Norwegian Meteorological Institute met.no

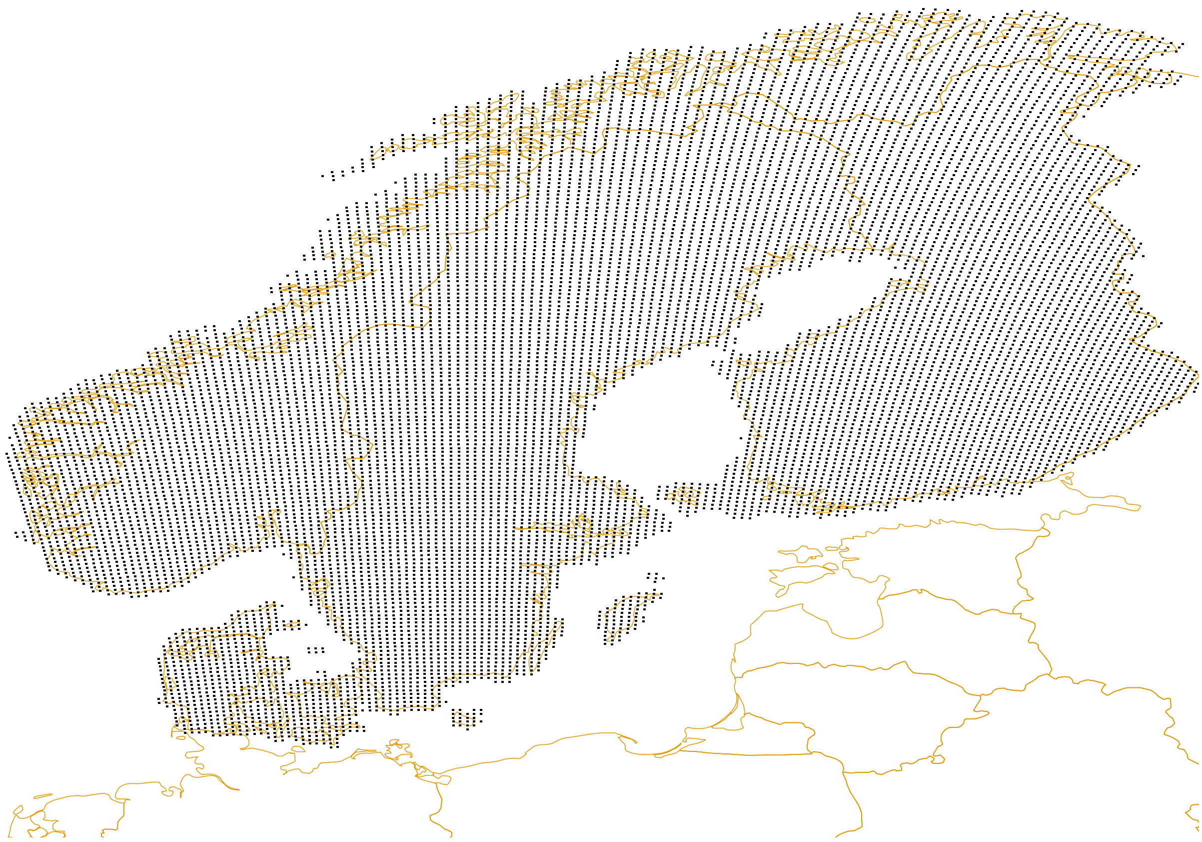
met.no report

no. 03/2007
Climate

NORDGRID
-a preliminary investigation on the potential for creation of a joint Nordic gridded climate dataset

by

Anna Jansson, Ole Einar Tveito, Pentti Pirinen, Michael Scharling



report

Title
NORDGRID - a preliminary investigation on the potential for creation of a joint Nordic gridded climate dataset

Date
08.02.2007

Section
Climatology Division, section for climate research.

Report no.
no. 03/2007 Climate

Author(s)
Anna Jansson (SMHI), Ole Einar Tveito (met.no), Pentti Pirinen (FMI) and Michael Scharling (DMI).

Classification
☒ Free ☐ Restricted

ISSN 1503-8025
e-ISSN 1503-8025

Client(s)
NOSC (NORDMET Steering Committee) and NORDMET Council.

Client's reference

Abstract

NORDGRID was established in 2004 and finished during autumn 2006, the main objectives of the project has been to establish a gridded dataset for model validation (both NWP and climate models), data quality control (objective spatial data quality control of observations), forming a basis for a new European Climatology, and offset for a future daily climatological map of Europe and to show the responsibility of the NMS's to produce such information!

The first phase of NORDGRID had three focuses: Creation of a test dataset (1995-1999) of gridded temperature and precipitation for Fennoscandia based on national gridded data, comparing the gridding methods applied in the different countries and a comparison of the grid estimates along the national borders.

The gridding methods applied in the four countries participating in the first phase of NORDGRID belongs to two different approaches. The methods applied in Denmark, Finland and Norway are traditional spatial interpolation methods which takes different terrain and other physiographical information into consideration. The method applied in Sweden is based on objective analysis utilizing a first guess field (NWP forecast or re-analysis) and observations. Both these two concepts have advantages and limitations.

For this first phase it was concluded that the gridding methods used in Sweden and Norway had the smallest biases when validated against a set of reference stations. The reference stations (the EWGLAM-stations) are though very limited in number, and strongly biased towards lower altitudes. The Danish method was not validated as its setup not allows such analysis. The validation procedure was not executed completely similarly in all countries. Further investigations and comparisons to assess the strengths and weaknesses of the methods are therefore needed before improving and selecting a method for producing a joint Nordic gridded dataset. The comparison of gridded values along the borders revealed large variations, especially in temperature.

Keywords
Gridded datasets, spatial interpolation, temperature, precipitation, Fennoscandia

Disiplinary signature **Responsible signature**

Eirik J. Førland

Postal address	Office	Telephone	Telefax	e-mail: met@met.no	Bank account	Swift code
P.O.Box 43, Blindern NO-0313 OSLO Norway	Niels Henrik Abelsvei 40	+47 22 96 30 00	+47 22 96 30 50	Internet: met.no	7694 05 00628	DNBANOKK

CONTENTS

1 Introduction

The kick-off meeting of NORDGRID-project was held at FMI in May 2004 with participants from Denmark, Finland, Norway and Sweden. The meeting ended up in a pilot project consisting of three modules:

Module A: Comparison of gridding methods used in the participating countries. In this part all participating countries (except Denmark) have made gridded maps of daily mean temperature and daily precipitation for Fennoscandia (Sweden, Norway, Finland and Denmark) for year 1999. The gridded data have also later been validated. In all gridding procedures in Finland, Norway and Sweden the same observations have been used. The observations are also quality controlled at each national institute. When daily mean temperatures are compared between countries, problems arise because of different daily mean temperature algorithms and the different time interval the mean temperature is calculated for. This problem has also been taken into consideration in module A.

Module B: Creation of a gridded dataset based on “National gridded data” from each participating country. Here, already existing gridded maps from each institute of daily mean temperature and daily precipitation have been patched together along the national borders for the years 1995-1999 to create a gridded dataset for the whole Fennoscandia.

Module C: In the last part, the dataset created in module B is being validated along the national borders.

This report describes the results of this pilot NORDGRID phase that ended during autumn 2006. An outlook for a more substantial re-analysis over Fennoscandia for many years is also presented and discussed in this report.

2 Gridding methods used in the participating countries

2.1 DENMARK

The Danish data which is presented in NORDGRID originates from Climate Grid - Denmark. Climate grid - Denmark is various interpolated climate data, that is quadratic area data situated in a fixed defined net (UTM coordinate system zone 32) and covers the surface of Denmark.

The interpolation routines are optimized to Danish climatic conditions and are not suitable in mountainous areas as information of elevations are not included in the calculations.

2.1.1 Method description for precipitation gridding

The number of manually operated precipitation stations in Denmark has since 1990 been approximately 450 ± 50. Geographically the stations are distributed evenly across the country, and they are monitored once a day (at 8:00 AM, local time), see figure 1.

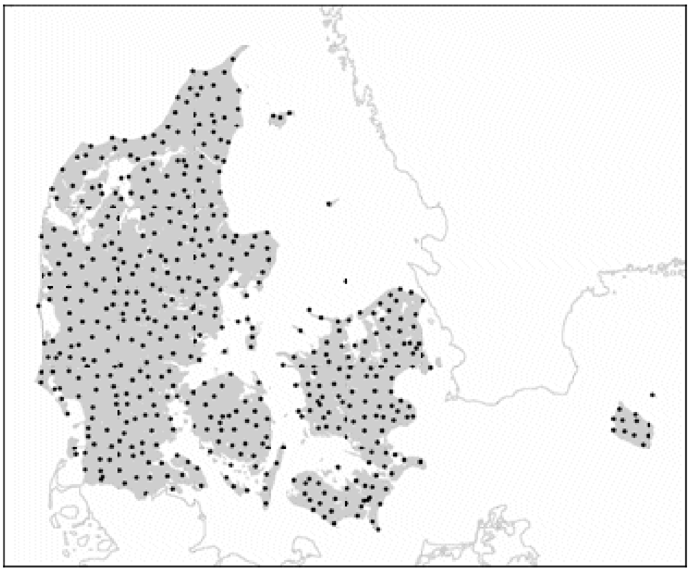


***Figure 1**. Precipitation stations (black dots) in Denmark 2000.*

The algorithm used for interpolating grid values is termed "inverse-distance", and the individual weights are given by:

$$\frac{1}{r^a} \tag{1}$$

where r is the geographical distance between the interpolation point and the station and a is the power.

What determines the value at an interpolated point is thus an average of the surrounding stations, and the importance of the individual stations is weighted with a given power which is calculated on the basis of the distance between the station and the interpolation point.

Based on knowledge about how the daily precipitation is generally distributed in Denmark as well as on the basis of test runs it was decided to use the power of a=2 in equation (1) for weighting the importance of the individual stations. During the interpolation this ensures that stations located close to the grid cell being calculated are weighted comparatively higher than more remote stations.

Instead of using a fixed search radius which would include all stations within a defined distance in the calculation, the program searches in four sectors and uses only the closest station in each sector. Consequently, the stations included in the calculation are the closest ones accessible, and at the same time they are geographically dispersed, see figure 2.

The inverse-distance algorithm is an exact interpolator. This means that if a precipitation station is located at the same (or very close to the) point for which the interpolation is carried out, the interpolated point will assume the station value, even if there are other stations in the neighbourhood with values which are significantly different from that station. This may lead to errors in the interpolation, because it is implicit in the calculation that the point for which the interpolation is carried out is representative of the entire 10x10 km² grid cell in which it is located. In order to avoid this effect the interpolation is carried out for four points within each 10x10 km² grid cell, corresponding to a 5x5 km² grid, and then the points are averaged, see figure 2.

2.1.2 Method description for temperature gridding

The number of weather stations in Denmark which measure temperature every hour or every sixth hour has since 1990 been approximately 65 ± 10. The stations are not distributed evenly across the country, see figure 3.

