R-12-03

Digital elevation model of Forsmark

Site-descriptive modelling

SR-PSU biosphere

Mårten Strömgren, Lars Brydsten
Umeå University

December 2013

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 R-12-03
ID 1404623

Digital elevation model of Forsmark

Site-descriptive modelling

SR-PSU biosphere

Mårten Strömgren, Lars Brydsten
Umeå University

December 2013

Keywords: Digital elevation model, DEM, Topography, GIS, Forsmark, Surface, Biosphere.

Abstract

A digital elevation model (DEM) describes the terrain relief. An accurate DEM is important for further modelling of the Forsmark region. The aim of this project was to improve the DEM for the Forsmark area using new elevation data and develop the method used for the construction of the previous DEM.

Data have been replaced in part of the terrestrial area by airborne laser scanned data. In the sea, most of the data used for the construction of the previous DEM have been replaced with new and reinterpreted data. All data used for the interpolation were combined into a dataset of approximately 1.9 million points unevenly spread over an area of about 900 km^2.

The software ArcGis 9.3.1 Geostatistical Analysis extension was used for the interpolation of data points. The interpolation was done in two domains demarcated by the sea shoreline. The interpolation method used was Ordinary Kriging. In the terrestrial domain an exact interpolation was chosen and in the marine domain a smooth interpolation was chosen. Both of these choices allow a cross validation and a validation before the interpolation is conducted. Cross validation with different Kriging parameters were performed and the model with the most reasonable statistics was chosen. Finally, a validation with the most appropriate Kriging parameters was performed in order to verify that the model fits unmeasured localities. The map projection used in the elevation model is RT 90 2.5 Gon W and the height system is RH 70. The DEM has a cell size of 20×20 m.

An analysis of the elevation model confirms existing knowledge of the area being extremely flat. The range in elevation is only approximately 106 metres with the highest point at 50.1 metres above sea level in the south-west part of the DEM, and the deepest sea point at –55.4 metres in the northern part of the so-called Gräsörännan.

The surface of the current DEM is smoother than for the previous DEM. The highest accuracy of the DEM is reached in the area covered with data from the airborne laser scanning and in the 10 km^2 large area where the Geological Survey of Sweden (SGU) have performed a very detailed survey. Only small errors are found in the shallow bays, within the extent of SKBs 10 m DEM and in SGUs detailed survey area. Errors larger than a few metres are found in the area covered with data from the 50 m DEM from the National Land Survey of Sweden (LMV) and the area covered with data from the digital nautical chart and SGUs measurements above 20 m depth.

The accuracy is lowest in the areas covered only with data from the digital nautical chart and data from SGU below 20 m depth. Errors of some metres size are quite usual in these areas and errors larger than 10 m can also be expected in both these areas.

Sammanfattning

En digital höjdmodell (DEM) beskriver reliefen i terrängen. En bra DEM är viktig för fortsatt modellering av Forsmarksområdet. Syftet med detta projekt var att förbättra den digitala höjdmodellen över Forsmarksområdet genom att använda ny höjddata och utveckla metoden som användes vid framställningen av tidigare DEM.

Data har i en del av det terrestra området ersatts med laserscannad data. I havet har största delen av den data som användes vid framställningen av föregående DEM ersatts med ny och nytolkad data. All data som användes vid interpoleringen slogs samman till ett dataset med sammanlagt cirka 1,9 miljoner punkter ojämnt spridda över ett cirka 900 km^2 stort område.

Interpolering mellan olika datapunkter utfördes i programmet ArcGis 9.3.1 Geostatistical Analysis extension. Interpoleringen utfördes i två olika domäner, avgränsade av kustlinjen. Ordinary Kriging användes som interpoleringsmetod. I den terrestra domänen valdes en exakt interpolering och i den marina domänen en mjuk interpolering. Båda dessa val ger möjlighet till att göra en korsvalidering och en validering av höjdmodellen innan interpoleringen genomförs. Korsvalideringar med olika Krigingparametrar utfördes och modellen med den mest rimliga statistiken valdes. Slutligen gjordes en validering med de mest lämpade parametrarna för att verifiera att modellen passar även där det inte finns några mätpunkter. Höjdmodellen har koordinatsystemet RT 90 2.5 Gon W och höjdsystemet RH 70 och har en cellstorlek om 20×20 meter.

En analys av höjdmodellen bekräftar vetskapen om att området är mycket flackt. Det totala höjdintervallet är enbart ungefär 106 m med den högsta punkten 50.1 meter över havsnivån i modellens sydvästra del och den lägsta punkten –55.4 m under havsnivån i Gräsörännans norra del.

Den här höjdmodellen har en mjukare yta än tidigare höjdmodell. Den högsta noggrannheten i höjdmodellen finns i området där bara laserscannad data har använts och i det 10 km^2 stora område där Sveriges geologiska undersökning (SGU) har gjort en mycket detaljerade undersökning. Enbart mindre fel finns i de grunda havsvikarna, inom utsträckningen för SKBs 10 m höjdmodell och i SGUs detaljerade mätområde. Fel större än några meter finns i området där endast data från Lantmäteriets 50 m höjdmodell och data från det digitala sjökortet i områden grundare än 20 m har använts.

Noggrannheten är lägst där enbart data från det digitala sjökortet och den del av SGUs regionala mätområde där det är 20 m och djupare har använts. Några meter stora fel är ganska vanligt förekommande i dessa områden och fel större än 10 m kan också förväntas.

Contents

1 Introduction

In order to take care of the low and intermediate waste generated during dismantling of closed nuclear facilities in Sweden an extension of the existing facility for low and intermediated radioactive waste, SFR (Slutförvaret för kortlivat radioaktivt avfall), is planned. The SFR facility is situated at the coast of the Baltic Sea in the vicinity of the Forsmark power plant. A proper digital elevation model (DEM), describing the terrain relief, is important as input for further modelling within the safety assessment, SR-PSU, for an extension of SFR. Such models include, e.g. hydrology, Quaternary deposits and land use. Since the previous DEM for the Forsmark area was constructed new data have been produced and some of the existing data have been reanalysed. The aim of this project was to improve the DEM for the Forsmark area using this new elevation data and develop the method used for the construction of the previous DEM.

DEM resolution is the size of DEM cells. A DEM is constructed by an interpolation of irregular spaced elevation data. In this model, Kriging interpolation was used. Kriging is a geostatistical interpolation method based on statistical models that include autocorrelation (a statistical relationship). 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.

Normally, a DEM has a constant value for sea surface and constant values for lake surfaces. For the Forsmark area, the DEM has negative values in the sea to represent water depth, but constant positive values for land or varying values to represent lake bottom elevations. Input data for the interpolation have many different sources, such as existing digital elevation models, airborne laser scanning measurements, elevation lines from digital topographical maps, digital nautical charts, and depth soundings in both lakes and the sea. The quality of these data sources is also different. The density and distribution of input data are different for the terrestrial and marine areas; most of the data in the terrestrial area are distributed in regular grids throughout the landscape. In the marine area, on the other hand, most elevation data are distributed densely along survey lines and there are no elevation data at all in large areas between the survey lines. The Kriging interpolation was performed using the ArcGis 9.3.1 Geostatistical Analysis extension.

2 Method

2.1 Data used for the construction of the DEM

Part of the data used for the construction of the previous DEM (Strömgren and Brydsten 2008) were also used for the construction of the current DEM. However, new data of higher quality have become available and some of the existing data have been reanalysed. The method and necessary calculations for creating the dataset used in the previous DEM are thoroughly described in Strömgren and Brydsten (2008). Therefore, this report mainly focuses on describing the new data used in the interpolation procedure and the method used for improvement of the DEM.

2.1.1 Data for the terrestrial area

Data covering the largest terrestrial areas are shown in Figure (2-1). Data from the terrestrial areas used in the previous DEM were collected from six different sources: the existing DEM from the National Land Survey of Sweden (LMV) with a resolution of 50 m, SKBs 10 m DEM (Wiklund 2002), measured values from lakes (Brunberg et al. 2004), fixed points from the digital localities map, and brook measurements (Brydsten and Strömgren 2005).

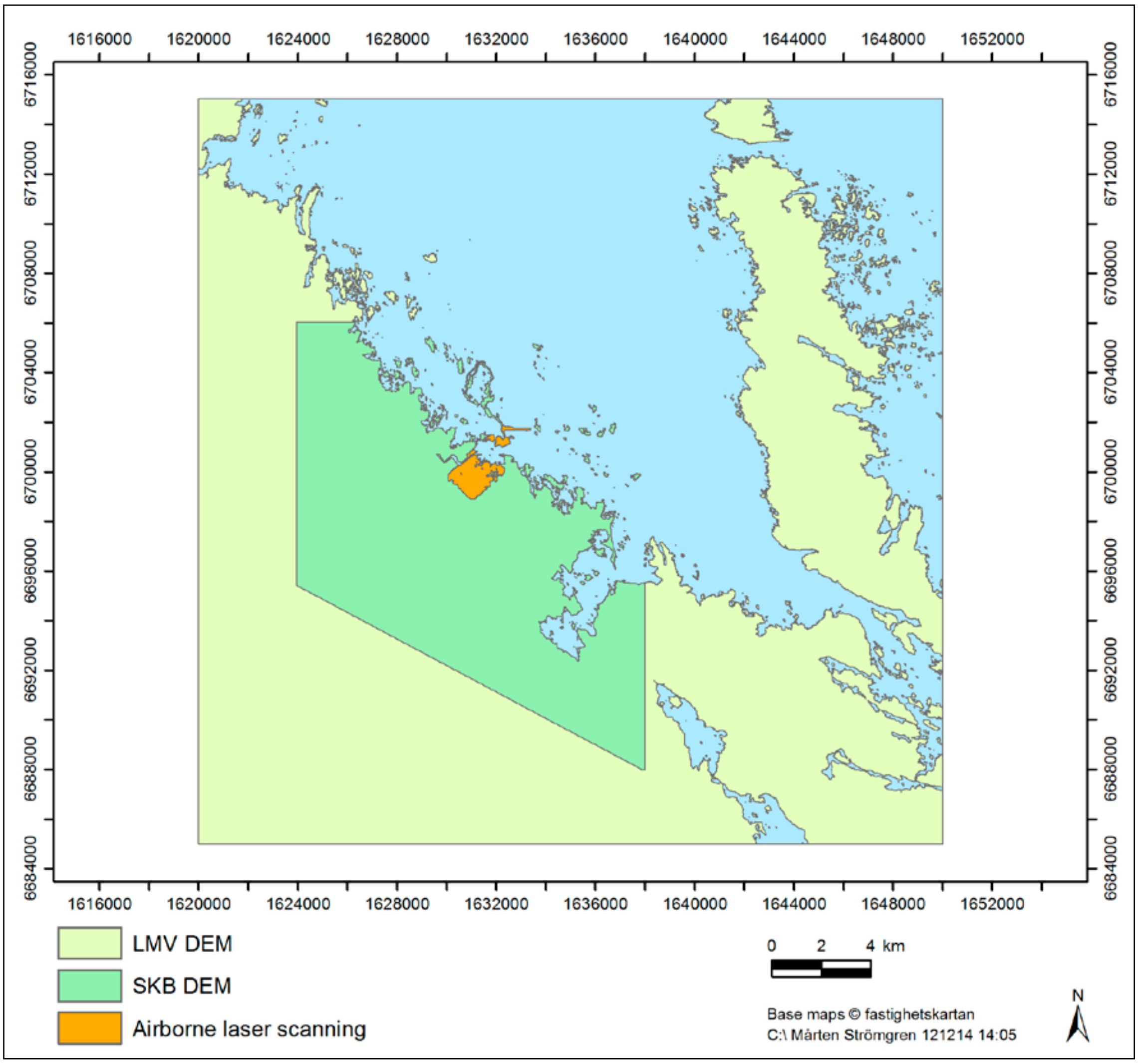

Figure 2-1. *Extensions of data sources covering the largest areas in the terrestrial part of the DEM.*

Data from airborne laser scanning performed in the Forsmark area (SKB GIS databas, C201 SDEADM.SKB_FM_HOJ_7882[1]) have been used to replace data from SKBs 10 m DEM in two small areas close to the Forsmark power plants. More than 800,000 points were produced from this measurement. These data were recalculated to mean values in 5-metre cells (Figure 2-2).

