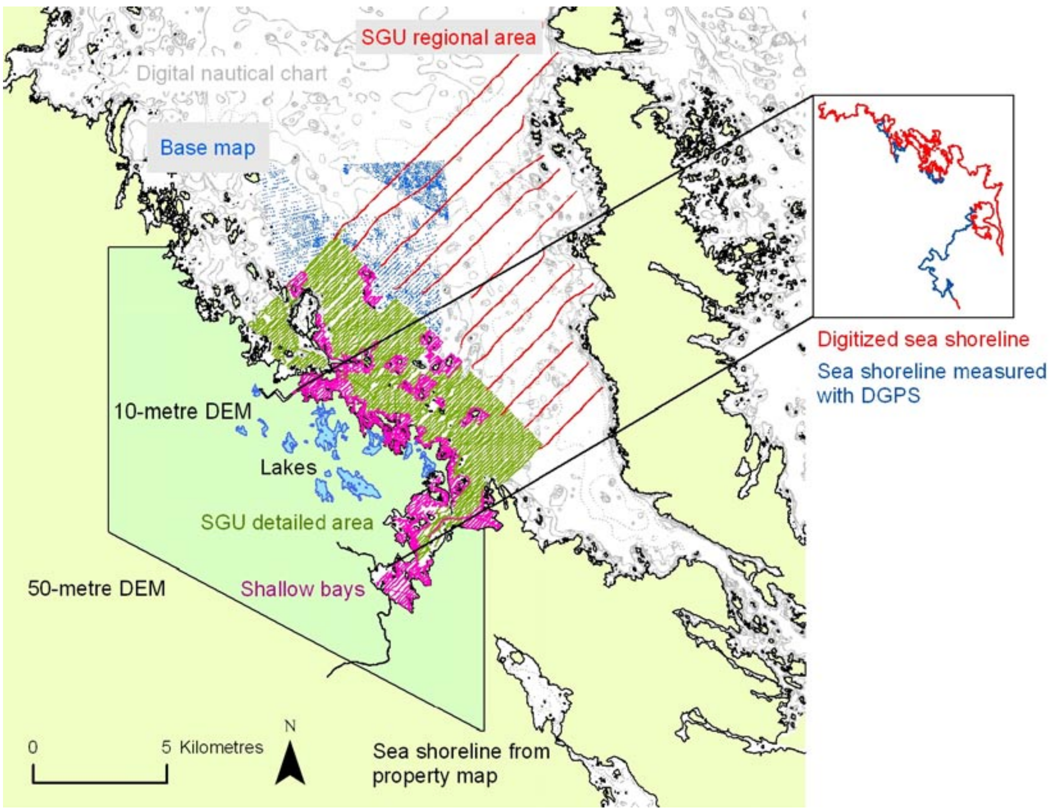

Figure 3-1. *Input data for the existing 20-metre digital elevation model for the Forsmark area.*

Table 3-1. The table is a summary of Table 2-4 and 2-5 and shows a statistical analysis of errors for different areas in the existing digital elevation model (DEM) for the Forsmark area. The unit is metre. No normal distribution test is performed. The last three columns show the 5, 25, 75, and 95 percentiles, respectively. The statistics for the 10 and 50-metre DEMs, the sea shoreline from the property map, the digitized sea shoreline, and the sea shoreline measured with DGPS is calculated for the whole area of these data sources. The statistics for the rest of the data is calculated for areas between the survey lines or for areas were data is missing. In the area of the sea shoreline data sources, the mean errors and the dispersion from the mean errors are small. In the areas of the 10 and 50-metre DEMs, the dispersion from the mean errors is larger. In the shallow bays and in the detailed area of the measurements from the Geological Survey of Sweden (SGU), the mean error and dispersion from the mean error are small. In the area of base map data, the mean error and dispersion from the mean error are larger. In the regional area of the measurements from the Geological Survey of Sweden (SGU), the dispersion from the mean error is even larger, and larger errors are likely to be found there. In areas with only data from the digital nautical chart/nautical chart, the mean error and dispersion from the mean error is very large. In this area error larger than 35 m has been calculated.