The algorithm used for interpolating temperature is basically the same as for precipitation. The power in equation (1) is here a=2, for weighting the importance of the individual stations but the program searches in eight sectors instead of only four as in the precipitation analysis. The number of weather stations are small compared to the many precipitation stations and as they also are irregular positioned it constitutes a problem in relation to interpolation. This problem is exacerbated if we try to interpolate a climate parameter such as temperature, humidity or wind speed, because these parameters often change when moving from a coast-dominated to a more inland-dominated climate. Test runs thus proved that it is impossible to obtain a satisfactory result when all station data were used as input for a single interpolation. Some of the grid cells in areas with poor station coverage were allocated interpolated climate values which were not in accord with the climate that was to be expected for that area. We therefore chose to carry out a double interpolation for the three climate parameters mentioned (temperature, humidity and wind speed). In a double interpolation two calculations are made – one that uses coast-dominated stations exclusively, and one that uses inland-dominated stations exclusively. Whether a specific station should be classified as a coast-dominated or an inland-dominated station is determined on the basis of its distance from the ocean and local climate circumstances.

The two interpolation results are then weighted as illustrated in figure 4.

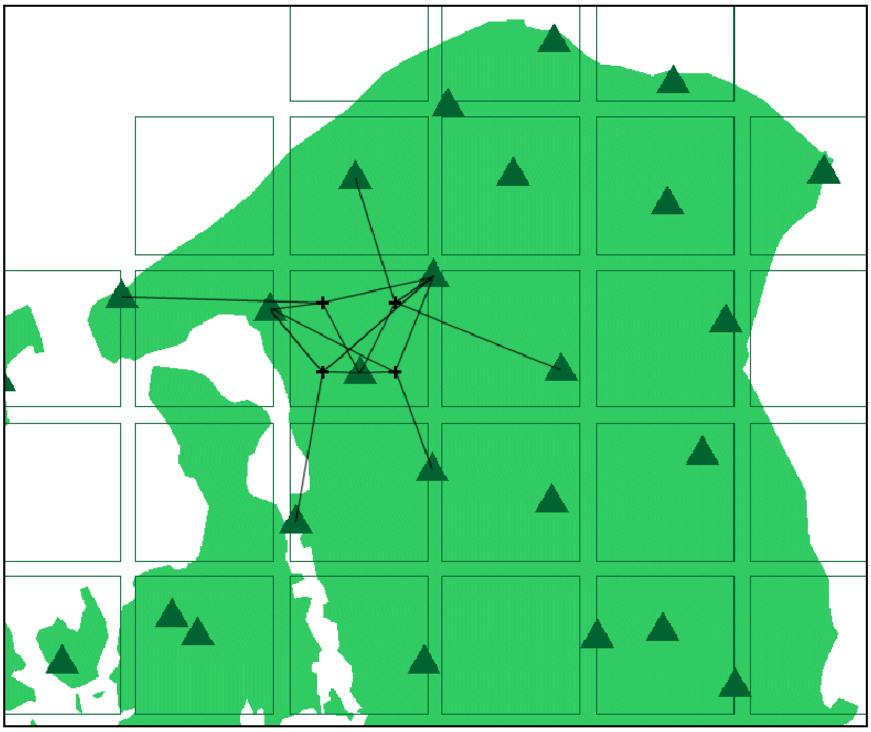


Figure 2. *Example illustrating which stations (dark green triangles) are included in the calculation of a 10x10 km² area grid (the quadratic squares) or four 5x5 km² interpolation points (black crosses) when the program searches in four sectors.*

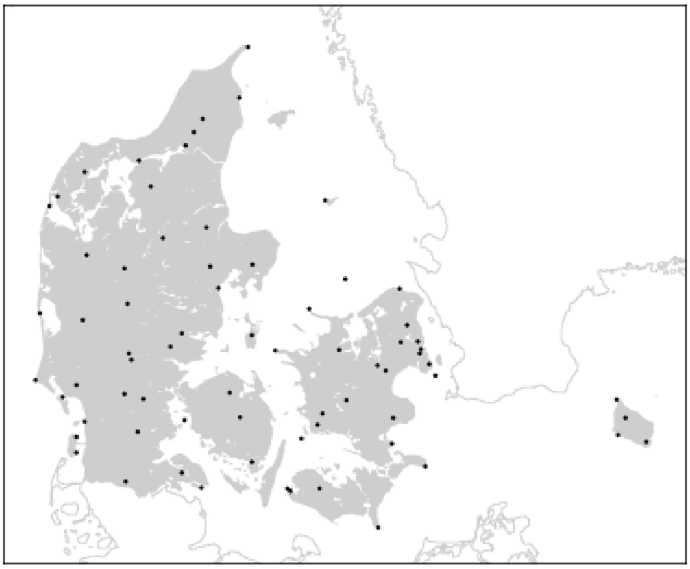


Figure 3. *Weather stations that measures temperature (black dots) in Denmark 2000.*

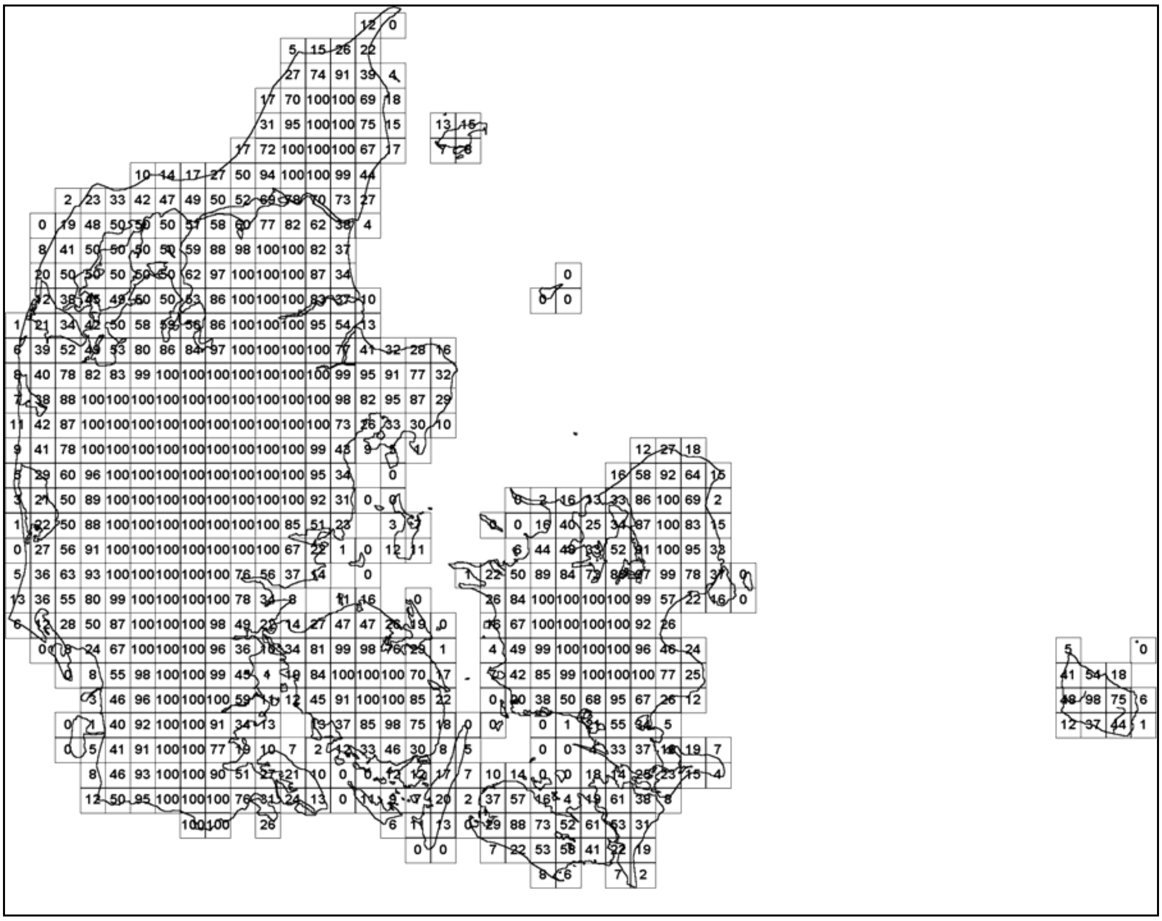


Figure 4. *Proportion of inland-dominated climate in percent for 10x10 km² grid cells.*

Figure 4 shows that in the central parts of Denmark the inland climate is 100 % dominating, whereas in the areas within a range of 10-40 km from the coast the climate begins to become affected by the ocean.

The procedure described above, in which a knowledge-based weighting of the inland-dominated and the coast-dominated interpolation is made, ensures that coastal stations will not have an excessively high weight in areas that are dominated by inland climate and vice versa. Afterwards four 10x10 km² temperature point values are aggregated to a 20x20 km² area grid cell.

2.2 FINLAND

The grid used in Finland is based on the Finnish National Coordinate system known as YKJ. The projection is the so-called Gauss-Kruger with a central-meridian at 27 °N. The westernmost coordinate of the area we have used is 3075000 and the easternmost 3735000. The most southern and the most northern grid squares have coordinates 6635000 and 7785000, respectively. The grid size is 10x10 km². Only those grid squares that are located inside or on the Finnish borders are used in the interpolation, see figure 5.

2.2.1 Meteorological station network

The number of stations making measurements has varied over time. Since 1966 the number of daily precipitation measurements available in the FMI climatological database has exceeded 400, whereas around 1961 there are only 100-200 stations available

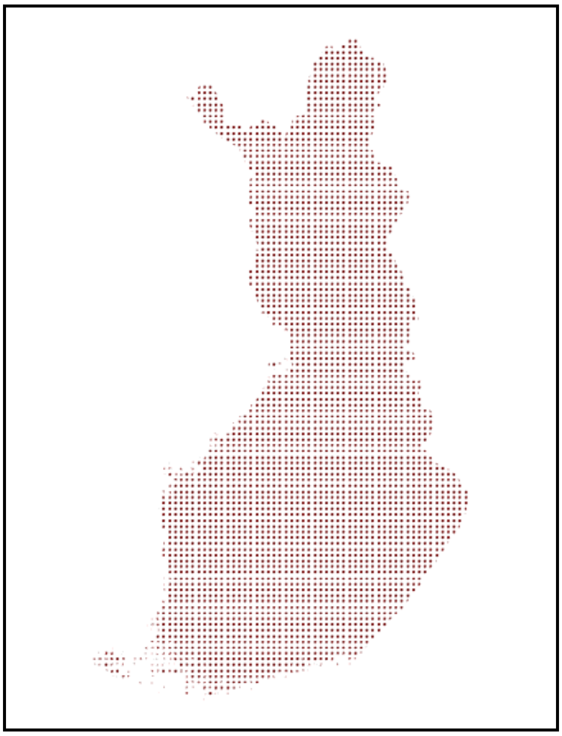


Figure 5. *The grid squares used for interpolated data in Finland.*

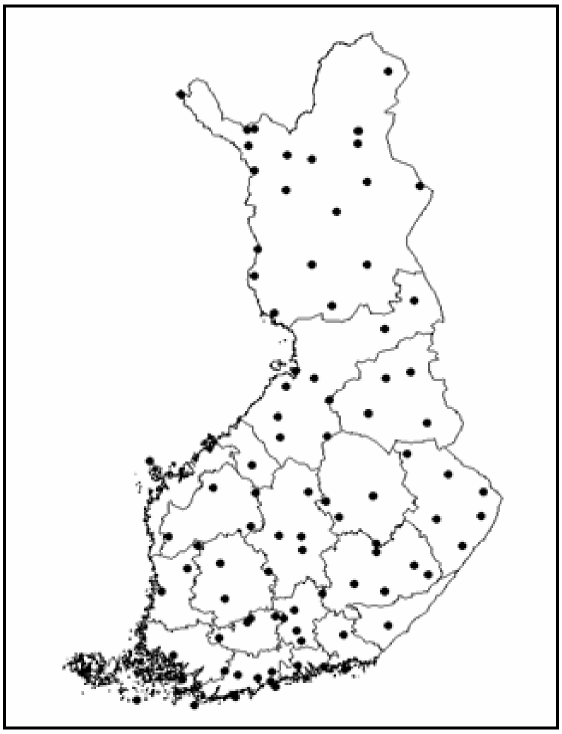


Figure 6. *Precipitation and temperature observations used in NORDGRID.*

in the database. For the comparison in NORDGRID roughly 100 stations situated evenly in different parts of Finland were selected, see figure 6.