All data available in the terrestrial area for the construction of the current DEM are shown in Table 2-1.

Table 2-2 in Strömgren and Brydsten (2008) shows the lake surface levels used for conversion of depth measurements to the height system RH 70. However, during the construction of the current DEM, it has been pointed out that there are some errors regarding lake surface levels for six of the lakes shown in Strömgren and Brydsten (2008). Table 2-2, in this report, shows the correct lake surface levels for these six lakes. These levels were also used for the construction of the previous DEM, no matter the table in Strömgren and Brydsten (2008) shows different values.

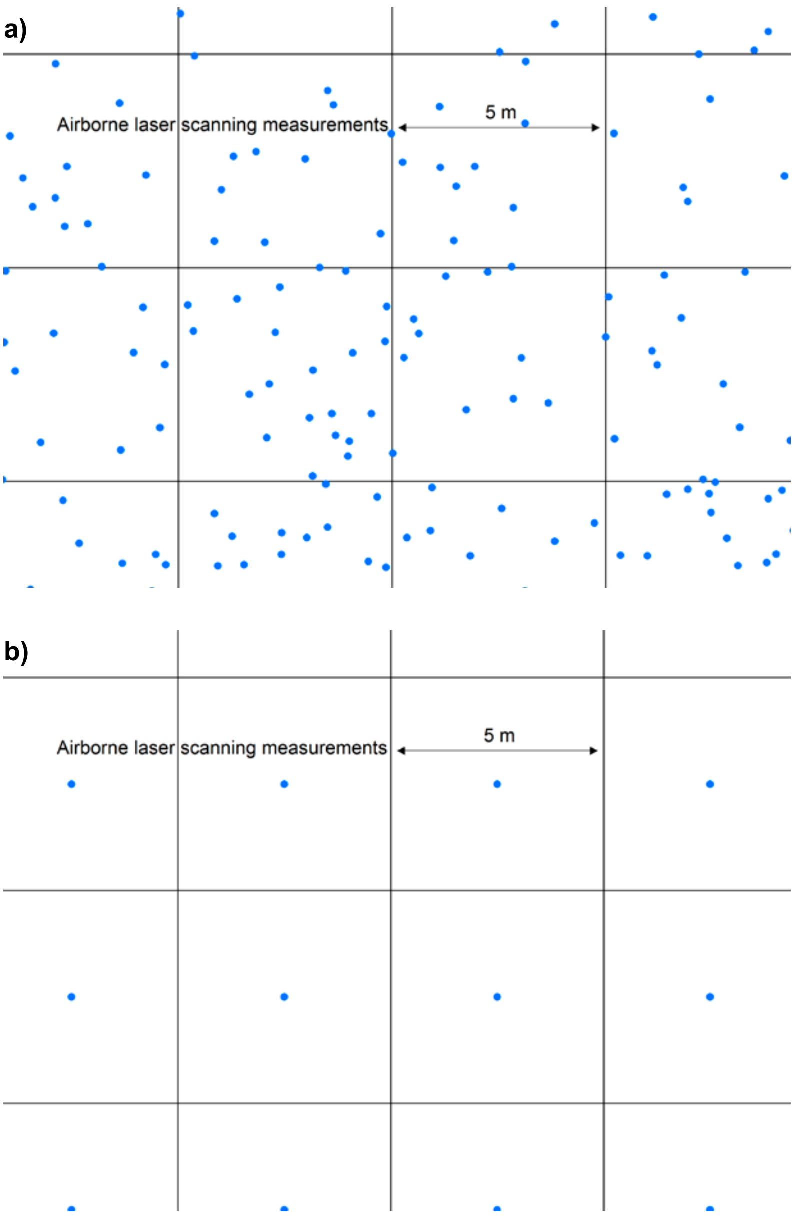

Figure 2-2. *a) Airborne laser scanning measurements from the Forsmark area (SKB GIS databas, C201 SDEADM.SKB_FM_HOJ_7882) and b) Mean values for airborne laser scanning measurements calculated for measurements within 5 m cells.*

[1] Data may be made available upon request.

Table 2-1. Available data for the terrestrial area for the construction of the DEM.

Data	No of points	Reference
SKBs 10 m DEM	1,289,049	(Wiklund 2002)
LMVs 50 m DEM	150,027	(Swedish national land survey, LMV)
Airborne laser scanning	93,686	(SKB GIS databas, C201 SDEADM.SKB_FM_HOJ_7882)
Measured values from lakes	21,293	(Brunberg et al. 2004)
Brook measurements	634	(Brydsten and Strömgren 2005)
Fixed points	68	(Property map)

Table 2-2. Lake surface levels (metre above RH 70) for six lakes where values shown in Strömgren and Brydsten (2008) (right column) differs from values used for the construction of the current and the previous DEM for the Forsmark area (left column). The values in the left column are also the correct values.

Lake	Lake surface levels used for the construction of the current and previous DEM	Lake surface levels shown in Strömgren and Brydsten (2008)
AFM000080	0.02	0.06
AFM000085	0.05	0.365
AFM000087	0.43	0.675
AFM000089	1.14	1.19
AFM000091	0.48	0.73
AFM000094	1.82	2.82

2.1.2 Data for the marine area

Most of the data for the marine area used for the construction of the previous DEM have been replaced by new and reanalysed data. Some of these data are situated outside the extension of the DEM, north of the DEM. However, in Figure 2-3 the extensions for the sources contributing with most elevation data for the marine area are only shown within the extension of the DEM. The data for the marine area used for the construction of the DEM have been obtained from the following sources:

1. the digital nautical chart (the Swedish Maritime Administration), area A in Figure 2-3,
2. the paper nautical chart (number 535 Öregrund – Grundkallen – Björn), area A in Figure 2-3,
3. the Geological Survey of Sweden, SGU, (Nyberg et al. 2011), area A (regional survey area), B (detailed survey area), and C (very detailed survey in 10 km^2 large area) in Figure 2-3,
4. SGUs interpreted depth data (Elhammer and Sandkvist 2005), area B in Figure 2-3,
5. depth soundings of shallow bays (Brydsten and Strömgren 2004), area D in Figure 2-3,
6. with DGPS measured shoreline points,
7. digitized shoreline points from IR orthophotos,
8. the sea shoreline from the digital localities maps from Lantmäteriet,
9. constructional drawings for the inlet channel to the nuclear power plant (Vattenfall 1977), area F in Figure 2-3,
10. depth chart for the Biotest lake from 1976 (SKB), area G in Figure 2-3, and
11. supporting depth data within shallow bays, area D and F in Figure 2-3.

A very detailed marine geological survey (Nyberg et al. 2011) was conducted in 2010 in a 10 km^2 large area outside Forsmark (area C in Figure 2-3), 100 m spacing between survey lines. Data from this area were delivered in ESRI Grid format, 1 m cell size. Nyberg et al. (2011) also reanalysed survey data retrieved in 2002 from the same area, and reanalysed survey lines retrieved in 2002, 2008 and 2009 during SGUs regular mapping program from a larger, adjacent area including Öregrundgrepen (area A in Figure 2-3).

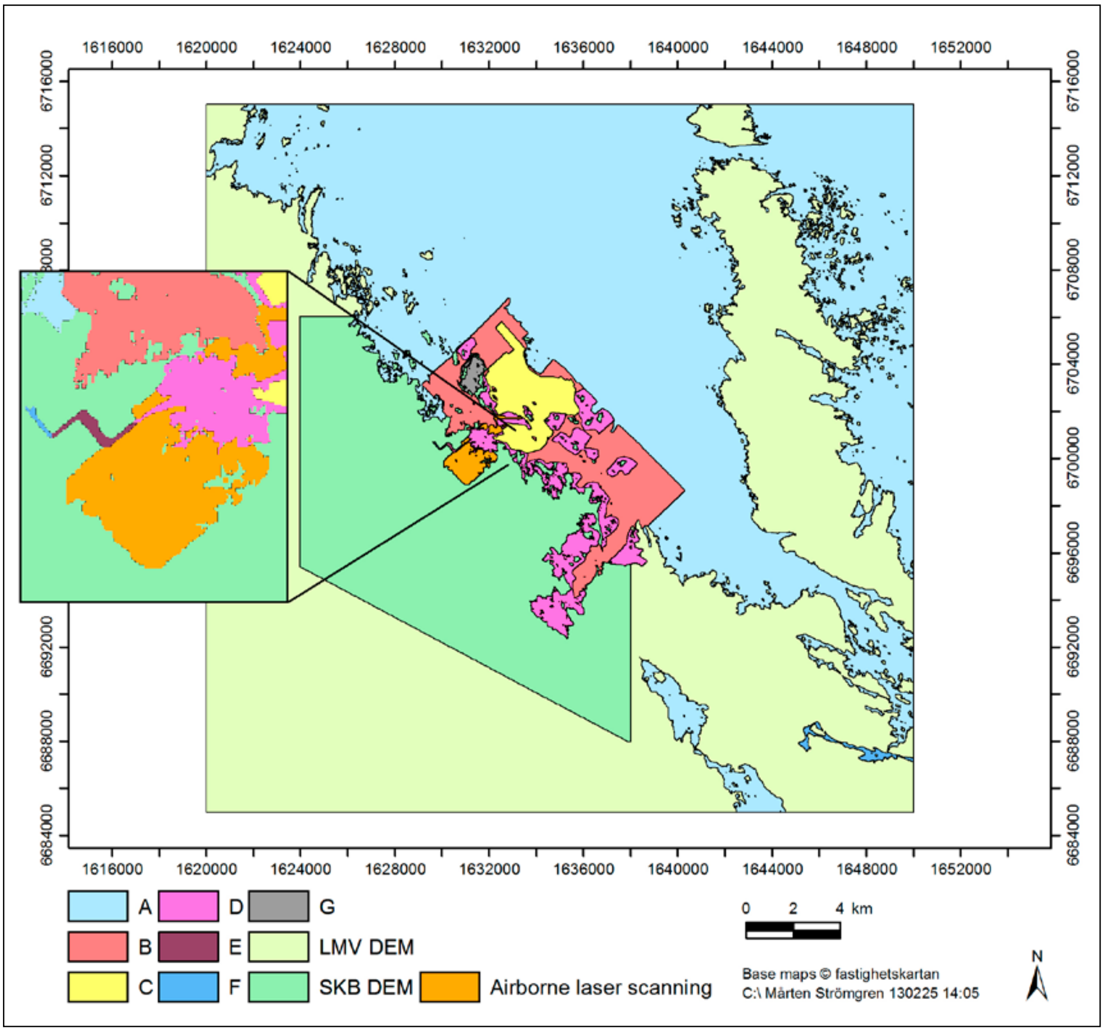

Figure 2-3. *Extensions of different data sources for the marine area in Forsmark. A = depth soundings performed by SGU and data from the digital nautical chart and the nautical chart, B = depth soundings and interpreted depth data from SGU, C = depth soundings performed by SGU (very detailed survey), D = depth soundings of shallow bays, E = constructional drawings for the inlet channel to the nuclear power plant (enlarged area), F = supporting depth data (enlarged area), and G = depth chart for the Biotest lake.*

Data from this area used in the previous and current DEM are shown in Figure 2-4.