Data source	Nr of observations	Mean	Median	Standard deviation	Minimum error	Maximum error	5%	25%	75%	95%
10-metre DEM	322,943	0.32	0.33	1.10	–11.10	9.24	–1.45	–0.21	0.89	2.02
50-metre DEM	972,088	0.35	0.22	1.45	–8.91	11.99	–1.71	–0.51	1.06	2.77
Sea shoreline from Property map	46,503	0.16	0.19	0.17	–2.02	0.85	–0.07	0.09	0.26	0.35
Digitized sea shoreline	2,685	–0.00	0.00	0.19	–0.82	3.92	–0.21	–0.01	0.00	0.17
Measured sea shoreline	819	–0.00	0.00	0.09	–0.74	0.37	–0.12	–0.00	0.00	0.10
SGU regional area	3,899	0.25	0.46	2.19	–9.89	7.31	–4.03	–0.72	1.48	3.36
SGU detailed area	502	–0.28	–0.16	0.70	–4.03	1.93	–1.74	–0.47	0.02	0.69
Shallow bays	265	0.06	0.07	0.47	–3.00	1.45	–0.63	–0.12	0.27	0.81
Digital nautical chart/ nautical chart	4,184	4.84	3.09	5.52	–5.44	35.55	–0.87	1.44	6.54	16.45
Base map	824	–0.69	–0.57	1.02	–4.48	3.13	–2.55	–1.25	–0.09	0.82

In the shallow bays, the mean error is only 0.06 m and the standard deviation 0.47 m. In the detailed and regional areas of the measurements from the Geological Survey of Sweden (SGU), the mean errors are –0.28 m and 0.25 m, and the standard deviations 0.7 and 2.2 m, respectively. The maximum and minimum errors in the shallow bays and the detailed area are ranging from almost 2 m to a little more than –4 m. For the regional area, these errors are 7.31 to –9.89 m, indicating that larger errors are more likely to be found there. The mean error and the standard deviation are quite small in the area of the base map. In the large area within elevation data from the digital nautical chart/nautical chart, the uncertainty of the quality in the Forsmark DEM must be considered high. The statistics for this area shows a mean error as high as 4.84 m, a standard deviation of 3.09 m, a minimum error of –5.44 m, and a maximum error of more than 35 m.

No calculation of the error in the lakes in the Forsmark DEM was possible to do since no proper data for validation exists. However, the distance between measurement points in lakes could give an indication of the expected errors. For most lakes, the maximum distance between measurement points is around 30–40 m. For the three largest lakes: Bolundsfjärden, Eckarfjärden, and Fiskarfjärden, the maximum distance between measurement points is between 70–100 m. The distance between points in the larger lakes are about the same as in the survey lines from shallow bays and the detailed area of the measurements from the Geological Survey of Sweden. Altogether, this would suggest the error for lakes in the Forsmark DEM to be at least as low as within the shallow bays. No error calculation was done for the inlet channel to the nuclear power plant and the areas where false depth values are placed.

Figure 3-2 show areas for the data sources for the Forsmark DEM ranked according to the percentage of the mean error ± the standard deviation within an interval of ± 0.5 m, which both take a low mean error and a small dispersion in consideration. The four highest ranked areas (the data sources for the sea shoreline and the shallow bays) all have 100 percent of the interval ± the standard deviation within the ± 0.5 m interval. The internal ranking of these areas is therefore based on the percentage of the mean error ± the standard deviation within a ± 0.1 m interval. The detailed area of the measurements from the Geological Survey of Sweden comes next in the ranking order followed by the areas for the base map, the 10-metre DEM, and the 50-metre DEM. The regional area of the measurement from the Geological Survey of Sweden and the area of the digital nautical chart/nautical chart are of lowest quality according to this calculation.

Further studies

The knowledge from this analysis can for example be used assessing the size of errors for the catchments in the Forsmark area. Using the hydrological model in ArcGis 9.2, catchments can be produced from the existing Forsmark DEM and from a DEM where the error distribution calculated for the existing Forsmark DEM is randomly placed. Comparing the catchments produced from different DEMs should give a measure of the size of the errors the in catchment caused by error in input data to the existing Forsmark DEM.