2.2.2 The spatial interpolation method

The spatial interpolation method currently used at FMI is known as kriging. This method is based on the theory presented by Ripley (1981) and programmed by Henttonen (1991) for climatological applications in forestry (see also Venäläinen and Heikinheimo, 1997, Venäläinen and Heikinheimo, 2002, Vajda and Venäläinen, 2003). This same method was used for the production of a basic gridded dataset aimed for the use of climate adaptation studies (Venäläinen et al., 2005).

The interpolated values of daily maximum and daily minimum temperature and daily precipitation sum as obtained using the kriging interpolation method were compared (Kuittinen et al., 1999) with results obtained using the interpolation method used in the crop growth monitoring system of the Joint Research Centre. According to that comparison, the results obtained using kriging were better, especially in the case of daily minimum temperature and precipitation. Vajda and Venäläinen (2003), using this same method, interpolated a number of climatological parameters onto a 1x1 km^2 grid in northern Finland.

They discovered that the error, e.g., in the case of average summer daily mean temperature, varied between -0.6 and +0.6 °C with a root mean square error of 0.3 °C. Thus, we may assume that the mean temperature values can be interpolated reasonably well.

Precipitation is more problematic due to large spatial variation especially in the case of rain showers and in these cases the interpolated value may differ remarkably from the measured one. Large spatial variation of precipitation seems to create also systematic differences between measured and interpolated values. The method tends to cut the peak values away and when the measured values were compared with the interpolated values it was found that the interpolated values were systematically lower than the measured values in the cases when the measured precipitation values were above 10 mm/day. If daily precipitation was less than 5 mm then there was no systematic difference between the interpolated and measured values. The data used in this comparison consisted of measurements made at about 70 stations in 1.1.2000-31.12.2000. The spatial variation of monthly precipitation sum values is smaller than that of the daily values. Earlier we have interpolated monthly precipitation sum values using the same method and same grid. When we calculated mean annual precipitation sums for each grid square based on 1.1.1991-31.12.2000 data, it was found that the values were on average 17 % higher if the monthly data was used as input in the interpolation. In this sense it is good to remember that if long term mean values are calculated based on interpolated daily precipitation data they tend to be systematically smaller compared with the case when monthly or annual data is used as input in the interpolation.

2.3 NORWAY

2.3.1 Method description for temperature gridding

In Norway a residual interpolation approach is applied to estimate daily high resolution temperature grids. The method is the same as used for deriving the Nordic monthly mean temperature maps (Tveito et al., 2000).

Residual interpolation is based on the assumption that the spatial distribution of temperature can be described by two components; a deterministic and a stochastic (Tveito and Førland, 1999). It is known that temperature is highly influenced by local and regional physiographical conditions like terrain, and vicinity to the sea or lakes. This influence, denoted as the deterministic component can be described by establishing relations between observed temperatures and variables describing these physiographical characteristics. The relations can be established a priori, or by statistical methods like linear regression. When removing the influence of these variables from the observed temperatures, the remaining residuals can be regarded as a stochastic field. These values are analysed by e.g. geostatistical methods, and the spatial interpolation of the residuals might be done by any spatial interpolation method.

Within the NORDKLIM project Tveito et al. (2000) established models for Fennoscandia for deriving mean monthly temperature maps for the current standard normal period 1961–1990. These models are based on an investigation of mean monthly temperatures from 1152 stations in Denmark, Finland, Norway and Sweden, and the deterministic component includes five independent predictors:

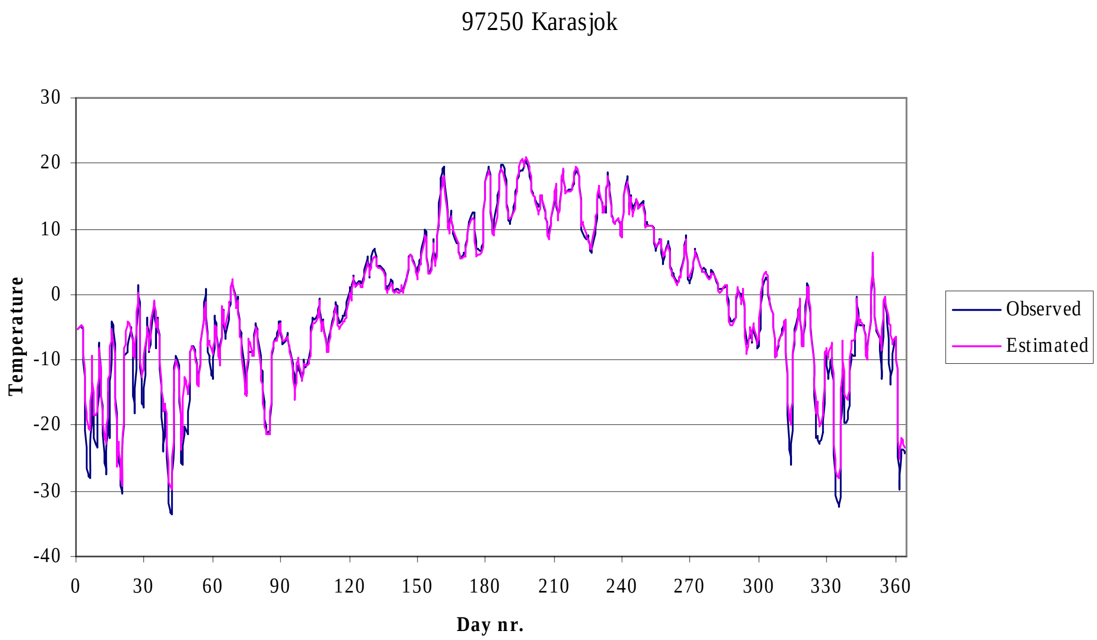
- Altitude of grid cell or point to be estimated
- Mean altitude within a 20 km circle
- Lowest altitude within a 20 km circle
- Latitude
- Longitude

Models are established for each month using linear regression, and the coefficients and their significance show large variations between the different months and seasons.

For establishing the daily grids of temperature the same method is applied. Despite developed for monthly mean values, it shows also to perform satisfactorily for most cases. The method fails however during some specific weather types, like e.g. inversion situations. Figure 7 shows the performance of the model for a single year at the weather station Karasjok in northern Norway, located in a shallow valley at Finnmarksvidda. This station is heavily exposed to inversions, and it can be seen that the model fails considerably when temperature is below -10 °C. Being based on mean monthly temperatures, the model coefficients are not estimated from a sample representing such conditions. To be able to describe such cases properly, the model has to be based on a dataset stratified to describe weather conditions like that. Tveito (2007) describes an attempt to make such a stratification applying weather types as conditional information.

2.3.2 *Method description for precipitation gridding*

Gridding of daily precipitation sums for Norway is a rather tricky task. The Norwegian precipitation climate is characterised by large variations and steep gradients due to the complex topography and vicinity to the sea. In some parts of the country mean annual precipitation varies between less than 300 and more than 3000 mm in just a few tenths of



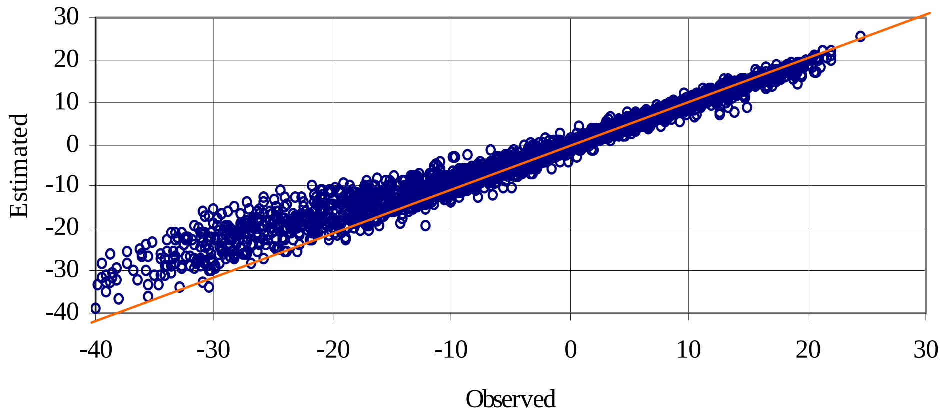


***Figure 7**. Observed and estimated temperatures (°C) at Karasjok, northern Norway in 1997. Time series for observed and estimated temperature (top) and scatter plot for observed versus estimated temperature (bottom), showing the large bias for the lower temperatures.*

kilometres. Precipitation distribution in Norway is a combination of two steep gradients; the vertical gradient where precipitation is expected to increase by 8-10% for each 100 m in altitude, and the horizontal gradients from the western coast with its maximum zone between 50 and 100 km from the coastline.

Spatial interpolation of precipitation is not a single problem, but rather a two-step procedure. First the occurrence of precipitation or not has to be estimated, and then, if precipitation is expected to occur, the amount of precipitation.

When introducing a new snow map service in 2003, met.no had to find a quick solution for estimating gridded precipitation. As a first approach triangulation was chosen. The reason for this is that it is a fairly quick method, and that it also in a fairly easy manner could include a transparent estimation of the vertical precipitation gradient.

The method of triangulation is originally a method developed to model surfaces, and creates a set of adjacent, none overlapping triangles between observation points (figure 8). The input are the x and y coordinates of the observation points and a z value, which for precipitation interpolation is the precipitation amount.

The interpolation process consists of several steps.

Step 1: Establishing a TIN (triangular irregular network) of precipitation based on all points with observations of precipitation.
Step 2: Transforming the established TIN to a regular precipitation grid.
Step 3: Establishing a new TIN based on the same points as in step 1, but using altitude as z-value.
Step 4: Transforming the altitude TIN into a regular grid.
Step 5: Take the difference between the grid derived in step 4 and a digital terrain model to find the difference between the terrain described by the precipitation surface derived in step 1 and 2 and the real terrain.
Step 6: Combine the terrain difference grid and the vertical precipitation gradient (10%/100 m) to adjust the precipitation grid derived in step 2.

The procedure is also shown graphically in figure 9.

One disadvantage with the triangulation technique is that it is limited to cover the area between observation points. That means that areas outside defined triangles will be cut off. It is therefore not able to estimate values outside these triangles. The technique is therefore purely an interpolation technique not allowing extrapolation. The technique is probably going to be abandoned by met.no in the near future, as a new method allowing anisotropy and local parameterisation is under development.