The numbers of depth values produced from the very detailed marine geological survey (Nyberg et al. 2011) (area C in Figure 2-3) are very large, more than 10 million depth values. All these depth values would not have been possible to use in the interpolation procedure. In order to reduce the number of depth values, but still use all, the same technique already used for the data from the laser scanning measurements (described on page 7) was also used here, i. e. the original measurements were recalculated to mean values for a cell with 5 m resolution. These 5 m cells were converted to points and added to the dataset used for the construction of the DEM.

The new and reanalysed data from SGU (Nyberg et al. 2011) were delivered to SKB as 52 files in SGU-format (text files), generally one file for each survey line. The columns in the files consist of X-coordinates with a resolution of 2 digits and Y-coordinates with a resolution of 3 digits and a Z-value with a resolution of 3 digits. The coordinate system was RT 90 and the height system was RH 70. These text files were imported to Excel and each text file was saved as an Excel document. These Excel documents were opened in ArcGis and exported to ESRI Shape-format. All depth data were merged to one shapefile for the detailed survey and the regional survey area, respectively (area B and A in Figure 2-3).

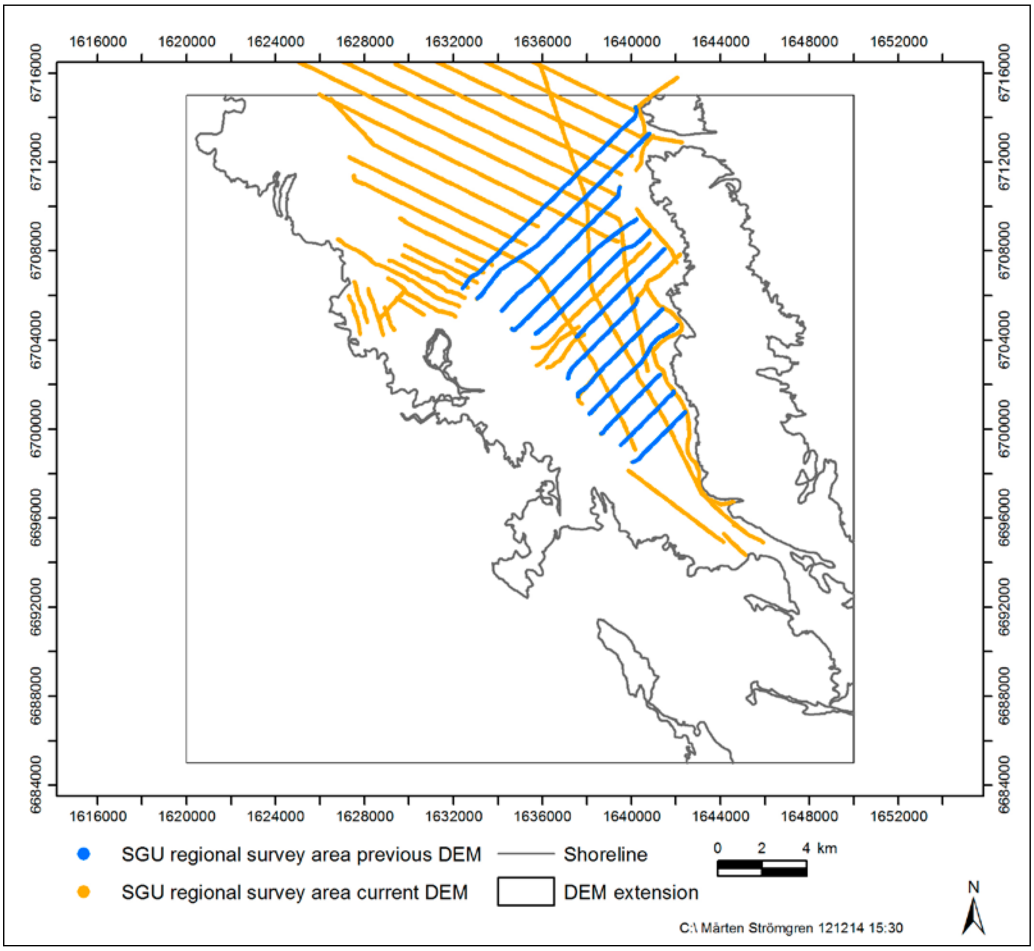

Figure 2-4. *Data from the blue survey lines were used in the previous DEM (Strömgren and Brydsten 2008) and data from the orange and blue survey lines were used in the current DEM.*

The new and reanalysed data from SGU (Nyberg et al. 2011) were compared with data from the base map to the nautical chart, and data from the digital nautical chart in an area where these different data sources overlapped. Data from the base map to the nautical chart were used in the previous DEM (Strömgren and Brydsten 2008). The water depth values from the base map to the nautical chart differed from the other two data sources. After some test interpolations it was decided that depth values from the base map had to be removed from the dataset used for the construction of the DEM.

In the dataset from the digital nautical chart used in the previous DEM, the distance between points showing water depth values along depth lines is 50 m for most depth lines. However, during the construction of this DEM, it was noticed that for some depth lines the distance between points showing water depth values is approximately 150 m. Data from the digital nautical chart were therefore supplemented with point depth values so that the distance between points for all depth lines was around 50 m. To the dataset was also single depth values from the digital nautical chart not used for the construction of the previous DEM added. In total 4,443 points showing the water depth were added compared to the dataset used in the previous DEM (Figure 2-5).

The depth values in the digital nautical chart refer to mean sea level 1970, so no adjustment was needed for mixing soundings and land elevation data in RH 70.

Depth lines were digitized from a depth chart from 1976 for the Biotest Lake in Forsmark obtained by personal communication with Peter Karås SLU, also stored at SKB_SVN\SFR\SR-PSU\Landscape\Indata[2] (area G in Figure 2-3). These depth lines were converted to points with 10 m distance between the points. These points were added to the dataset used for the construction of the DEM.

[2] Data may be made available upon request.

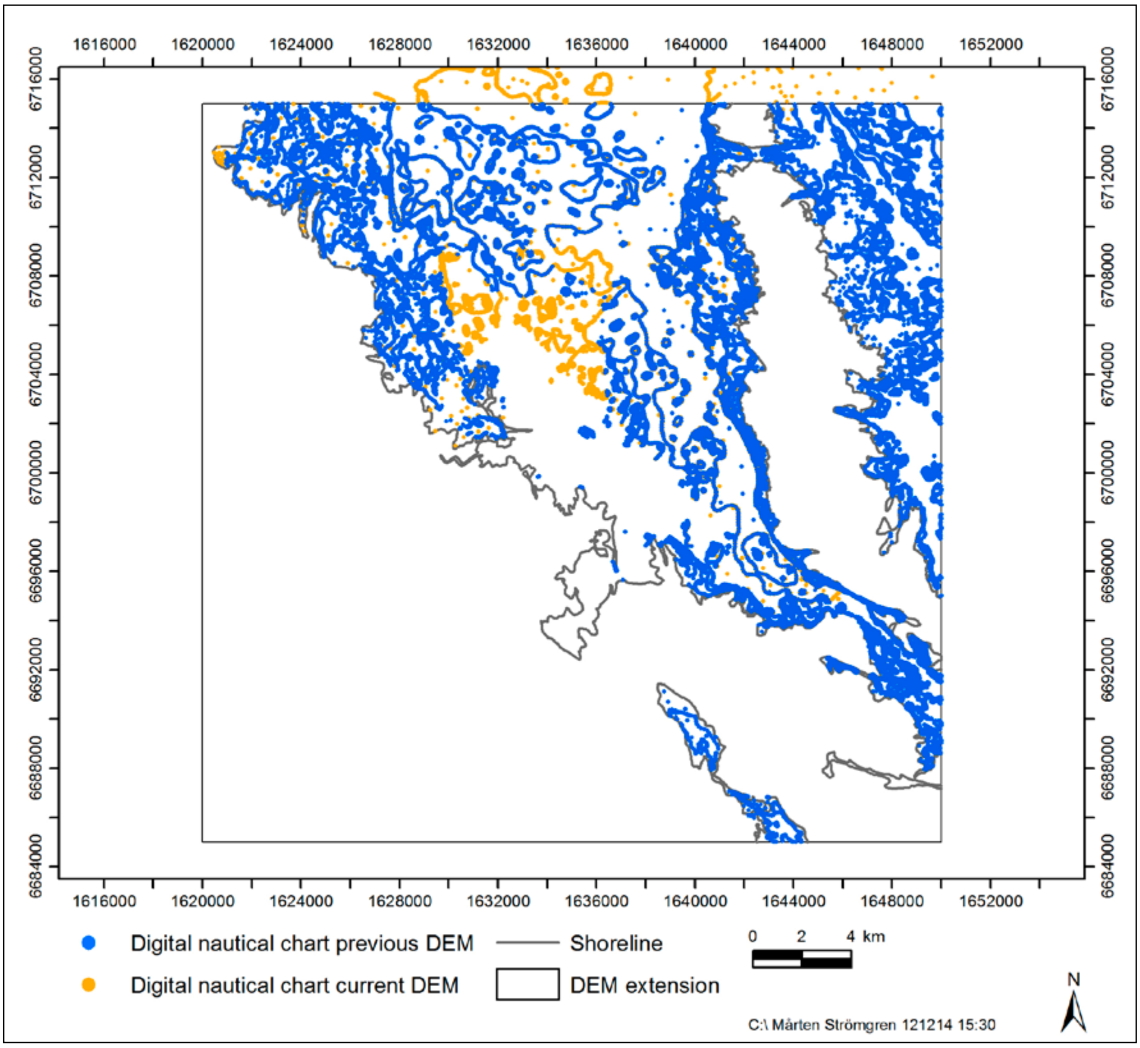

Figure 2-5. Data from the digital nautical chart/the paper nautical chart marked in blue were used in the previous DEM (Strömgren and Brydsten 2008). Data marked in orange and blue were used in the current DEM.

A mapping of reed in the sea in the Forsmark survey area was done in September 2010 (Strömgren and Lindgren 2011). During this mapping water depth was measured. Water depth was also measured during transportation between areas where reed were mapped. These depth measurements (area B and D in Figure 2-3) were recalculated to RH 70 and added to the dataset used for interpolation of the new DEM.

All data available for the marine area for the construction of the DEM are shown in Table 2-3.

Table 2-3. Available data for the marine area for the construction of the DEM.