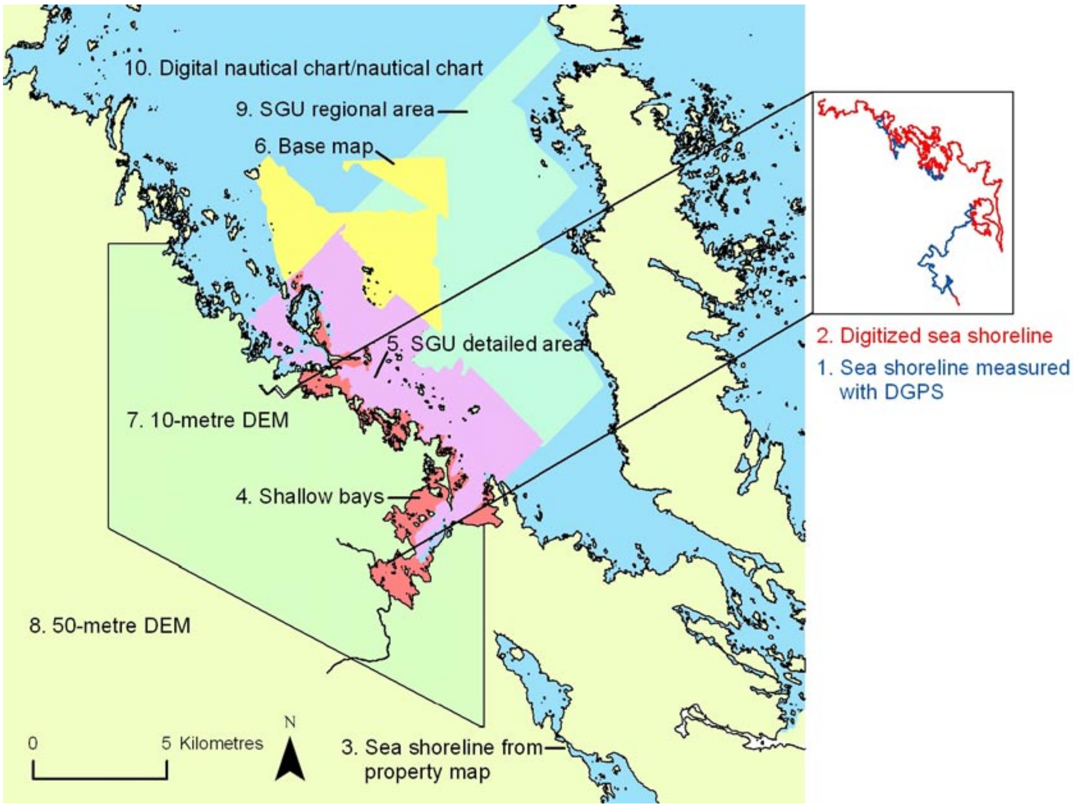

Figure 3-2. *Ranking order of quality for areas of data sources in the existing Forsmark DEM (numbered 1–8). The areas within the data sources for the sea shoreline and shallow bays are of highest quality. The detailed area of the measurements from the Geological Survey of Sweden comes next in ranking order followed by the areas for the base map, the 10-metre DEM, and the 50-metre DEM. The regional area of the measurements from the Geological Survey of Sweden and the area of the digital chart/nautical chart are of lowest quality.*

4 References

Brunberg A-K, Carlsson T, Blomqvist P, Brydsten L, Strömgren M, 2004. Identification of catchments, lake-related drainage parameters and lake habitats. SKB P-04-25, Svensk Kärnbränslehantering AB.

Brydsten L, Strömgren M, 2004. Water depth soundings in shallow bays in Forsmark. SKB P-04-125, Svensk Kärnbränslehantering AB

Brydsten L, Strömgren M, 2005. Measurement of brook gradients and lake threshold. SKB P-04-141, Svensk Kärnbränslehantering AB.

Elhammer A, Sandkvist Å, 2005. Detailed marine geological survey of the sea bottom outside Forsmark. SKB P-03-101, Svensk Kärnbränslehantering AB.

Hedenström A, Sohlenius G, Strömgren M, Brydsten L, 2008. Depth and stratigraphy of regolith at Forsmark. SKB R-08-07, Svensk Kärnbränslehantering AB.

Nyborg M, 2005. Aerial photography and airborne laser scanning Laxemar-Simpevarp. The 2005 campaign. SKB P-05-223, Svensk Kärnbränslehantering AB.

Strömgren M, Brydsten L, 2008. Digital elevation models of Forsmark. Site descriptive modelling. SDM-Site Forsmark. SKB R-08-62, Svensk Kärnbränslehantering AB.

Strömgren M, Lindgren F, 2010. Total station measurements of DEMs in the Laxemar-Simpevarp area, in prep. Svensk Kärnbränslehantering AB.

Wiklund S, 2002. Digitala ortofoton och höjdmodeller. Redovisning av metodik för platsundersökningsområdena Oskarshamn och Forsmark samt förstudieområdet Tierp Norra. SKB P-02-02, Svensk Kärnbränslehantering AB.