2.4 SWEDEN

The gridding method developed and used at SMHI is called MESAN (Häggmark et al., 2000). The MESAN model is based on optimal interpolation (OI) where a background field (also called a first guess field) is modified by irregular distributed observations. In OI the structure functions (background error correlation functions) plays an important role in the analysis. The structure functions have been calculated from a large statistical dataset for each analysed parameter separately in MESAN. For example, the structure functions for precipitation analyses are reflecting the precipitation climate, i.e. precipitation observations from dry areas will not reduce the amount of precipitation in wet areas. Also the accumulation

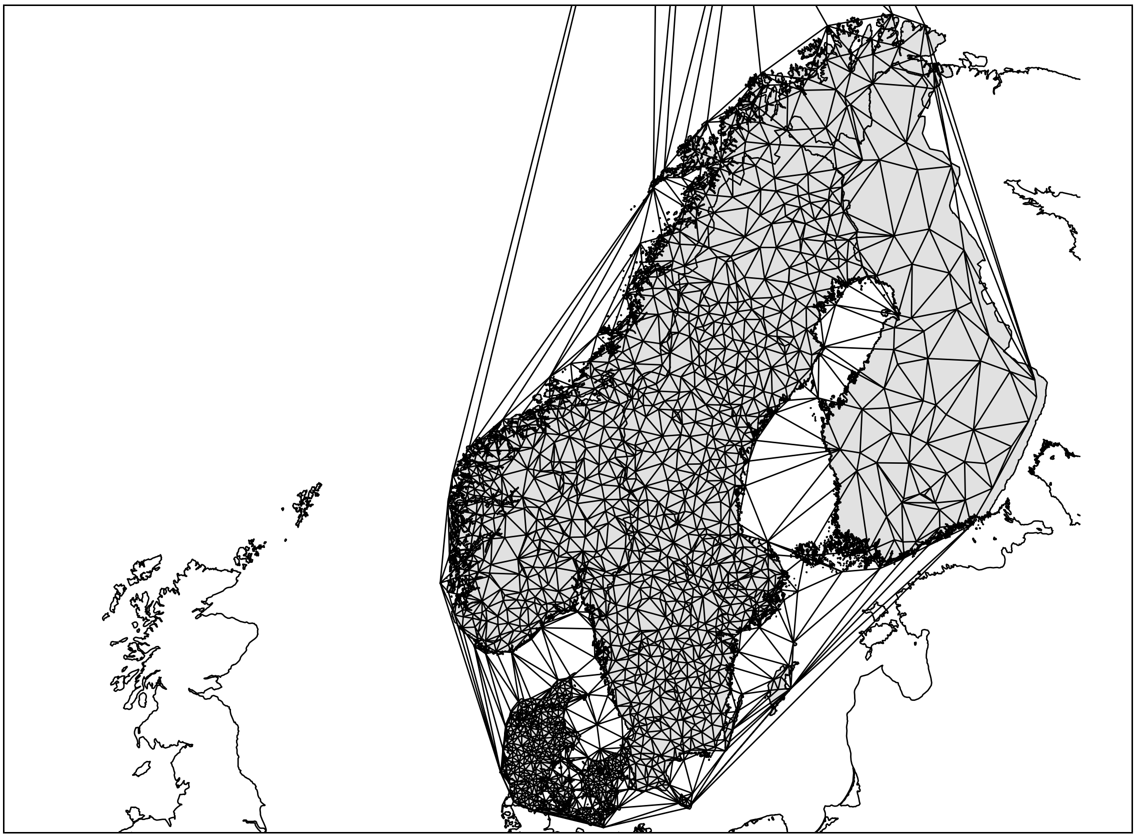


Figure 8. *TIN (triangular irregular network) for the NORDGRID precipitation network. Each triangle node represents a precipitation observation station.*

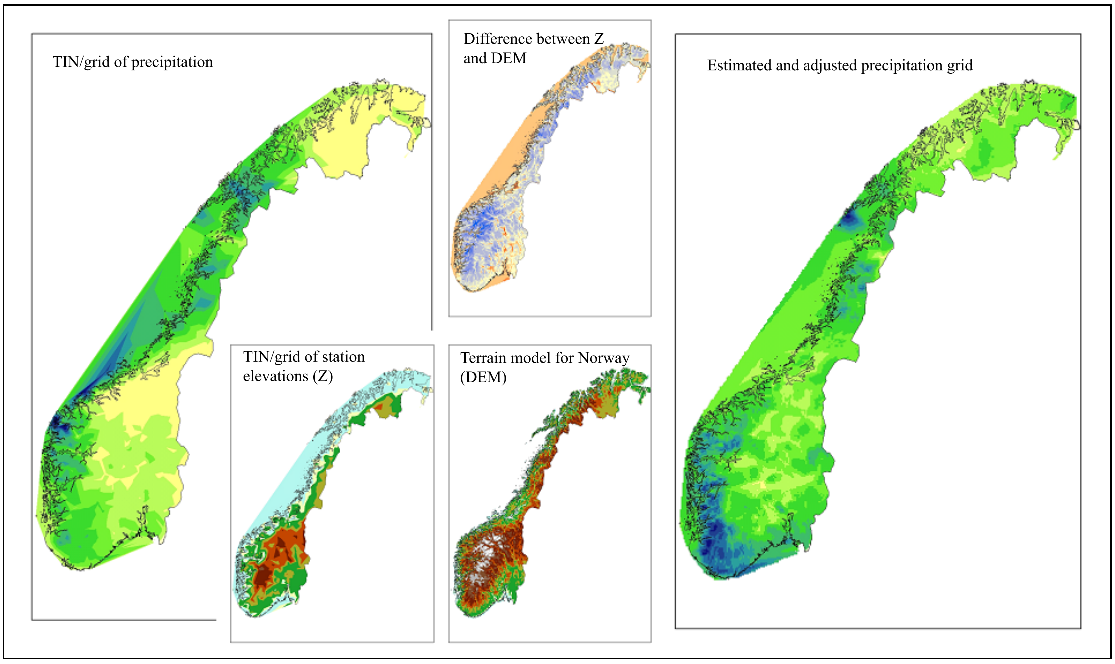


Figure 9. *The procedure for deriving precipitation grids.*

time is considered in the structure functions. The first guess field in MESAN is a HIRLAM-forecast (HIgh Resolution Limited Area Model) (Undén et al., 2002) and the observations that are used are all available synoptic observations from the whole MESAN area (figure 10), automatic weather observations in Sweden, observations from the Swedish Road Authorities, and climate observations in Sweden. Also radar and satellite pictures are converted into observations and used in the precipitation and cloud analyses.

MESAN has been operational at SMHI since 1998 and cover northwest Europe (figure 10). The MESAN model has a spatial resolution of 22 km and 11 km and the analyses are made on hourly and 3 hour basis. Examples on analysed parameters in MESAN are 2 m temperature, 2 m max- and min-temperature, 2 m wet bulb temperature, pressure at mean sea level, 2 m relative humidity, visibility, 10 m u- and v-component of the wind and gust, total cloud cover, low cloud cover, cloud top and base, precipitation 1, 3, 12 and 24 hours accumulated and fresh snow, 1 and 3 hours accumulated.

The MESAN-analyses are directly used to guide forecasters but it also provides input to other models such as, radiation, dispersion, fire risk and atmospheric chemistry models.

In this project, the creation of a joint dataset that cover Fennoscandia was decided for the years 1995-1999 (module B). These years are not available from MESAN, instead a different application of MESAN have been used in the NORDGRID-project. This application is called the ERAMESAN dataset (Jansson et al., 2007). The aim of the ERAMESAN-project was to try to create a time consistent dataset that covers most of Europe and has a fine spatial resolution and also cover a relatively long time period. The dataset was to be used in atmospheric chemistry modelling for air quality studies over Europe. In MESAN the first guess field, i.e. a HIRLAM-forecast has changed model version, geographical area, and spatial resolution and so on during the years and is only valid for some years back in time, the same holds for MESAN. In the ERAMESAN-project a pilot study was made to find out the best first guess field that was appropriate to use in MESAN for a long time period that is time consistence and covers most of Europe. The first guess field chosen was the ERA-40 re-analysis (Uppala et al., 2005) from the European Centre for Medium-Range Weather Forecasts (ECMWF). The observations used in the ERAMESAN dataset are selected from the observations archive at SMHI. All observations from Sweden are quality controlled while foreign observations are not quality controlled. The ERAMESAN-project was limited to not include any further development of the model to fit ERA-40 instead of HIRLAM-forecasts as first guess field.

The analysed parameters in ERAMESAN are 2 m temperature and wind at 00, 06, 12 and 18 UTC, 12 hour accumulated precipitation at 06 and 18 UTC and 24 hour accumulated precipitation at 06 UTC with a spatial resolution of 11 km. The first guess field in the ERAMESAN temperature and wind analyses are ERA-40 analyses with a 125 km spatial resolution that have been interpolated to 44x44 km^2. In the ERA-40 analyses all available synoptic observations at ECMWF have been used as input expect in the wind analyses, here only observations over the sea have been used. For the precipitation analyses an ERA-40 forecast has been used, since precipitation analyses not exists in the ERA-40 dataset. The ERAMESAN dataset covers the years 1980-2004, and the area (figure 11) is 4.3 times larger than the MESAN area (figure 10).

Figure 10. *The MESAN area.*

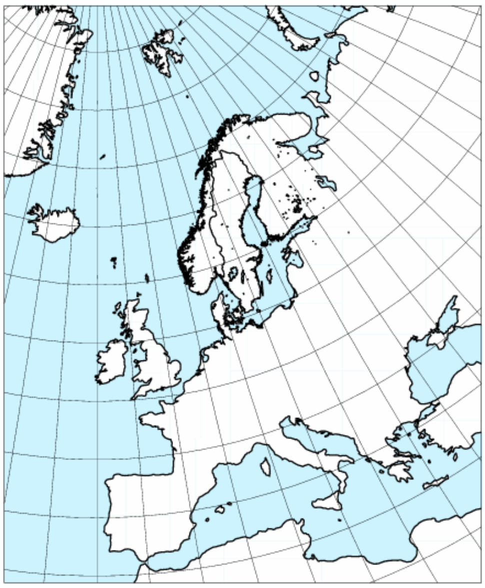


Figure 11. *The ERAMESAN area.*

3 Module A: Comparison of gridding methods used in the participating countries

The grid methods in Norway, Finland and Sweden have been applied over Fennoscandia for the year 1999 and then been compared against independent observations. The Danish method has not been applied over Fennoscandia partly because of the coarser grid resolution in the Danish temperature analysis, but also because of the difficulties in applying the model over mountainous areas since the elevation is not included in the analysis algorithm (chapter 2.1). Usually Norway and Finland only makes gridded maps over its own land area but in this comparison the whole Fennoscandia is analysed. The parameters that have been analysed and compared against observations are 24 h accumulated precipitation and 24 hour daily mean temperature.

3.1 Observations used in the validation

In this validation all countries have used the same observations from Denmark, Finland, Norway and Sweden that are quality controlled at each institute. The difficulties in comparing daily mean temperatures from several countries are that each country has its own daily mean temperature algorithm, also the time period for the daily mean temperature to be valid can differ. A short description of the daily mean temperature algorithms applied in each country is given below. An alternative algorithm that can be applied in all participating countries is also described.

3.1.1 Daily mean temperature algorithms

3.1.1.1 Temperature algorithm in Denmark

The temperature algorithm in Denmark differs depending on the number of observation times during day and night. All hours are given in UTC.