Data	No of points	Reference
SGUs detailed survey area, 10 km^2	10,694,592	(Nyberg et al. 2011)
SGUs detailed survey area, reanalysed data	174,937	(Nyberg et al. 2011)
SGUs detailed survey area, interpreted data	27,480	(Elhammer and Sandkvist 2005)
SGUs regional survey area, reanalysed data	155,301	(Nyberg et al. 2011)
Shallow bays	84,122	(Brydsten and Strömgren 2004)
Digital nautical chart/paper nautical chart	28,221	(the Swedish Maritime Administration/ number 535 Öregrund – Grundkallen – Björn)
Mapping of reed	12,639	(Strömgren and Lindgren 2011)
With DGPS measured shoreline	8,470	(Strömgren and Brydsten 2008)
Digitized shoreline from IR orthophotos	12,018	(Strömgren and Brydsten 2008)
Sea shoreline from the Property map	95,264	(Lantmäteriet)
Depth chart for the Biotest Lake from 1976	3,874	(obtained by personal communication with Peter Karås SLU, also stored at SKB_SVN\SFR\SR-PSU\Landscape\Indata)
Constructional drawings for the inlet channel to the nuclear power plant	471	(Vattenfall 1977)
Supporting depth data in shallow bays	5,857	(Strömgren and Brydsten 2008)

2.2 Interpolation of data

The previous DEM (Strömgren and Brydsten 2008) is produced from an interpolation of irregularly spaced point values. Some of these point values are gathered along survey lines with almost no elevation data in between them. The maximum distance between some of these survey lines is sometimes more than 1,000 m and the distance between points along survey lines only around 3 m. In the previous DEM, a diffuse pattern is seen along some of these survey lines compared to the areas in between them. The uncertainty estimation of the previous DEM (Strömgren and Brydsten 2009) shows a higher accuracy in areas where the distance between depth data in survey lines used in the interpolation is short (around 100–200 m and less) and lower accuracy in areas where survey lines are separated by long distance.

Using Kriging (Davis 1986, Isaaks and Srivastava 1989) as interpolation method, regularly spaced digital elevation models were produced testing different models in ordinary Kriging, using the method described in Strömgren and Brydsten (2008). No matter new data and reanalysed data were added to the dataset used for the interpolation, a diffuse pattern was seen along some of the survey lines separated by long distance in all these new digital elevation models. Despite point values from survey lines perpendicular to the data used for the previous DEM had been added the distance between points along the survey lines seems to be too short and the distance between the survey lines too long, resulting in this pattern. It became clearly that another method had to be used resulting in a smoother DEM without this pattern. Instead of showing all tests done during the development of the new method, the final method for producing the DEM is described.

The method for producing the current DEM mainly differs in the following aspects from the method used for the previous DEM:

1. the interpolation was done in two domains (Figure 2-6). A natural border between these two domains are the sea shoreline, since most data on land are arranged in regularly spaced grids with different resolution and most data in sea are arranged along survey lines. In the interpolation of the previous DEM no domains were used,
2. the data from survey lines in the regional survey area were thinned out to a distance of approximately 50 m compared to around 3 m in the previous DEM,
3. a buffering distance of 125 m was used between data from survey lines in the regional survey area and the data from the digital nautical chart, and
4. different interpolation techniques in ordinary Kriging were used in the two domains. In the land area an exact interpolation was performed, precisely as in the construction of the previous DEM. In the sea area a smooth interpolation (Gribov and Krivoruchko 2004) was performed instead.

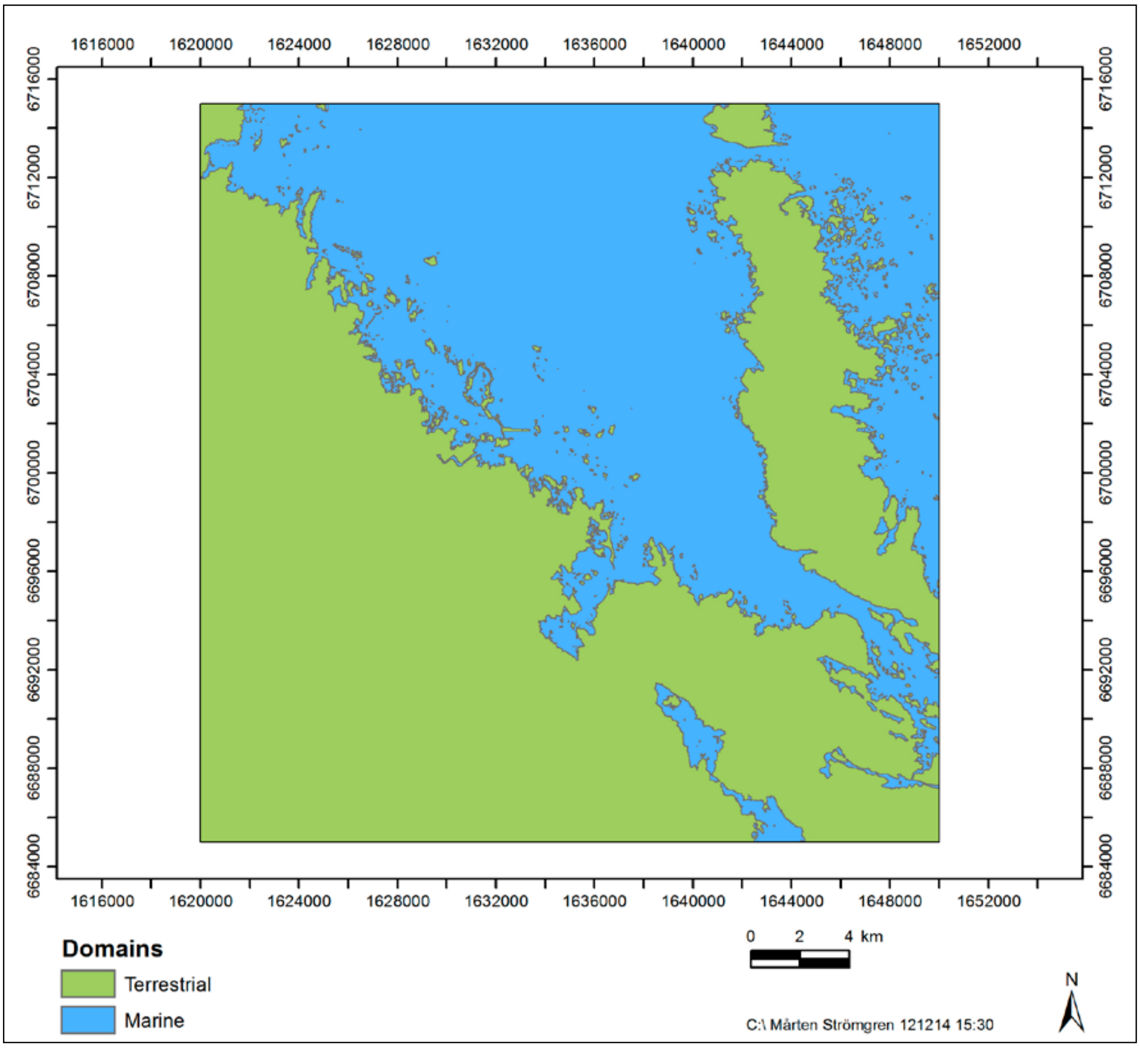

Figure 2-6. *Domains used for the construction of the DEM. The sea shoreline is used for demarcation of the two domains.*

All elevation point values used for the interpolation procedure in the construction of the DEM (Table 2-4) were merged to a dataset with approximately 1.9 million points.

From this dataset two different datasets for producing digital elevation models for the terrestrial and marine domains were created. All elevation point values for land (including lakes) and elevation data from the sea within 200 m from the terrestrial domain were used for interpolation of data in the terrestrial domain. All elevation point values from the sea and data from land within 200 m from the marine domain were used for the interpolation of data in the marine domain. The overlap between these datasets was necessary to generate a smoother transition in the DEM in the border between the two domains.

With these datasets a DEM representing land surface and lake bottoms, and a DEM representing sea bottom were constructed using the extension for the previous DEM (Strömgren and Brydsten 2008), 1,619,990 west, 1,650,010 east, 6,715,010 north, and 6,684,990 south. The RT 90 2.5 Gon W map projection and the height system RH 70 was used in these elevation models.

The interpolation from irregularly spaced point values to a regularly spaced DEM in the terrestrial and marine domains was done using the software ArcGis 9.3.1 Geostatistical Analysis extension. Ordinary Kriging was chosen as the interpolation method (Davis 1986, Isaaks and Srivastava 1989). The choosing of theoretical semivariogram model and the parameters scale, length, and nugget effect were done with the extension. An exact interpolation was performed in the terrestrial domain.

Table 2-4. Data used for the interpolation procedure in the construction of the DEM.

Data	No of points	Reference
SGUs detailed survey area, 10 km^2	430,558	(Nyberg et al. 2011)
SGUs detailed survey area, reanalysed data	127,739	(Nyberg et al. 2011)
SGUs detailed survey area, interpreted data	19,179	(Elhammer and Sandkvist 2005)
SGUs regional survey area, reanalysed data	3,813	(Nyberg et al. 2011)
Shallow bays	78,358	(Brydsten and Strömgren 2004)
Digital nautical chart/paper nautical chart	28,221	(the Swedish Maritime Administration/ number 535 Öregrund – Grundkallen – Björn)
Mapping of reed	12,639	(Strömgren and Lindgren 2011)
With DGPS measured shoreline	8,470	(Strömgren and Brydsten 2008)
Digitized shoreline from IR orthophotos	12,018	(Strömgren and Brydsten 2008)
Sea shoreline from the Property map	95,264	(Lantmäteriet)
Depth chart for the Biotest Lake from 1976	3,874	(obtained by personal communication with Peter Karås SLU, also stored at SKB_SVN\SFR\SR-PSU\Landscape\Indata)
Constructional drawings for the inlet channel to the nuclear power plant	471	(Vattenfall 1977)
Supporting depth data in shallow bays	5,857	(Strömgren and Brydsten 2008)
SKBs 10 m DEM	1,262,747	(Wiklund 2002)
LMVs 50 m DEM	150,027	(Swedish national land survey, LMV)
Airborne laser scanning	93,686	(SKB GIS databas, C201 SDEADM.SKB_FM_HOJ_7882)
Measured values from lakes	21,293	(Brunberg et al. 2004)
Brook measurements	634	(Brydsten and Strömgren 2005)
Fixed points	68	(Property map)

The resolution was set to 20 m. The same procedure was repeated for the marine domain, but in Ordinary Kriging a smooth interpolation (Gribov and Krivoruchko 2004) was performed instead of using the exact interpolation choice.

The models used for interpolation of point values in the terrestrial and marine domains were validated with cross-validation (one data point is removed and the rest of the data are used to predict the removed data points) and ordinary validation (part of the data are removed and the rest of the data are used to predict the removed data). Both the cross-validation and ordinary validation goals produce a standardised mean prediction error near 0, small root-mean-square prediction errors, average standard error near root-mean-square prediction errors, and standardised root-mean-square prediction errors near 1.

Cross validations with different combinations of Kriging parameters were performed until the standardised mean prediction errors were close to zero, but not necessarily the lowest value was always chosen. Because the aim was to determine the most valid model for both measured and unmeasured locations, care was taken to produce low values for the root-mean-square prediction errors and minimise the difference between the root-mean square prediction errors and the average standard errors. Different models were compared and the ones with the most reasonable statistics were chosen.

Validations were performed with the most appropriate Kriging parameters in order to verify that the models fit unmeasured locations. The final choice of parameters is presented in Appendix 1.

The terrestrial and marine domains were used to separate the digital elevation models for the land and sea areas. Finally, the digital elevation models for the terrestrial and marine domains were merged to one DEM, representing the Forsmark area.