1) For stations with hourly observations (automatic stations) the daily mean temperature is calculated as:

$$T_m = (T_{00} + T_{01} + \dots + T_{23})/24 \qquad (2)$$

Equation (2) is used if missing values in the first place are interpolated using a calculation routine. If more than 3 observations still are missing the daily mean temperature is calculated as:

$$T_m = (T_{max} + T_{min})/2 \qquad (3)$$

If T_{max} and T_{min} are present.

2) For synoptic stations with 8 observations the daily mean temperature is calculated as:

$$T_m = (T_{00} + T_{03} + T_{06} + T_{09} + T_{12} + T_{15} + T_{18} + T_{21})/8 \qquad (4)$$

If more then 2 observations are missing the value will be calculated as in equation (3) if T_{max} and T_{min} are present. Otherwise, if T_{max} or T_{min} or both values are missing the daily mean temperature will not be calculated.

3) For climate stations with 3 observations the daily mean temperature is calculated as equation (5), if all observations are present:

$$T_M = M - C \qquad (5)$$

Where:

$M = (T_{07} + T_{13} + T_{20})/3$
$C = (s \cdot (dt07 + dt13 + dt20)/3) \cdot (dt13 - (dt07 - dt20)/2)$
$s = T_{13} - (T_{07} + T_{20})/2$

The values of *dt* depends on the month:

	*dt*07	*dt*13	*dt*20
Jan	-0.33	0.63	-0.10
Feb	-0.63	1.38	-0.24
Mar	-0.66	2.00	-0.45
Apr	-0.64	2.54	-0.73
May	-0.56	3.03	-0.76
Jun	-0.16	2.84	-0.75
Jul	-0.08	2.86	-0.76
Aug	-0.28	2.90	-0.99
Sep	-0.28	2.41	-0.88
Oct	-0.43	1.69	-0.48
Nov	-0.31	0.88	-0.24
Dec	-0.26	0.59	-0.11

If some of the 3 observations are missing the mean temperature is calculated as in equation (3) if $T_{\max}$ and $T_{\min}$ are present.

3.1.1.2 Temperature algorithm in Finland
Daily mean temperatures are calculated in the following order depending on number of observations / day:
1) When number of observations is 8:

$$T_m = (T_{00} + T_{03} + T_{06} + T_{09} + T_{12} + T_{15} + T_{18} + T_{21})/8 \qquad (6)$$

2) When number of observations is 4:

$$T_m = (T_{00} + T_{06} + T_{12} + T_{18})/4 + TCOR(month) \qquad (7)$$

where *TCOR* is a correction factor depending on the month:

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
-0.04	0.01	0.17	0.00	-0.12	-0.07	-0.15	-0.11	0.08	0.05	-0.03	-0.03

3) Otherwise the so-called Kolkki's formula is used when night temperature T_{00} is missing.
For January, February and October-December:

$$T_m = (T_{18pre} + T_{06} + T_{12} + T_{18} + T_{06next})/5 + TCOR(month) \qquad (8)$$

Otherwise, for March-September:

$$T_m = \left(T_{18pre} + T_{06} + T_{12} + T_{18}\right)/4 + TCOR(month)\cdot\left(T_{18pre} - T_{\min 06}\right) + 0.1 \qquad (9)$$

Where:
T_{18pre} = temperature at 18 UTC on the preceding day.
T_{06next} = temperature at 06 UTC on the next day.
T_{min06} = minimum temperature at 06 UTC.

In equation (8) and (9) *TCOR* is also a correction factor depending on the month:

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
-0.01	0.05	-0.114	-0.180	-0.227	-0.227	-0.237	-0.216	-0.136	0.06	0.01	-0.01

3.1.1.3 Temperature algorithm in Norway

Different formulas were used in Norway during the period 1876 to 1889, but from 1890 a formula attributed to Köppen was introduced. Later T_m was recalculated by Köppen's formula also for the period 1876-1889. Thus from 1876 to present the same formula has been used in Norway, equation (10).

$$T_m = T_f - k\left(T_f - T_n\right) \qquad (10)$$

Here T_f is the mean of the three observations at fixed hours (morning, midday, and evening), T_n is the daily minimum temperature, and k is the so called Köppen's constant. The magnitude of k depends on the location, month, and the time of the observation. The k-values were calculated from hourly observations in Oslo, Bergen, Trondheim, Alta, Vardø, and Spitsbergen. For the other stations the k-values was established by map interpolations. There is an ongoing project at The Department of Climatology on testing of the old k-values by use of the recently established network of automatic stations.

Since 1876 observations have been taken at following hours. (A few stations have had a different observational program)

Jan 1876	-	Jun 1910:	8, 14 and 20 local time
Jul 1910	-	Jun 1920:	7, 13 and 19 UTC
Jul 1920	-	Dec 1938:	7, 13 and 18 UTC
Jan 1949	-	Jun 1949:	7, 12 and 18 UTC
Jul 1949	-	Present:	6, 12 and 18 UTC

Thus k had to be changed in 1920, 1938, and two times in 1949. Minor changes also took place in 1894 when a new procedure for observing minimum temperature was introduced. With the present observation hours 06, 12 and 18 UTC, is the k-value ≤ 0.25, and in northern Norway $k \leq 0.2$. Maximum occurs in midsummer, minimum in midwinter.

3.1.1.4 Temperature algorithm in Sweden

In Sweden the daily mean temperature is calculated with a formula called the Ekholm-Modéns fomula, equation (11), and is valid for 00-00 UTC. All hours are given in UTC.

$$T_m = aT_{06} + bT_{12} + cT_{18} + dT_{\max} + eT_{\min} \qquad (11)$$

Here *a*, *b*, *c*, *d*, *e* are constants depending on month and longitude.
If the T_{12} observation is missing it is interpolated according to:

$$T_{12} = aT_{06} + bT_{18} + cT_{\max} + d \qquad (12)$$

Here *a*, *b*, *c*, *d* are other constants depending on the month and the latitude (not the longitude). The coefficients have been calculated for longitude 11-25 ºE and latitude 55-69 ºN. The coefficient in equation (11) has been recalculated when observation times have changed and before 1914 another formula was used (Nordli et al., 1996).

3.1.1.5 A common temperature algorithm for all participating countries

As seen in part 3.1.1.1-3.1.1.4 the algorithm for daily mean temperature in the participating countries differs a lot, if the already existing daily mean temperatures should be used in different gridding methods that covers all countries, discrepancies may occur along the national borders. Thus, a common temperature algorithm has to be used, to be able to compare different gridding methods. The daily mean temperature is here valid for 06-06 UTC, and originates from Førland and Tveito (1997).

$$T_m = \left(T_{06_I} + 2T_{18_I} + T_{06_{II}} + 2T_{\max} + 2T_{\min}\right)/8 \qquad (13)$$

Here $T_{\max}$ and $T_{\min}$ consider the maximum and minimum temperature during the time interval 06-06 UTC. *I* and *II* indicates from which day during 06-06 UTC the observation originates from. If some observation is missing, apart from $T_{\max}$ and $T_{\min}$, the mean temperature is instead calculated as:

$$T_m = \left(T_{\max} + T_{\min}\right)/2 \qquad (14)$$

So in the first place equation (13) has been used in all participating countries and in the second place equation (14). Each country has done this calculation for the temperature and then distributed the common calculated daily mean temperature to the other countries. Figure 12 shows the total amount of calculated daily mean temperatures in all participating countries for the 1. August 1999.

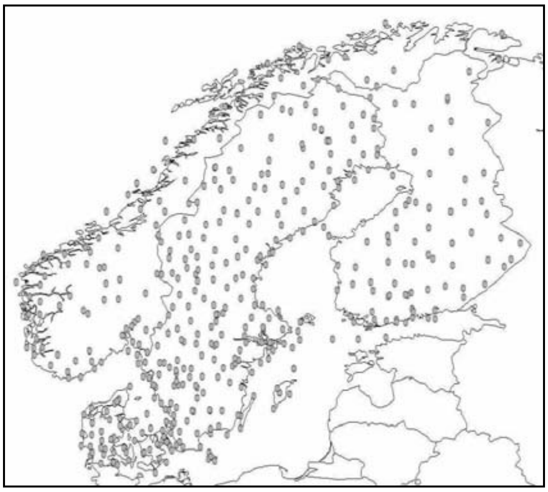


***Figure 12**. All available calculated daily mean temperatures at the 1. August 1999.*

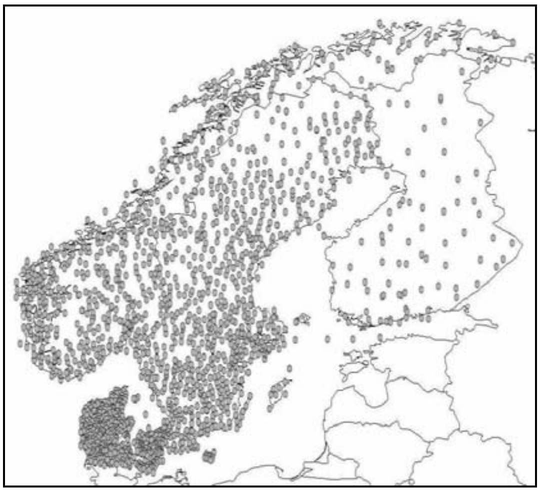


***Figure 13**. All available daily precipitation observations at the 1. August 1999.*

3.1.2 *Daily precipitation*

The daily precipitation observations are the same in Sweden and Finland. They are both observed at 06 UTC. The Danish observations available in this project are only the climate observations, here the daily precipitation are measured 07 UTC. The Norwegian observations are a mixture of observations at synoptic stations at 06 UTC and climatic stations at 07 UTC. The one hour delay in the Danish and Norwegian observations has a minor impact on the validation. Therefore also the 07-07 UTC observations are used. Figure 13 shows all precipitation observations used in the validation for the 1. August 1999.

3.2 MODEL SETUPS

Since a new daily mean temperature algorithm has been applied, some adjustments in each model had to be done. For Finland and Norway the model also had to be applied on a new larger area. A short description will be given for all adjustments and changes that have been done for all participating models.

1) Model adjustments in Finland

In the Finnish model the three main interpreters are:

- elevation
- sea percentage in grid point (10x10 km² area)
- lake percentage in grid point

For the interpolation in Finland the routine interpolation system was used and no adjustments were needed.

The background data for Fennoscandia were integrated to 10x10 km² grid using the underlying datasets and the ArcView software.

- Sea and lake percentages were taken from the "Global Land Cover 2000 database", European Commission, Joint Research Centre, 2003, http://www-gem.jrc.it/glc2000.
- Elevation calculation was based on Globe dataset from ftp-site: ftp://ftp.ngdc.noaa.gov/GLOBE_DEM/data/elev/esri .

2) Model adjustments in Norway

Daily mean temperatures for 06-06 UTC were estimated using the algorithm presented in equation (13). The calculation uses all available observations at any time. Beside that no other adjustments were necessary for calculating the national grids.

For producing the grids for all the four countries, the necessary topographic information was derived from the GTOPO30 terrain model (USGS, 1996) for the entire Fennoscandia region. In addition the analysis window (the geographical extent of the grids) had to be increased.

3) Model adjustments in Sweden

In Sweden the first guess field for daily mean temperature had to be changed, since the new daily mean temperature is valid 06-06 UTC and otherwise from 00-00 UTC. The reading of observations has been changed to the new observations distributed from each country.

3.3 VALIDATION

The stations that have been used as validation stations are the EWGLAM-stations (European Working group for Limited Area Models) used in verification of weather prediction models

(Bjørge et al., 2003). The stations placed in Fennoscadia are shown in figure 14. In the appendix also a list of all EWGLAM-stations in Fennoscandia is given. In the precipitation validation the Danish EWGLAM-stations are excluded, since only climate observations from Denmark are valid. The ambition for all who participated was to neglect one station at the time of the EWGLAM-stations and then compare the results between the participating countries. However, due to lack of time all the EWGLAM-stations were neglected during the entire analysis cycle in each country. The routines for neglecting observations in the analysis cycle were unfortunately different in each country, so the number of neglected stations from the EWGLAM-stations is not the same. The differences in neglected observations in the validation between Sweden and Norway are only one station for both temperature and precipitation, so this difference is probably negligible in the comparison. But the numbers of neglected observations in the Finnish validation are instead about 14 stations fewer for temperature and about 8 stations fewer for precipitation. Therefore the results from Finland are not directly comparable with the results from Norway and Sweden. Also the first guess fields in the daily temperature grids from Sweden include observations and are therefore not totally independent.

3.4 RESULTS

Statistical calculations between the EWGLAM-stations and the gridded data from Finland, Norway and Sweden have been done for daily and monthly values. Also some attention has been paid on the extreme values both for temperature and precipitation. Calculated quantities are MBD (mean bias deviation), MAD (mean absolute deviation) and RMSD (root mean square deviation).

3.4.1 Daily mean temperature

The daily mean temperature from the EWGLAM-stations has been validated against the gridded data from Finland, Norway and Sweden. In all gridding procedures the EWGLAM-stations have been neglected and not used as input. All calculated quantities (MBD, MAD and RMSD) have been calculated as observation minus estimated value. In figure 15, MBD, MAD and RMSD have been calculated for all EWGLAM-stations during 1999, both for the daily

***Figure 14**. The EWGLAM-stations used as validation stations.*

mean temperature (left) but also for the monthly mean temperature (right) (the monthly mean of the daily mean temperatures). The MBD in the daily mean temperature is for Finland 0.43 ºC, for Norway 0.16 ºC and for Sweden 0.17 ºC. The largest differences between the models are seen in the RMSD both for daily (figure 15, left) and monthly (figure 15, right) values. Since it is the monthly mean values that are validated and not the monthly sums it is natural that the MAD and RMSD is much lower for the monthly values (figure 15, right) compared to the daily values (figure 15, left) and that the MBD is about the same for both daily and monthly values.

The MBD, MAD and RMSD have also been calculated for each month separately (figure 16), here it is clearly evident that the large errors from Finland seen in figure 15 only occurs during October-December. A common feature for all countries gridded data is that the largest MBD (figure 16, top) occurs during the summer months (expect from the large errors in the Finnish data during October-December), i.e. the gridded data from, Finland, Norway and Sweden are colder than the observations.

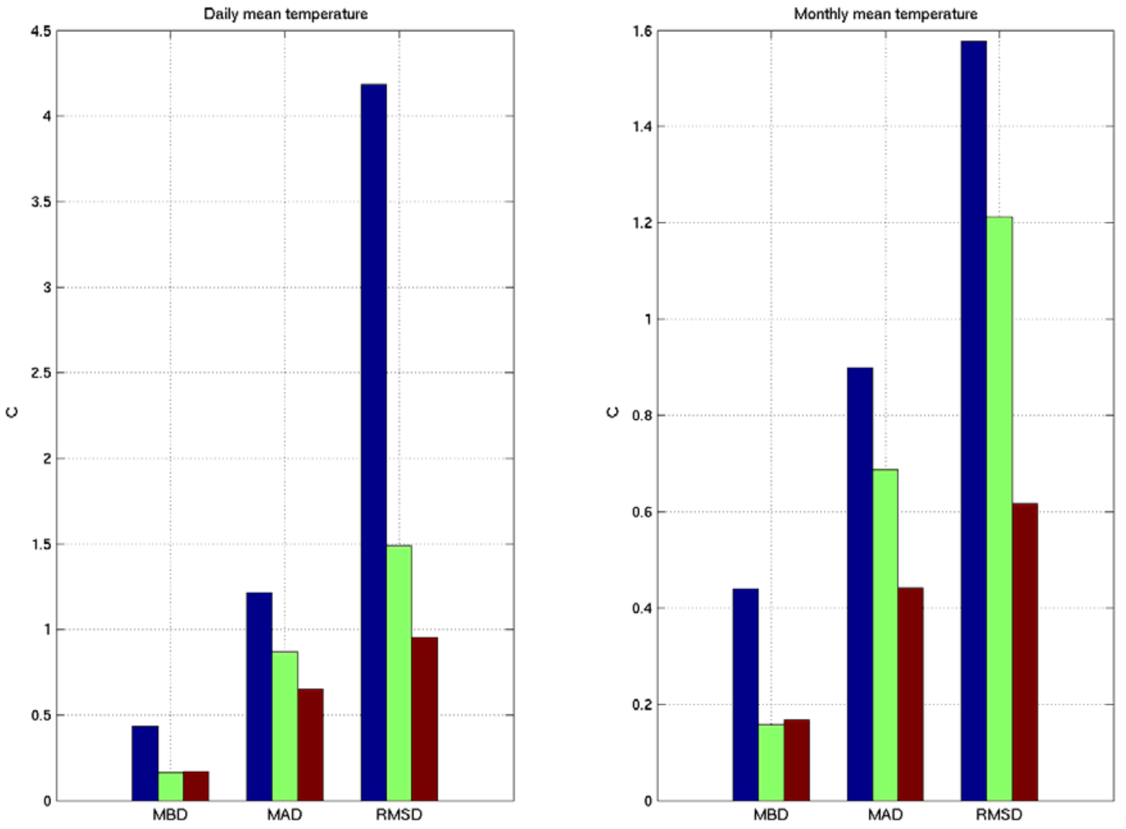


Figure 15. *Histograms for MBD, MAD and RMSD for the daily mean temperature (left) and for the monthly mean temperatures (right) during 1999. The calculations have been done between the EWGLAM-stations and the gridded data from Finland (blue bars), Norway (green bars) and Sweden (red bars).*

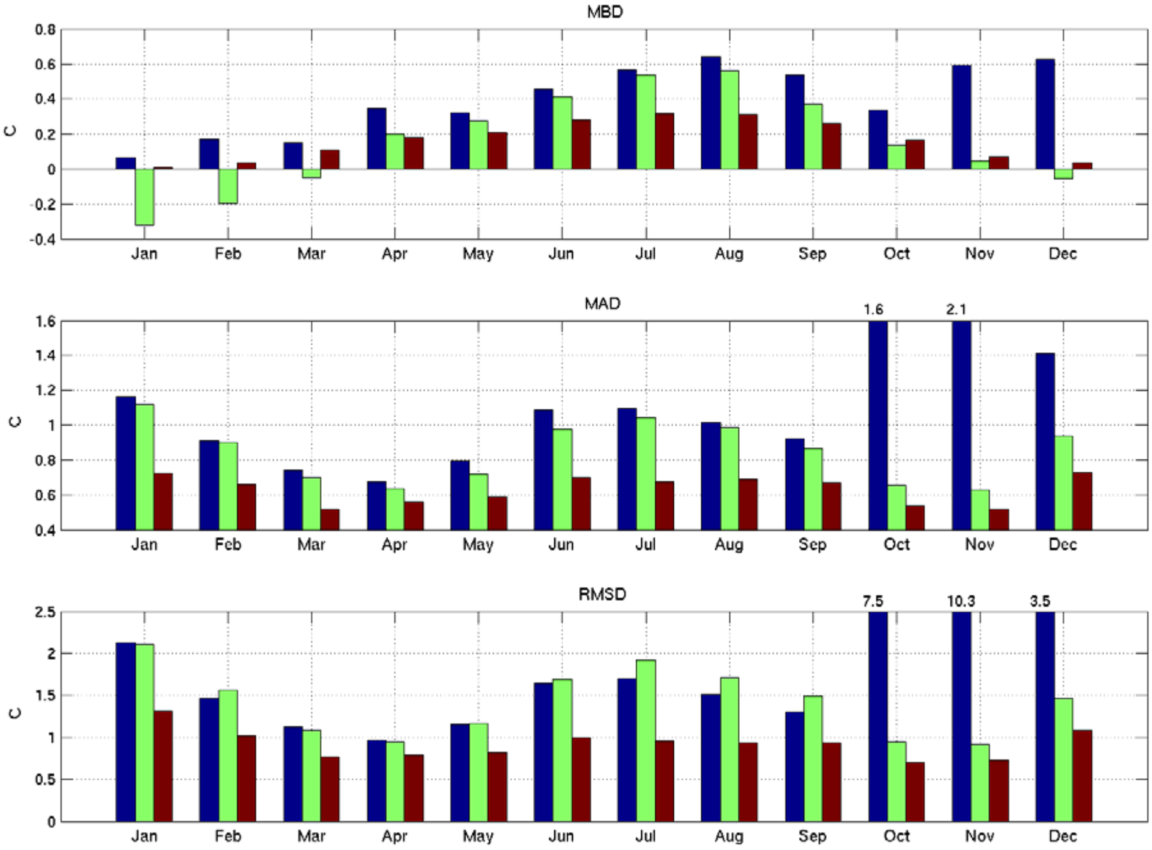


Figure 16. *Histogram for MBD (top), MAD (middle) and RMSD (bottom) calculated for each month separately during 1999. The calculations are done between the EWGLAM-stations and the gridded data from Finland (blue bars), Norway (green bars) and Sweden (red bars). For the MAD (middle) the blue bars for Oct reaches 1.6 ºC and for Nov 2.1 ºC. For the RMSD (bottom) the blue bars for Oct reaches 7.5 ºC, for Nov 10.3 ºC and for Dec 3.5 ºC.*

The MAD (figure 16, middle) is for all months lowest in the Swedish data and highest in the Finnish data. Also for the RMSD (figure 16, bottom) the lowest values originates from Sweden and the highest from Finland (during 6 months) and from Norway (during 6 months).

Some checks have also been done on the extreme values, the 5 lowest and highest temperature observations in each month during 1999 have been validated separately against the gridded data from Finland, Norway and Sweden (figure 17). As expected the gridded data are colder (positive MBD) when the maximum temperatures (figure 17, left) are validated and warmer (negative MBD) when minimum temperatures (figure 17, right) are validated. In both maximum and minimum temperatures Sweden has the smallest magnitude in MBD, MAD and RMSD, and Finland has the largest ones.

3.4.2 Daily precipitation

The daily precipitation amounts from the EWGLAM-stations have been validated against the gridded data from Finland, Norway and Sweden, in all gridding procedures the EWGLAM-stations have been neglected and not used as input. All calculated quantities (MBD, MAD and RMSD) have been calculated as observed minus estimated value.