2.3 Uncertainty estimation of the DEM

The uncertainty estimation of the DEM in both the terrestrial and marine areas is not performed against single elevation values, but instead against 20 m cells produced in areas covered with data of high accuracy and high point density.

The calculation of uncertainty of different parts in the DEM was mainly done using the method presented in Strömgren and Brydsten (2009). However, some new data have become available since this study was performed and one other method for the uncertainty calculation has instead been used for part of the DEM. It should also be noted that the uncertainty of the DEM is only calculated for areas with data sources covering larger parts of the DEM and therefore is supposed to be more important for further modelling with the DEM used as input data. All 20 m cells used for uncertainty calculations are shown in Figure 2-7.

2.3.1 The terrestrial domain

Measurements using total station have shown that airborne laser scanning data are very accurate (Strömgren and Brydsten 2009). The point density in part of the DEM covered with data from the airborne laser scanning measurements is high (25 points/20 m cell). In Strömgren and Brydsten (2009) some tests were performed regarding the importance of point density for the accuracy of 20 m cells produced using Kriging interpolation. These tests showed that using only 9 points/20 m cell results in 20 m cells with values very close to the cells produced using 121 points/20 m cell.

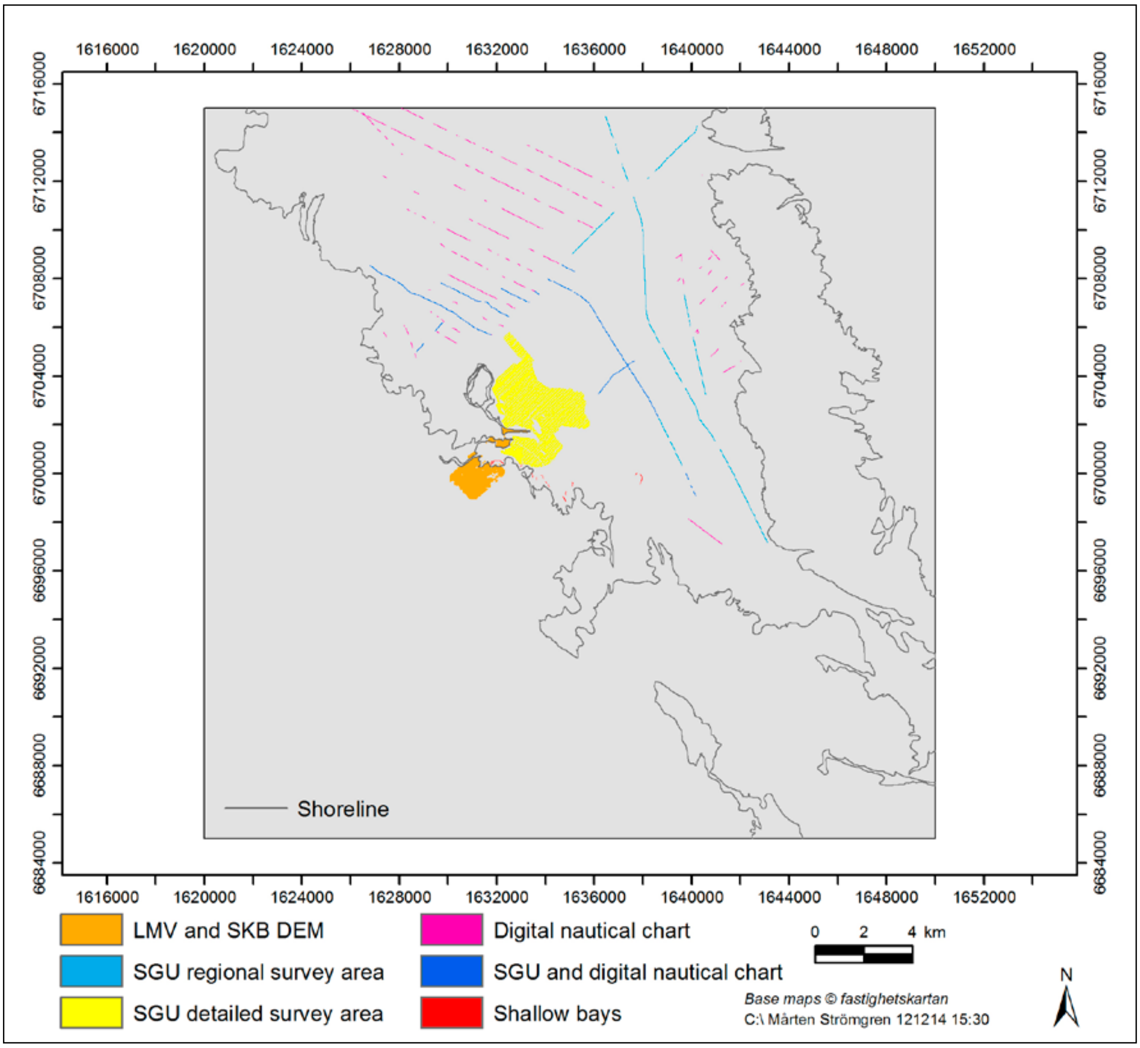

Figure 2-7. *20 m cells used for uncertainty calculations of different parts of the DEM.*

Based on these tests, the 20 m cells produced in the area covered with airborne laser scanning measurements in the DEM should result in 20 m cells of high accuracy and consequently could be used for uncertainty estimation of terrestrial areas covered with other data. This was done using the following method: Data from SKBs 10 m elevation model were used to replace data from the airborne laser scanning measurements in the dataset used for the construction of the terrestrial part of DEM (Figure 2-8 a and b). A raster layer was produced using the same Kriging parameters that were used for the construction of the corresponding layer for the DEM. The same procedure was repeated using data from LMVs 50 m DEM instead of data from the airborne laser scanning measurements (Figure 2-8 a and c). The difference between the raster layers produced using data from the 10 and 50 m elevation models and the DEM was calculated for all 20 m cells covered with data from the airborne laser scanning measurements (Figure 2-7). Descriptive statistics from these calculations are presented in Table 2-5.

Table 2-5. Statistical analysis of the accuracy of 20 m cells covered with different input data sources used for the DEM. No normal distribution test is performed. N = number of 20 m cells used for the calculation, RMS = root means square, STDV = standard deviation, CV = coefficient of variation, Min = minimum error, and Max = maximum error. The last three columns show the 5, 25, 75, and 95 percentiles, respectively. The unit is metre except for CV, which is shown in percent.

Data source extension	N	Mean	Median	RMS	STDV	CV	Min	Max	5%	25%	75%	95%
SKBs 10 m DEM	6,557	0.19	0.14	0.70	0.68	358	–2.97	5.21	–0.84	–0.19	0.54	1.32
LMVs 50 m DEM	6,557	–0.2	–0.23	1.56	1.55	775	–7.19	7.88	–2.73	–1.17	0.76	2.48

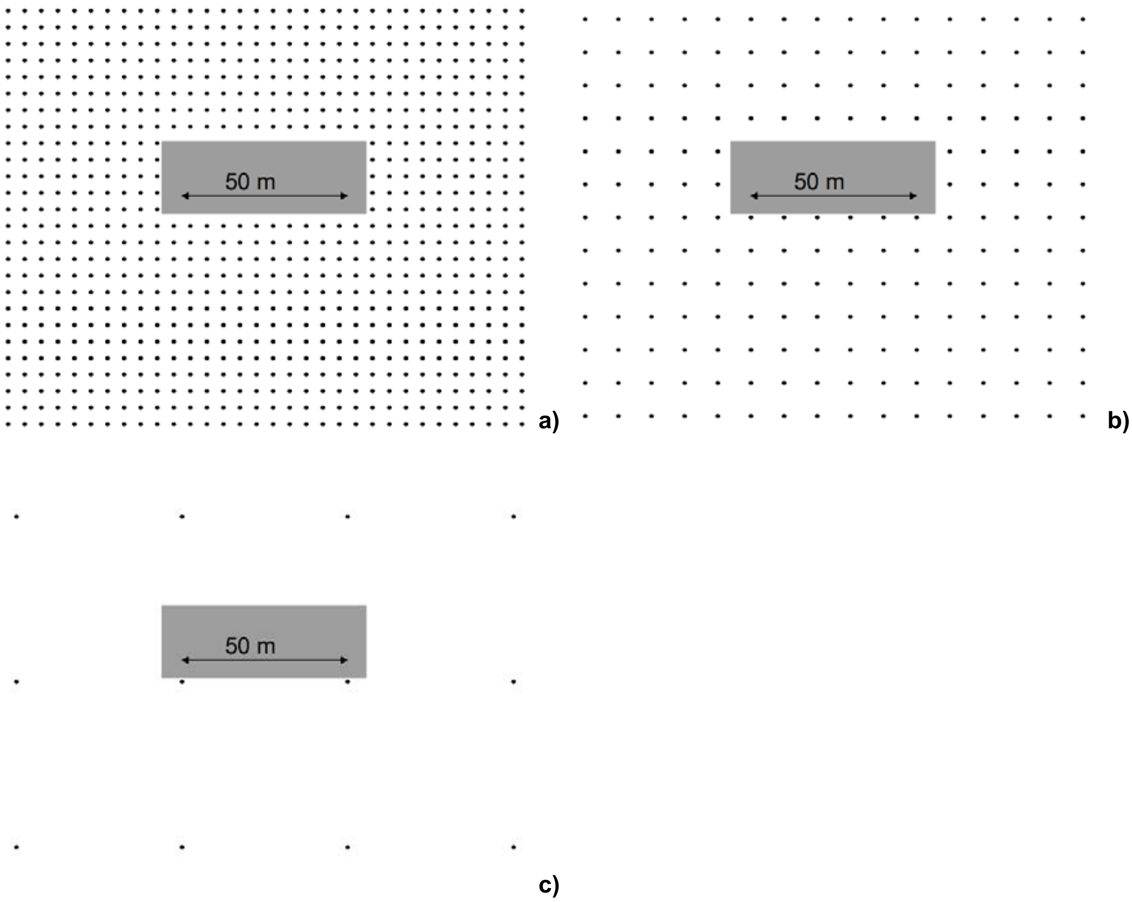

Figure 2-8. *Example of area for uncertainty calculation of the DEM. a) Part of the DEM covered with data from the airborne laser scanning, b) replaced with data from SKBs 10 m elevation model, and c) replaced with data from LMVs 50 m elevation model. All 20 m cells used for this calculation are shown in Figure 2-7.*

2.3.2 The marine domain

In the sea, areas with depth data produced from a survey performed in 2010 (Nyberg et al. 2011), reanalysed data from the same survey area, and data from Strömgren and Lindgren (2011) were used for uncertainty calculations of different parts of the DEM. These data sources should be of high accuracy and the point density in 20 m cells covered with data from these sources are high. The DEM were divided in five areas depending on data sources covering the area or the water depth.

Some data used for the interpolation in the marine domain were removed in order to imitate the spatial distribution of data in these five areas. Datasets with elevation data within 200 m from the marine domain were created for these five areas. Finally, five raster layers were produced from these datasets using the Kriging parameters that were used for the interpolation in the marine domain.

The principle for the uncertainty calculations of different part of the DEM is only illustrated for some of the marine areas. Below follows a short description of how the datasets for each one of these five areas were produced:

1. The uncertainty calculation for part of the DEM mostly covered with data from shallow bays (Brydsten and Strömgren 2005) was done using 20 m cells covered with data from crossing survey lines from a mapping of reed performed in 2010 (Strömgren and Lindgren 2011) (Figure 2-9).

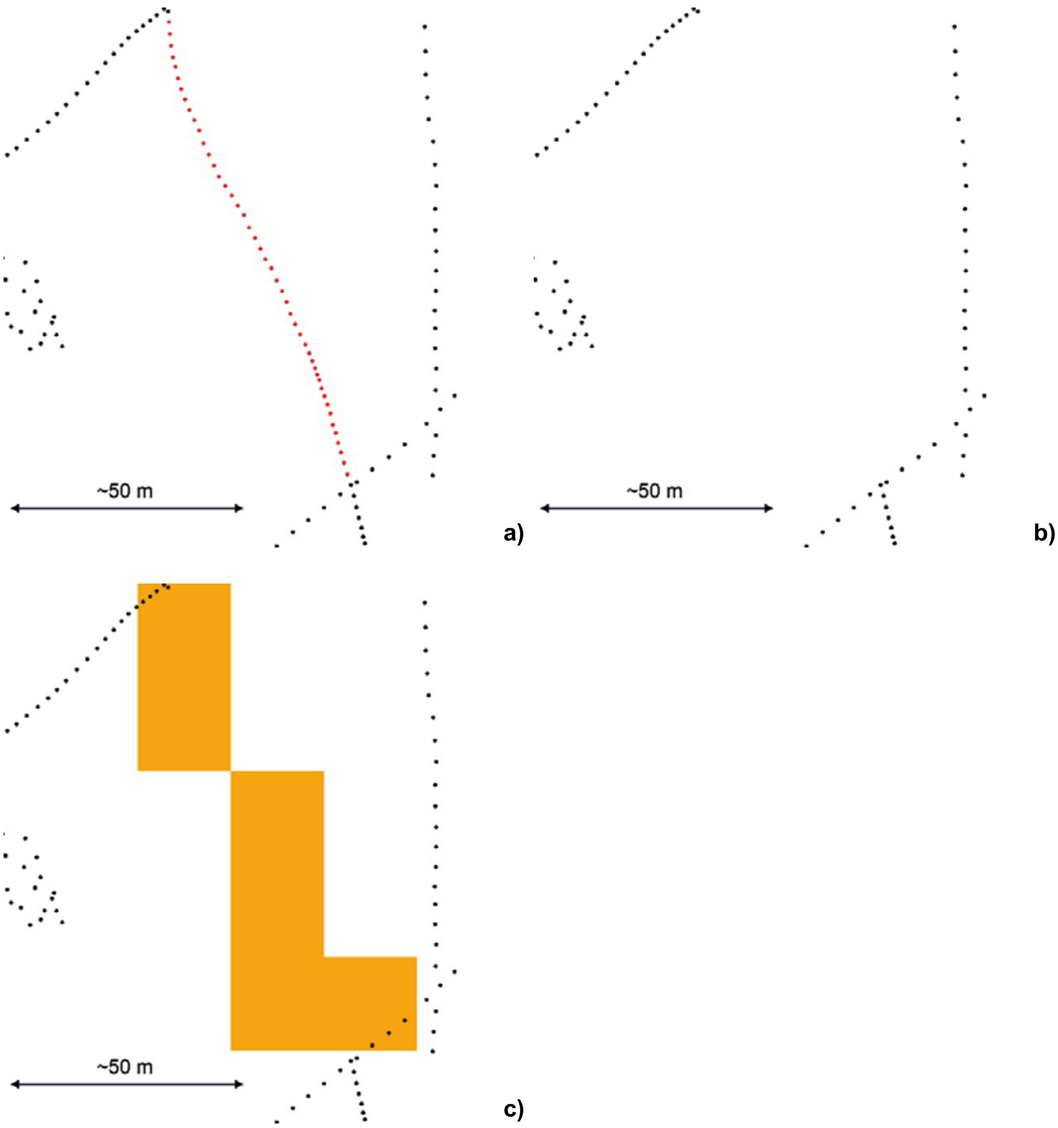

Figure 2-9. Example of area for uncertainty calculation of shallow bays in the DEM. a) Part of the dataset used for the interpolation procedure in the construction of the DEM, b) data removed from a crossing survey line, and c) 20 m cells used for uncertainty calculation (marked in orange). All 20 m cell used for uncertainty calculation of shallow bays are shown in Figure 2-7.

Mostly 20 m cells between survey lines were used for the uncertainty calculation. Some of the data from the mapping of reed in the dataset used for the interpolation procedure were removed, to imitate areas with data from the measurements of shallow bays where no crossing survey lines exist.

2. An approximately 10 km^2 large area (Area C in Figure 2-3), where a very detailed marine geological survey was performed in 2010 and reanalysed survey data retrieved in 2002 for the same area also were used (Nyberg et al. 2011). This data should be of high accuracy and the point density is high, 16 points/20 m cell, producing 20 m cells that can be used for uncertainty calculation of areas covered with measurements from SGUs detailed survey area (Area B in Figure 2-3), in which survey lines are separated with 100–200 m distance between the lines. Most of the data within Area C in Figure 2-3 used for the construction of the DEM were removed and instead only data from survey lines from the reanalysed data in this area were used, to imitate the detailed survey area (Area B in Figure 2-3). This is illustrated in Figure 2-10.

3. The uncertainty calculation of part of the DEM only covered with data from the digital nautical chart was done using 20 m cells covered with data from SGUs regional survey area (Nyberg et al. 2011). All data from the regional survey area were used and not only the data used for the construction of the DEM. This was only done for areas less than 20 m deep according to the depth lines from the digital nautical chart, since most areas of the DEM covered only by data from the digital nautical chart are above the 20 m depth lines. All points from the regional survey area used for the construction of the DEM within the 20 m cells chosen above were removed, to imitate areas where only data from the digital nautical chart exist (Figure 2-11). A raster layer was produced from this dataset.

4. The uncertainty calculation of the part of the DEM covered with data from the digital nautical chart and survey lines from SGUs regional survey area (Nyberg et al. 2011) was done using 20 m cells covered with data from SGUs regional survey area. All points from the regional survey area used for the construction of the DEM within the 20 m cells chosen above were removed, to imitate other areas where less data from the digital nautical chart and SGUs regional survey area exist. A raster layer was produced from this dataset.

5. The uncertainty calculation of the part of the DEM covered with data from SGUs regional survey area (Nyberg et al. 2011), in areas deeper than 20 m was done using 20 m cells covered with data from SGUs regional survey area. 20 m cells covered with data from SGUs regional survey area, from survey lines perpendicular to other survey lines were chosen. All points from the regional survey area used for the construction of the DEM within the 20 m cells chosen above were removed, to imitate other areas where less data in the regional survey area exist. A raster layer was produced from this dataset.

The uncertainty in areas 1 and 2 described above was calculated as the difference between the raster layers produced for these areas and the DEM, within the chosen 20 m cells (Figure 2-7).

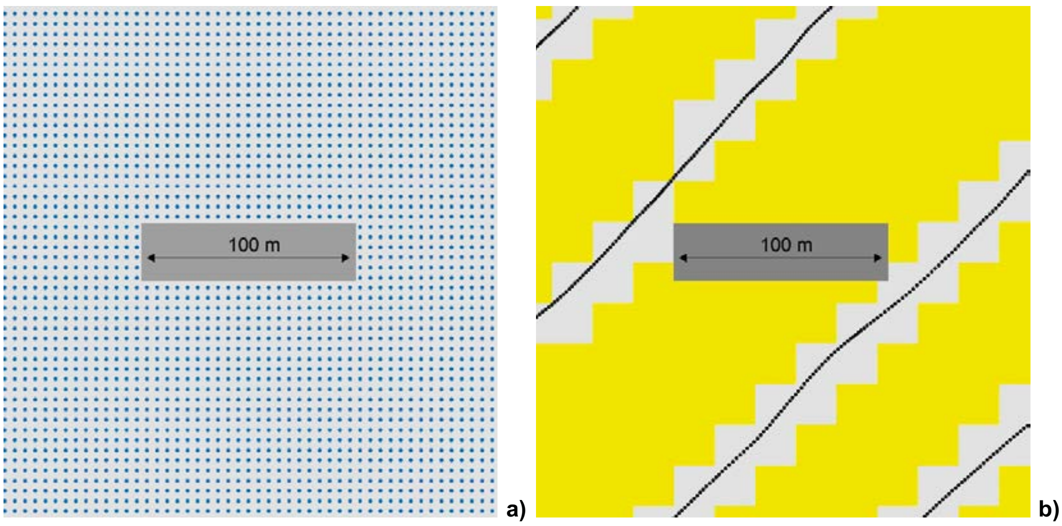

Figure 2-10. *Example of area for uncertainty calculation of part of the DEM covered with data from SGUs detailed survey area. a) Part of the dataset used for the interpolation procedure, and b) data removed and replaced by reanalysed data for the same area (black points). The 20 m cells used for uncertainty calculation are marked in yellow. All 20 m cell used for uncertainty calculation of this area are shown in Figure 2-7.*

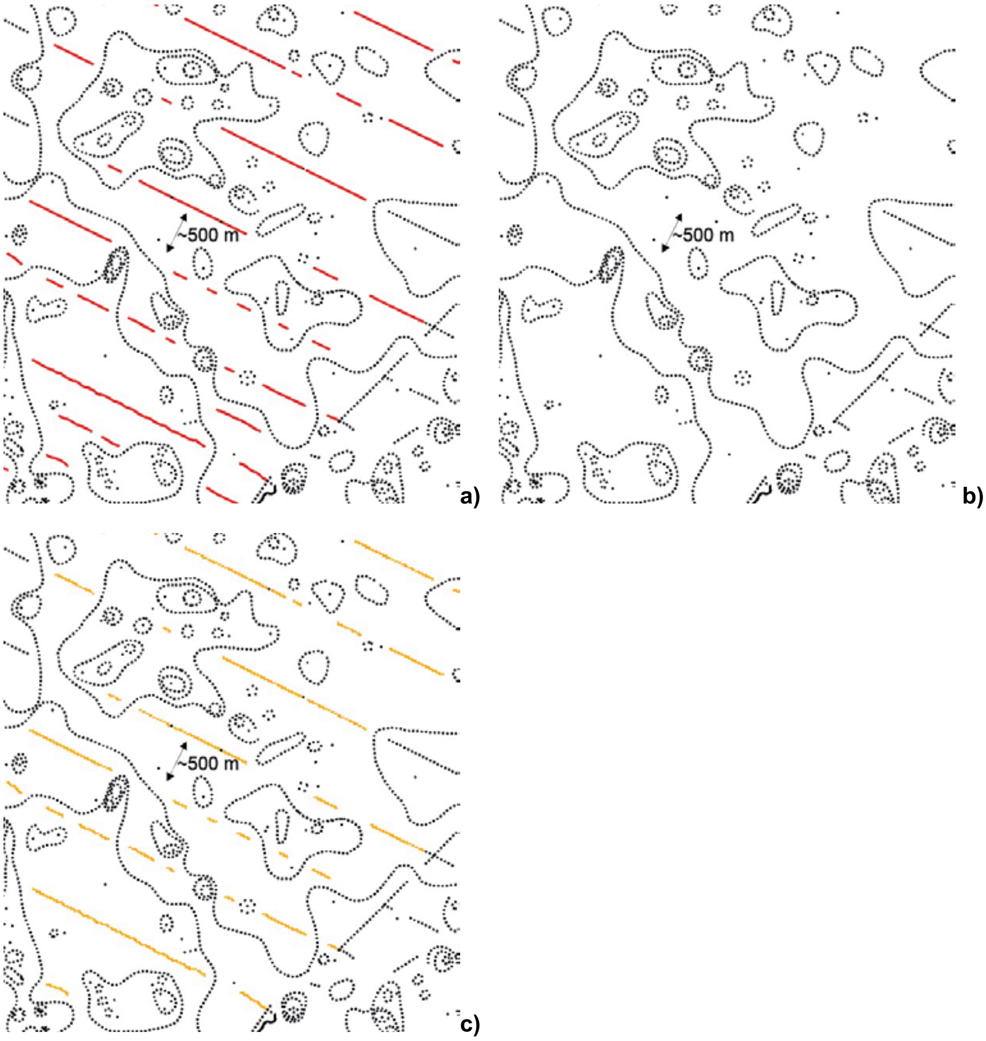

Figure 2-11. Example of area for uncertainty calculation of part of the DEM covered with data from the digital nautical chart. a) All data from SGUs regional survey area and other data used in the sea domain were used for the interpolation. b) Data from SGUs regional survey area were removed to imitate the condition in areas covered with only data from the digital nautical chart. c) The 20 m cells used for the uncertainty calculation are marked in orange. All 20 m cells used for this calculation are shown in Figure 2-7.