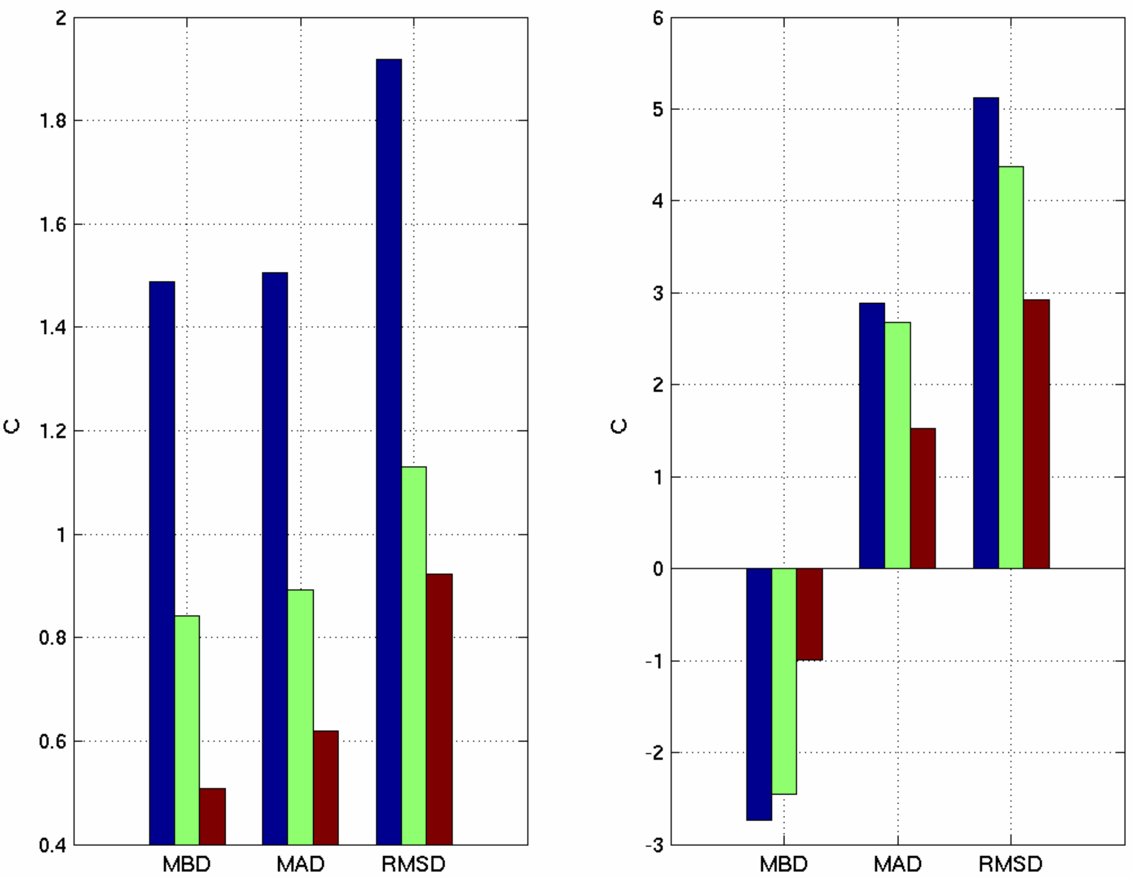


Figure 17. *MBD, MAD and RMSD have been calculated between the 5 highest (left) and the 5 lowest (right) observed daily mean temperatures in each month and the corresponding gridded data from Finland (blue bars), Norway (green bars) and Sweden (red bars).*

In figure 18, MBD, MAD and RMSD have been calculated for all EWGLAM-stations during 1999, both for the daily precipitation amounts (left) and for the total monthly precipitation amounts (right) from each observation (summarized from the daily precipitation). A common feature for all gridded data is that the estimates give higher values then the observations, i.e. the MBD is negative. It is well known that gridded estimates often gives higher values then the observations for small precipitation amounts and the opposite for large precipitation amounts, because of the difficulties in catching large local precipitation amounts. In the precipitation validation the differences between the models are not that significant as in the temperature validation. The lowest magnitude in the MBD for both daily (figure 18, left) and monthly (figure 18, right) values are found in the Norwegian model and the largest magnitude in the MBD is found in the Swedish model. The smallest MAD and RMSD are found in the Swedish model. But the differences between the models are very small.

In figure 19, the MBD, MAD and RMSD have been calculated for each month separately during 1999. The smallest error, MBD, MAD and RMSD, occurs for each model during spring time (January - May). The largest MBD is in most of the months found in the Swedish model and the smallest MBD is shifting between the Norwegian and Finnish model. From the MAD and RMSD the differences between the models are not that significant as for the temperature. The lowest values are shifting between the models, in most of the months (8 months) the lowest MAD is found in the Swedish model. The RMSD is for 7 month lowest in the Swedish model.

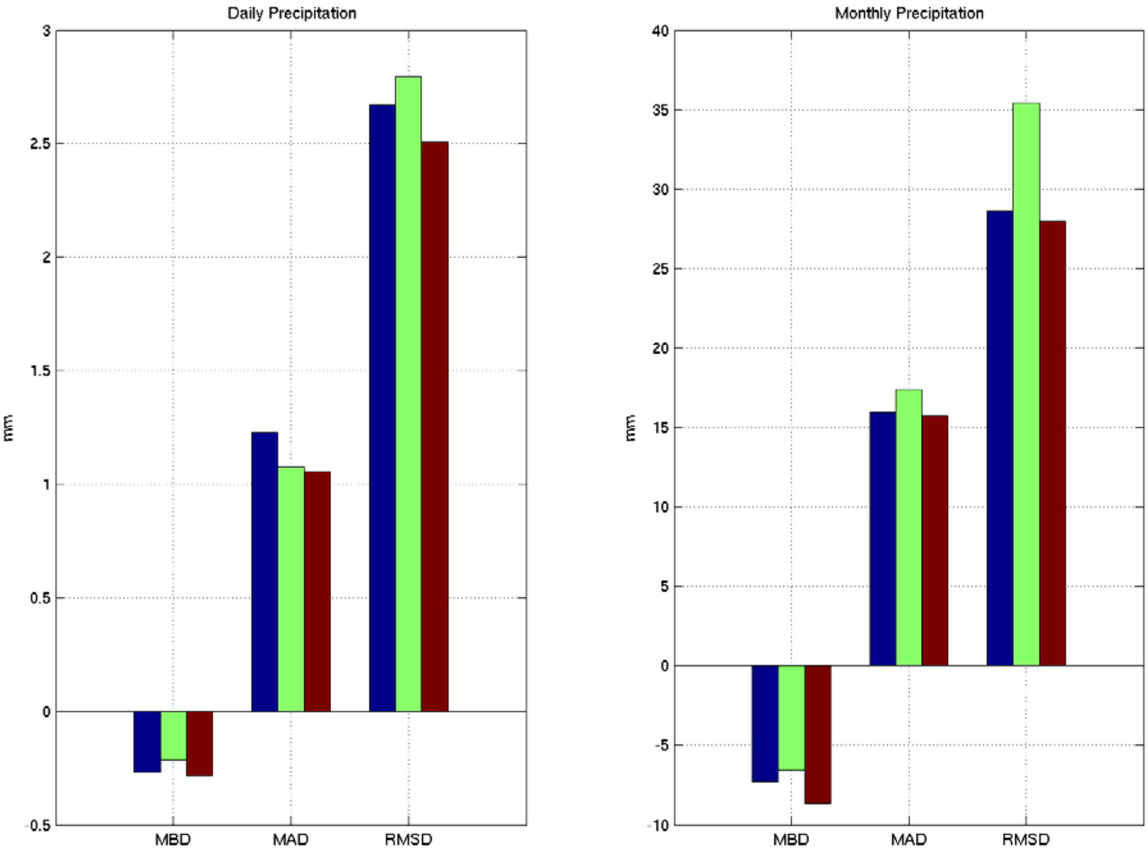


Figure 18. *Histograms for MBD, MAD and RMSD that have been calculated for the daily precipitation (left) and for the monthly precipitation (right) during 1999. The calculations have been done between the independent observations and the gridded data from Finland (blue bars), Norway (green bars) and Sweden (red bars).*

For the 5 highest precipitation observations (figure 20) in each month the largest errors are (for MBD, MAD and RMSD) found in the Swedish data and the smallest errors in Norway (MBD) and Finland (MAD and RMSD).

3.4.3 Discussion of the validation results

In this validation it is important to remember that the EWGLAM-stations which the gridded data have been validated against only constitutes about 9 % of the total amount of temperature observations and about 2 % of the total amount of the precipitation observations used in the creation of the gridded data. But it still constitutes an independent dataset that is well used in the validation of the NWP-models (Schyberg et al., 2003). The ultimate validation is of course a cross-validation, i.e. where one observation at the time (from the total amount of observations) is neglected in the gridding procedure. Such a procedure is very time consuming and therefore not used in this pilot project. Unfortunately the procedure of neglecting the EWGLAM-stations in each country has not been carried out in exactly the same way, and therefore the validation data from Finland, Norway and Sweden are not directly comparable. The difference in the number of neglected observations in Sweden and Norway is only one station and should not have any influence on the comparison. The Finnish data on the other hand comprises 14 fewer stations compared to the Swedish data in the temperature validation and 8 fewer stations in the precipitation validation. This leads to both fewer validated stations and also to more used observations in the gridded data from Finland.
Another issue is that the Swedish gridding method does not only use observations as the gridding methods in Finland and Norway does, it also includes a first guess field. In this case

the ERA-40 re-analyses are used as first guess fields (chapter 2.4). In the Swedish temperature analysis an ERA-40 analysis is used as first guess, i.e. observations are included also in the first guess, but in the precipitation analyses an ERA-40 forecast is instead used (no available precipitation analyses in ERA-40). So the Swedish temperature analyses are not totally independent, the neglected EWGLAM-stations might be used in the first guess field instead. A comparison between the Swedish gridded data and ERA-40 is described by Jansson et al. (2007).

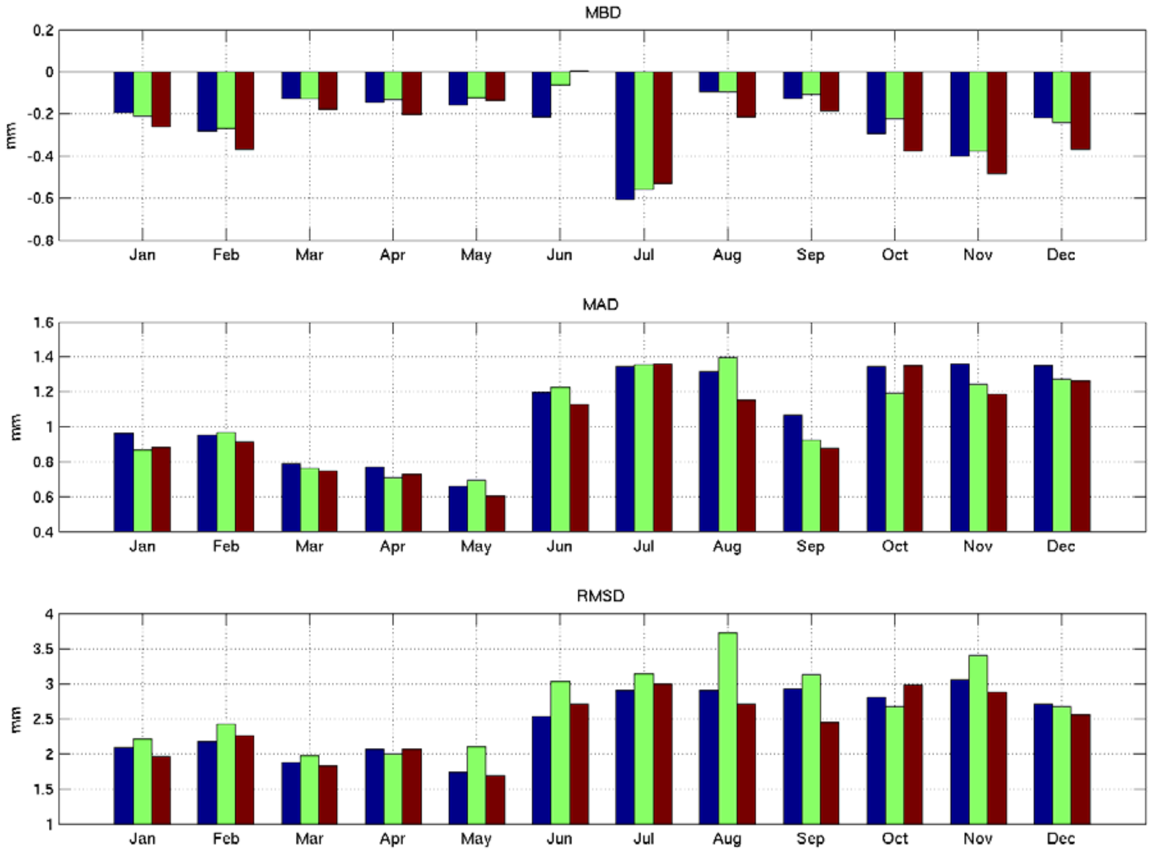


Figure 19. *Histogram for MBD (top), MAD (middle) and RMSD (bottom) calculated for each month separately during 1999 for daily precipitation. The calculations are done between the EWGLAM-stations and the gridded data from Finland (blue bars), Norway (green bars) and Sweden (red bars).*

In the comparison regarding extreme values (figure 17 and 20) different extreme values are probably used in the calculations since the validation data from Finland differ from the Swedish and Norwegian data.