The uncertainty in areas 3–5 described above was calculated as the difference between the raster layers produced for these areas and a raster layer produced using all data from SGUs regional survey area (i.e. 3 m between measurement points compared to 50 m in the dataset used for the construction of the DEM) and all other data used for the marine domain in the construction of the DEM, within the chosen 20 m cells (Figure 2-7).

Descriptive statistics from all uncertainty calculations are presented in Table 2-6.

Table 2-6. Statistical analysis of the accuracy of 20 m cells covered with different input data sources for the DEM. No normal distribution test was performed. N = number of 20 m cells used for the calculation, RMS = root mean square, STDV = standard deviation, CV = coefficient of variation, Min = minimum error, and Max = maximum error. The last three columns show the 5, 25, 75, and 95 percentiles, respectively. The unit is metre except for CV, which is shown in percent.

Data source extension	N	Mean	Median	RMS	STDV	CV	Min	Max	5%	25%	75%	95%
SGUs regional survey area below 20 m	1,700	2.49	1.84	4.21	3.39	136	–6.5	12.6	–2.8	0.28	4.6	8.96
SGUs detailed survey area	18,872	0.17	0.09	0.80	0.79	465	–5.5	5.82	–0.9	–0.2	0.5	1.45
Shallow bays	193	0	–0.01	0.00	0.02	17	–0.2	0.01	0	0	0	0
SGUs regional survey area/ digital nautical chart above 20 m	1,410	1.58	1.52	2.35	1.74	110	–4.2	8.79	–0.9	0.41	2.5	4.59
Digital nautical chart	2,492	2.99	2.59	3.83	2.41	81	–4	15	–0.4	1.38	4.6	7.01

3 Results

3.1 The DEM

The DEM describing land surface, sediment level at lake bottoms, and sea bottom is illustrated in Figure 3-1.

The DEM has a size of approximately 30×30 km a cell size of 20 m, 1,501 rows, and 1,501 columns: a total number of DEM cells of 2,253,001 and a file size of approximately 8.8 MB (ESRI Grid format). The extension is 1,619,990 west, 1,650,010 east, 6,715,010 north, and 6,684,990 south in the RT 90 coordinate system, and the height system is RH 70. The area is extremely flat so the range in elevation is only approximately 105 m with the highest point at 50.1 m above sea level at the south-west part of the DEM, and the deepest sea point at –55.4 m in the northern part of the so-called Gräsörännan.

The mean elevation in the DEM is only 1.8 m. The model area is covered by 59% land and 41% sea. The flat landscape is also shown in the statistics of the slope where the mean slope is 1.28 degrees. The slope is lower than 5 degrees in 98.6% of the cells and the slope is between 5 and 10 degrees in 1.3% of the cells. Almost all of the cells with a slope steeper than 10 degrees (0.03%) are man-made such as the inlet channel to the nuclear power plant or piers and wharfs close to SFR.

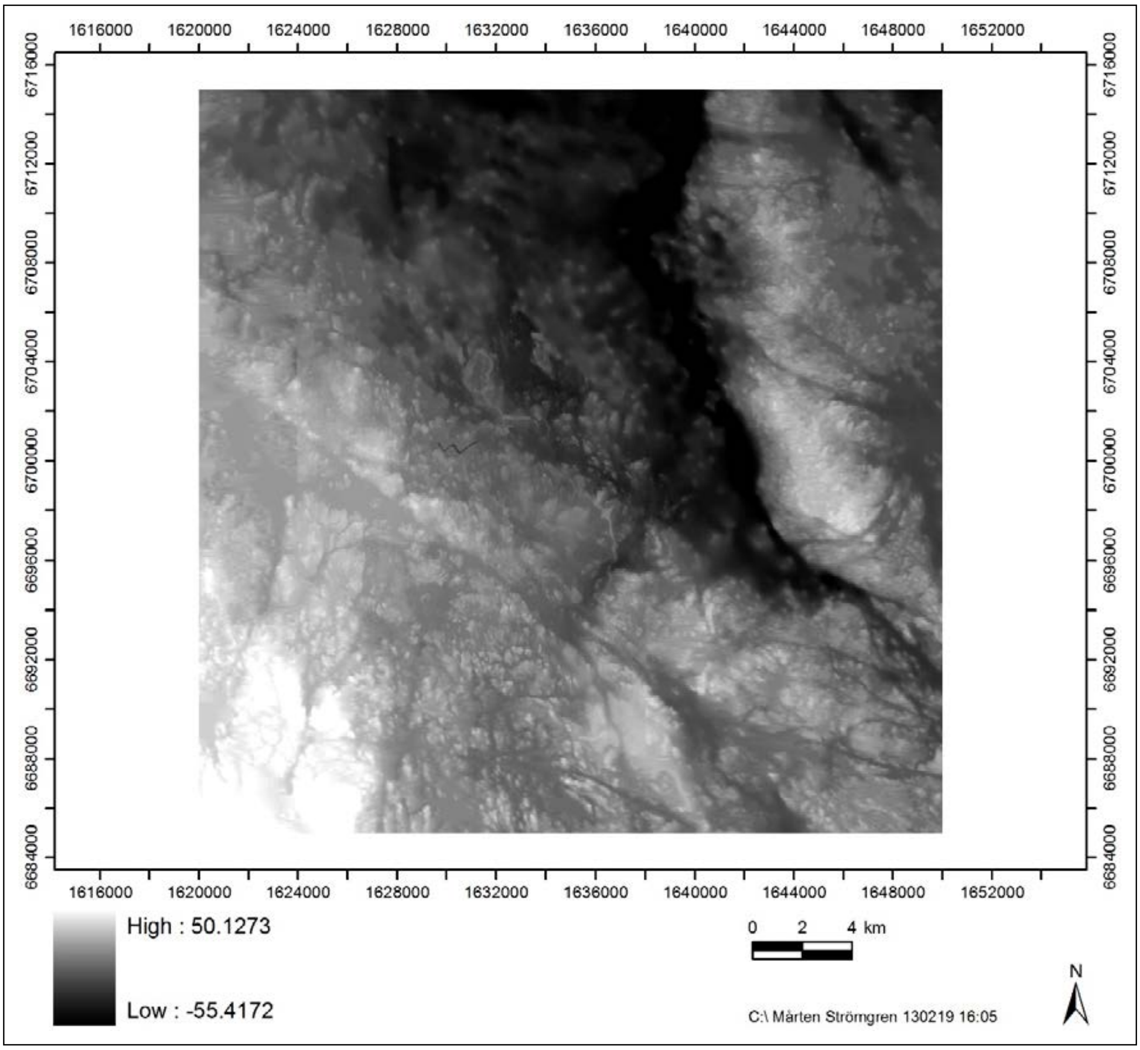

Figure 3-1. The 20 m DEM describing land surface, lake bottoms, and lake sediment surfaces.

The corresponding statistics for the previous DEM (Strömgren and Brydsten 2008) suggests that the current DEM is slightly smoother. Figure 3-2 shows a part of the previous DEM (a) and the current DEM (b) where the distance between surveys lines from SGU is approximately 1 km with no crossing survey lines in the previous DEM. Both models are stretched and hillshade effect is used. Darker shades indicate deeper areas. The current DEM shows a much smoother surface compared to the previous DEM where a ribbed pattern is visible along the survey lines.

3.2 Accuracy of the DEM

Statistics from uncertainty calculations of different parts of the DEM is shown in Table 3-1. The root mean square (RMS) value for shallow bays, 0.00 m, is lowest followed by 0.70 m for SKBs 10 m DEM and 0.80 m in SGUs detailed survey area. However, the coefficient of variation (CV) is much lower in shallow bays compared to the two other areas and no larger error is likely to be found in shallow bays. Based on the statistics in Table 3-1 errors larger than 2 m are unusual in SKBs 10 m DEM and in SGUs detailed survey area.

The root mean square for the area covered with data from LMVs 50 m DEM is 1.56 m and 2.35 m for the area covered with data from SGU and the digital nautical chart below 20 m depth. However, the coefficient of variation is much higher in the area covered with data from the 50 m DEM, 775 percent compared to 110 percent. Errors larger than a few metres can be expected in both these areas.

The root mean square for the area only covered with data from the digital nautical chart is 3.83 m and 4.21 m for the area covered with data from SGU below 20 m depth. The coefficient of variation is quite similar for these areas, 81 percent and 136 percent respectively, and errors larger than 10 m can be expected in both these areas.

Figure 3-3 shows areas in the DEM that are divided depending on data sources covering the area or the water depth. These areas are primarily ranked from the root mean square (RMS) shown in Table 3-1, but also the other statistics shown in the table are taken into consideration in the ranking. The lowest ranking number shows the area of highest accuracy. The areas used for the uncertainty calculations, i.e. the area covered with data from the airborne laser scanning and SGUs 10 km^2 detailed survey area, are ranked as number 1 and 2, respectively. This is based on the knowledge of input data and the point density in these areas.

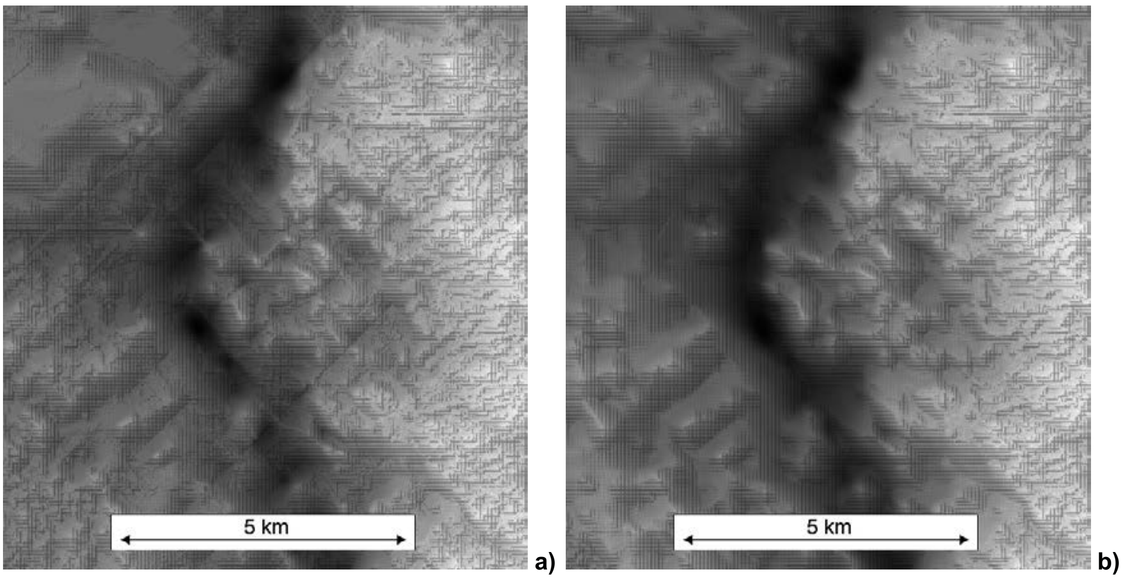

Figure 3-2. Part of the previous DEM (a) and the current DEM (b) where the distance between survey lines from SGU is approximately 1 km with no crossing survey lines in the previous DEM. Both models are stretched and hillshade effect is used. Darker shade indicate deeper areas. The current DEM shows a smoother surface compared to the previous DEM.

Table 3-1. The table is a summary of Table 2-5 and 2-6 and shows a statistical analysis of the accuracy of the DEM in 20 m cells covered by different input data sources. No normal distribution test was performed. N = number of 20 m cells used for the calculation, RMS = root means square, CV = coefficient of variation, Min = minimum error, and Max = maximum error. The last three columns show the 5, 25, 75, and 95 percentiles, respectively. The unit is metre except for CV, which is shown in percent.

Data source extension	N	RMS	CV	Min	Max	5%	25%	75%	95%
SKBs 10 m DEM	6,557	0.70	358	–2.97	5.21	–0.84	–0.19	0.54	1.32
LMVs 50 m DEM	6,557	1.56	775	–7.19	7.88	–2.73	–1.17	0.76	2.48
SGUs regional area below 20 m	1,700	4.21	136	–6.5	12.6	–2.8	0.28	4.6	8.96
SGUs detailed survey area	18,872	0.80	465	–5.5	5.82	–0.9	–0.2	0.5	1.45
Shallow bays	193	0.00	17	–0.2	0.01	0	0	0	0
SGUs regional survey area/digital nautical chart above 20 m	1,410	2.35	110	–4.2	8.79	–0.9	0.41	2.5	4.59
Digital nautical chart	2,492	3.83	81	–4	15	–0.4	1.38	4.6	7.01

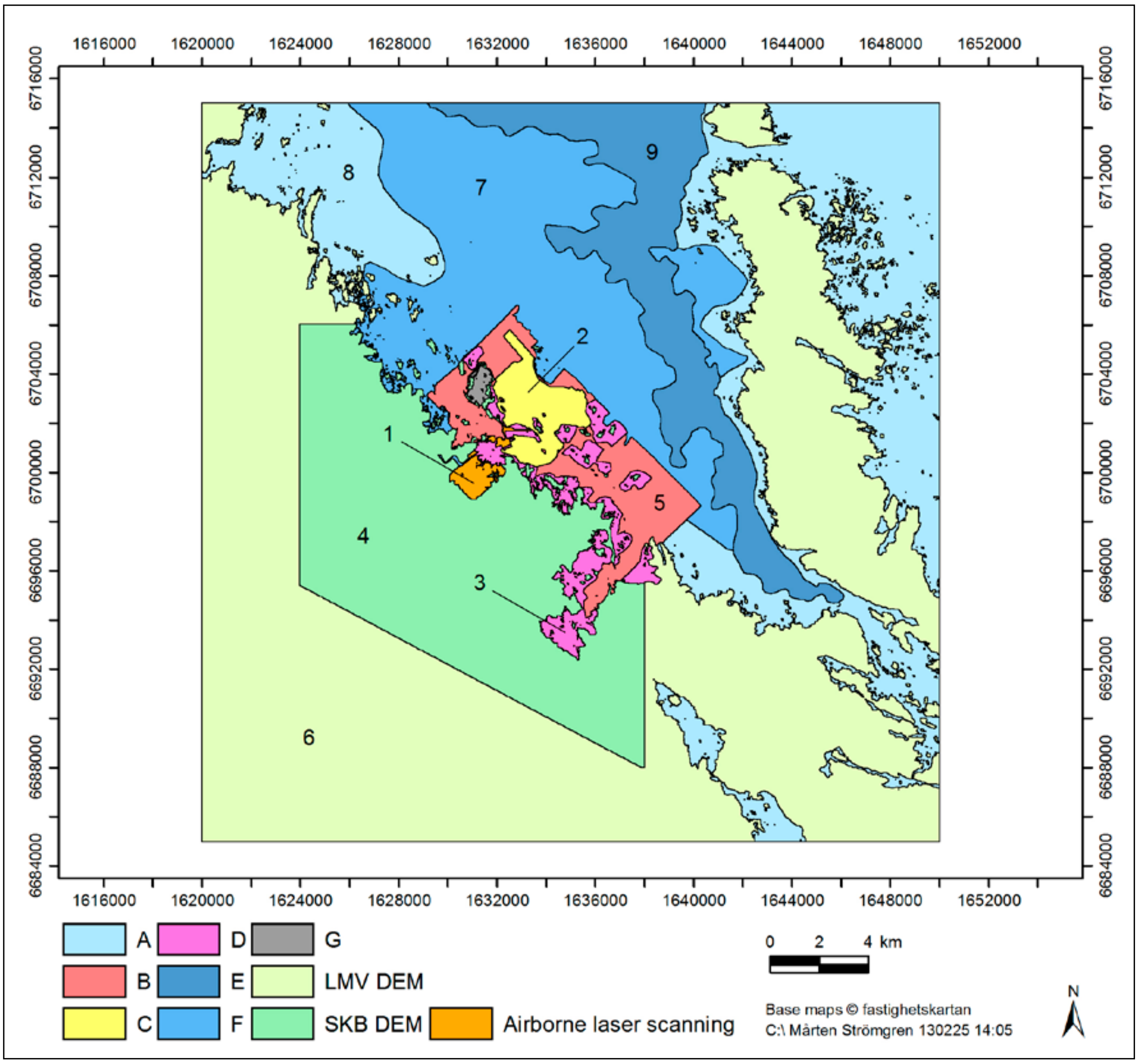

Figure 3-3. *Ranking of accuracy for areas in the DEM divided depending on data sources covering the area or water depth (numbered 1–9). The area covered with data from the airborne laser scanning and SGUs 10 km^2 detailed survey area (C) are ranked highest based on the knowledge of input data and the point density. These areas are also used for the uncertainty calculation of some of the other areas. The shallow bays (D) comes next followed by SKBs 10 m DEM, SGUs detailed survey area (B), LMVs 50 m DEM, and the area covered with data from the digital nautical chart and measurements performed SGU above 20 m depth (F). The accuracy is lowest in the area covered with data from the digital nautical chart (A) and the area covered with data from SGU below 20 m depth (E). No uncertainty calculation was possible for the Biotest Lake (G).*

The shallow bays comes next in ranking followed by SKBs 10 m DEM, SGUs detailed survey area, the 50 m DEM from LMV, and the area covered with data from the digital nautical chart and measurements performed by SGU above 20 m depth. The accuracy is lowest in the area covered with data from the digital nautical chart and the area covered with data from SGU below 20 m depth.

No uncertainty calculation was possible for the lakes in the Forsmark area since no proper validation data exist. However, the distance between measurements points in lakes and the depth could give an indication of the expected errors. For most lakes, the maximum distance between measurements 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 is about the same as in the survey lines from shallow bays and SGUs detailed survey area. Altogether, this would suggest the accuracy for lakes in the DEM to be the same as in the shallow bays. No uncertainty calculation was possible for the Biotest Lake.

3.3 Data files delivered to SKB

Following data files are delivered to SKB:

FM_DEM_110429 ESRI Grid format, land surface, lake bottoms, and sea bottoms.

Elevation data used for interpolation of DEM ESRI Shape format

References

SKB's (Svensk Kärnbränslehantering AB) publications can be found at www.skb.se/publications.

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. Forsmark site investigation. Water depth soundings in shallow bays in Forsmark. SKB P-04-125, Svensk Kärnbränslehantering AB.

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

Davis J C, 1986. Statistics and data analysis in geology. 2nd ed. New York: Wiley.

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

Gribov A, Krivoruchko K, 2004. Geostatistical mapping with continuous moving neighborhood. Mathematical Geology 36, 267–281.

Isaaks E H, Srivastava R M, 1989. An introduction to applied geostatistics. Oxford: Oxford University Press.

Nyberg J, Elhammer A, Sohlenius G, Kjellin B, Nordgren P, 2011. Results from marine geological investigations outside Forsmark. SKB P-11-39, 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, Brydsten L, 2009. Uncertainty aspects of the digital elevation model for the Forsmark area. SKB R-09-46, Svensk Kärnbränslehantering AB.

Strömgren M, Lindgren F, 2011. Mapping of reed in shallow bays. SFR-Site Forsmark. SKB P-11-09, Svensk Kärnbränslehantering AB.

Vattenfall, 1977. Constructional drawing "Forsmarks kraftstation, Block 1, Kylvattenanläggning, Tilloppskanal", Drawings number 1-954511, 1-954512, 1-954513, 1-785661, 1-785662 and 1-797448.

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.

Appendix 1

Cross validation of model

Domain 1

Lag size	Number of Lags	Regression function	Mean	RMS	Average SE	Mean stand	RMS stand	Samples
55	12	0.98 × x + 0.020	0.0001647	0.3588	0.8259	0.000383	0.3738	1,897,140

Domain 2

Lag size	Number of Lags	Regression function	Mean	RMS	Average SE	Mean stand	RMS stand	Samples
200	14	0.996 × x + –0,016	0.00026	0.437	1.381	0.0001602	0.2926	987,942

Validation of model

Domain 1

Lag size	Number of Lags	Regression function	Mean	RMS	Average SE	Mean stand	RMS stand	Samples
55	12	0.997 × x + 0.028	0.001118	0.4359	0.8636	0.001051	0.429	948,570

Domain 2

Lag size	Number of Lags	Regression function	Mean	RMS	Average SE	Mean stand	RMS stand	Samples
200	14	0.994 × x + –0.025	0.0003345	0.5316	1.404	–0.00002351	0.3473	493,971

Model parameters

The model equation should be read as follows:

Partial sill × Theoretical Semiovariogram (Major Range, Minor Range, Anisotropy Direction) + (Nugget value × Nugget).

Domain 1

Points	Modell	MS[1)]	Me[1)]	N[1)]	A[1)]
1,897,140	7.0603×Spherical (651.93, 519.23, 295.3)+0.43872×Nugget	0.43872 (100%)	0 (0%)	5/2	4

[1)] MS = Microstructure, Me = Measurement error, N = Searching Neighbourhood and A = Angular Sectors.

Domain 2

Points	Modell	MS[1)]	Me[1)]	N[1)]	SF[1)]
987,942	25.633×Circular (2,789.3, 1,981, 314.7)+1.6893×Nugget	1.6893 (100%)	0 (0%)	Smooth	1

[1)] MS = Microstructure, Me = Measurement error, N = Searching Neighbourhood and SF = Smoothing factor.