The MAD and RMSD are the most important calculated quantities to study, since the MBD removes the absolute error since positive and negative error can cancel each other. The MBD is instead an important quantity in order to discover possible systematic errors in the model.

A) Temperature

In the temperature validation there is no unexpected validation results. The systematic errors calculated for all daily mean temperatures during 1999 is very low in all models, 0.16 ºC colder in Norwegian model, 0.17 ºC colder in the Swedish model and 0.43 ºC colder in the Finnish model (figure 15, left). The systematic errors in the monthly mean temperatures (figure 15, right) are almost the same as for the daily mean temperatures. The cold systematic error in all models can of course also be influenced by the location of the stations used in the validation, for example if some station is very cold in comparison to the neighbouring stations.

In all validations, for the whole year (figure 15), monthly (figure 16) and for the extremes values (figure 17), the smallest magnitude of MBD (except during Mars, October and November, then the Norwegian model has the smallest magnitude), MAD and RMSD is found in the Swedish model. The largest errors, MAD are in all cases found in the Finnish model and the largest RMSD is mostly found in the Finnish model, except during February and May-September when it is instead found in the Norwegian model.

In the validation of extreme temperatures, all models give colder temperatures in the maximum temperature and warmer values in the minimum temperature validation in comparison to the observations. But the systematic warmer errors in the minimum temperature are much larger than the cold systematic errors in the maximum temperature validation. There is a joint problem to describe really cold temperature, especially during the winter months when deep temperature inversions are likely to occur. Extreme warm temperatures within limited areas are not that common as really cold temperatures, therefore the systematic error is not of the same magnitude as for the minimum temperature. Also, when extreme high temperatures occur, the temperature lapse rate is better described in the gridding methods than for inversion situations, which appear at a more local scale.

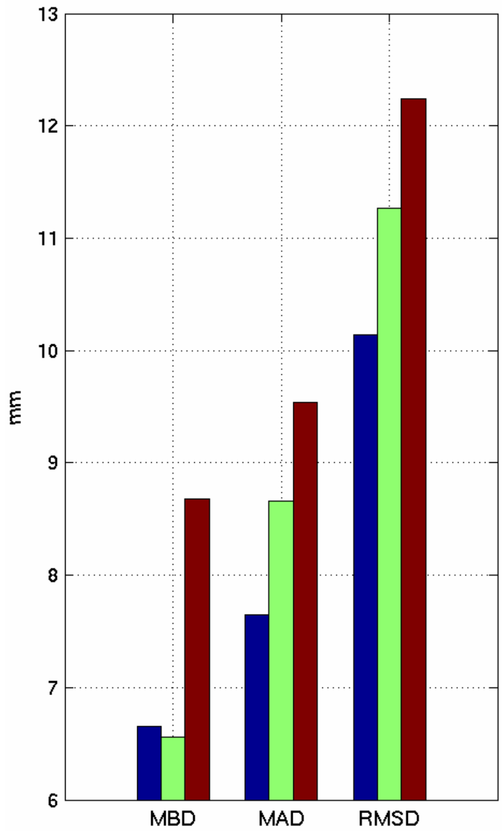


Figure 20. *Histograms for MBD, MAD and RMSD that have been calculated between the 5 highest precipitation observations during each month and the corresponding gridded data from Finland (blue bars), Norway (green bars) and Sweden (red bars).*

B) Precipitation

For gridded precipitation a wet systematic error (negative MBD) occur (figure 18). This is a common feature for small precipitation amounts, for large precipitation amounts the opposite often occur instead.

When the whole year is validated the smallest magnitude of systematic errors are found in the Norwegian model, and the largest in the Swedish model. The smallest MAD and RMSD are instead found in the Swedish model and the largest in the Finnish model (MAD) and in the Norwegian model (RMSD).

In the monthly validation the comparison between the models are not straight forward as for the monthly validation of daily mean temperature. The MAD is lowest in the Swedish model during 8 months, and the RMSD is lowest in the Swedish model during 7 months. Otherwise the Finnish models have the lowest MAD during July and the lowest RMSD during February, June and July. The Norwegian model has the lowest MAD during January, April and October and the lowest RMSD during April and October.

In the validation of extreme precipitation amounts (figure 20), the smallest errors (MBD, MAD and RMSD) is found in the Norwegian model, except for the MBD where the smallest magnitude instead originates from the Finnish model. The largest errors are found in the Swedish model. The systematic errors in all models are of course drier than the observations in the extreme precipitation amount validation (figure 20), due to the difficulties in describing large local precipitation amounts during short periods.

4 Module B: Creation of a gridded dataset based on ”National gridded data” from each participating country

4.1 DATA FROM EACH PARTICIPATING COUNTRY

For this module, each country has delivered already existing gridded data for daily mean temperature and daily precipitation for 1995-1999 over their own land area. The temperature algorithms are in this part the original in each country described in the chapters 3.1.1.1-3.1.1.4

Each country has submitted daily grid point files in a specific predefined data format. These points are representing points in the map projection used in each country (chapter 2.1-2.4) and are given in decimal degrees. Figure 21 shows the grid points for Fennoscandia based on the national grid points.

4.2 CREATION OF A COMMON DATASET, NORDGRID V0

To solve the problem with different grid projections in the participating countries some adjustments have been done on the original temperature and precipitation data.

The following gridding procedure is simply a pure GIS task, performed in the GIS-software ArcInfo (ESRI, 2001).

1. Establish ArcInfo point coverage in geographic projection (decimal degrees). The points included in and the shape of this coverage is the same as shown in figure 21.
2. Exclude grid points outside the territory of the country from where the estimated grid point originate (only applied on the Norwegian data).
3. Converting daily grid values from grid points to a regular grid.

The latter point is not as trivial as expected. If all grid points are in the same map projection and spatial resolution, this is a straightforward procedure to do (convert point values to grid values) that is included in most GIS-software’s. Due to the different projections applied in the participating countries the joint grid over Fennoscandia contains in gaps and therefore another procedure has to be used. Since the dataset is evenly distributed, though somewhat irregular at the borders between the countries, a nearest neighbour approach is a natural choice.

The nearest neighbour technique applied here is the Thiessen method (Shaw, 1983), well known from hydrology. The principle is also known as the Voronoï method. The method generate polygons around each point in a point dataset in the way that any location within the polygon is closer to the polygon point than any other point in the dataset. The result is a polygon dataset, which can be transformed to a continuous gridded dataset. The resulting grid values are area-weighted sums of the portion of the Thiessen polygons contributing to each grid cell. If the grid cell and the Thiessen-polygon are totally adjacent, the polygon and grid value will be equal.

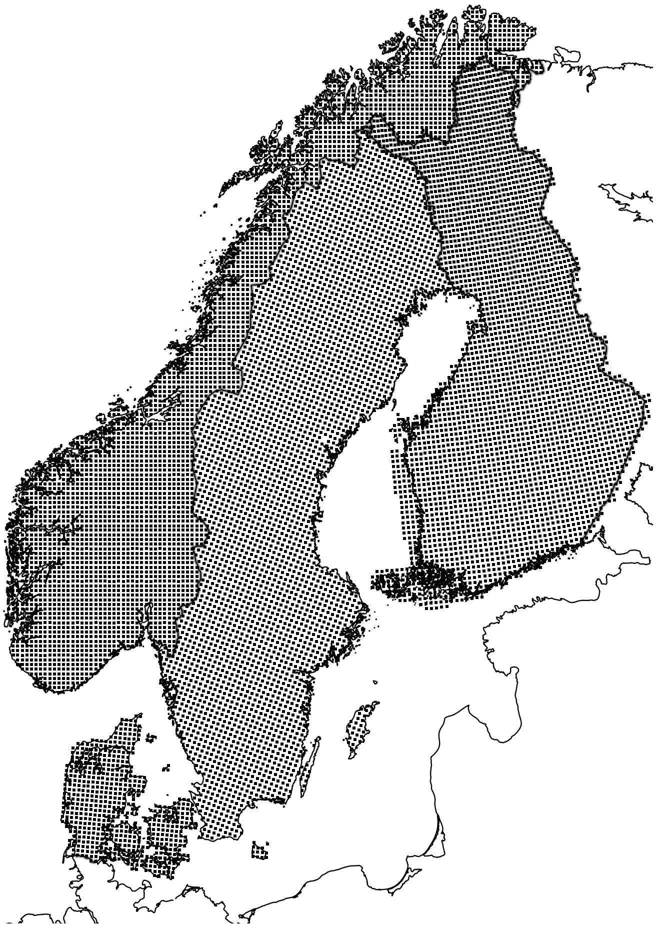


Figure 21. *The grid points for Fennoscandia based on the national grid points, used as a base for NORDGRID v0.*

The point to grid procedure is carried out in 4 steps, and these are visualized in figure 22. In order to get a better impression of the details, focus is given to the region in Northern Fennoscandia where the three national borders between Finland, Norway and Sweden meet (figure 22A). Then the effect of three different map projections can be seen, as well as the problems with the rugged Norwegian coastline.

Step 1: Project the point coverage into desired map projection (figure 22B). Here is WGS84 UTM zone 33 applied since this probably is the most applied projection for environmental applications for Fennoscandia. Other projections can also easily be applied by changing the projection parameters.

Step 2: Establish Thiessen polygons around each data point (figure 22C). A buffer (10 kilometres) around the Fennoscandia land area is used to clip the polygons to prevent gridded values in the seas between e.g. Norway and Denmark, Sweden and Finland etc.

Step 3: Transforming the polygons into a regular 10x10 km² grid using the `polygrid` command in ArcINFO (figure 22D).

Step 4: Identify grid cells with possible missing values in the grids, and assign these cells with the missing value code.

Examples of the resulting grid for daily mean temperature is shown in figure 23 and for daily precipitation in figure 24.

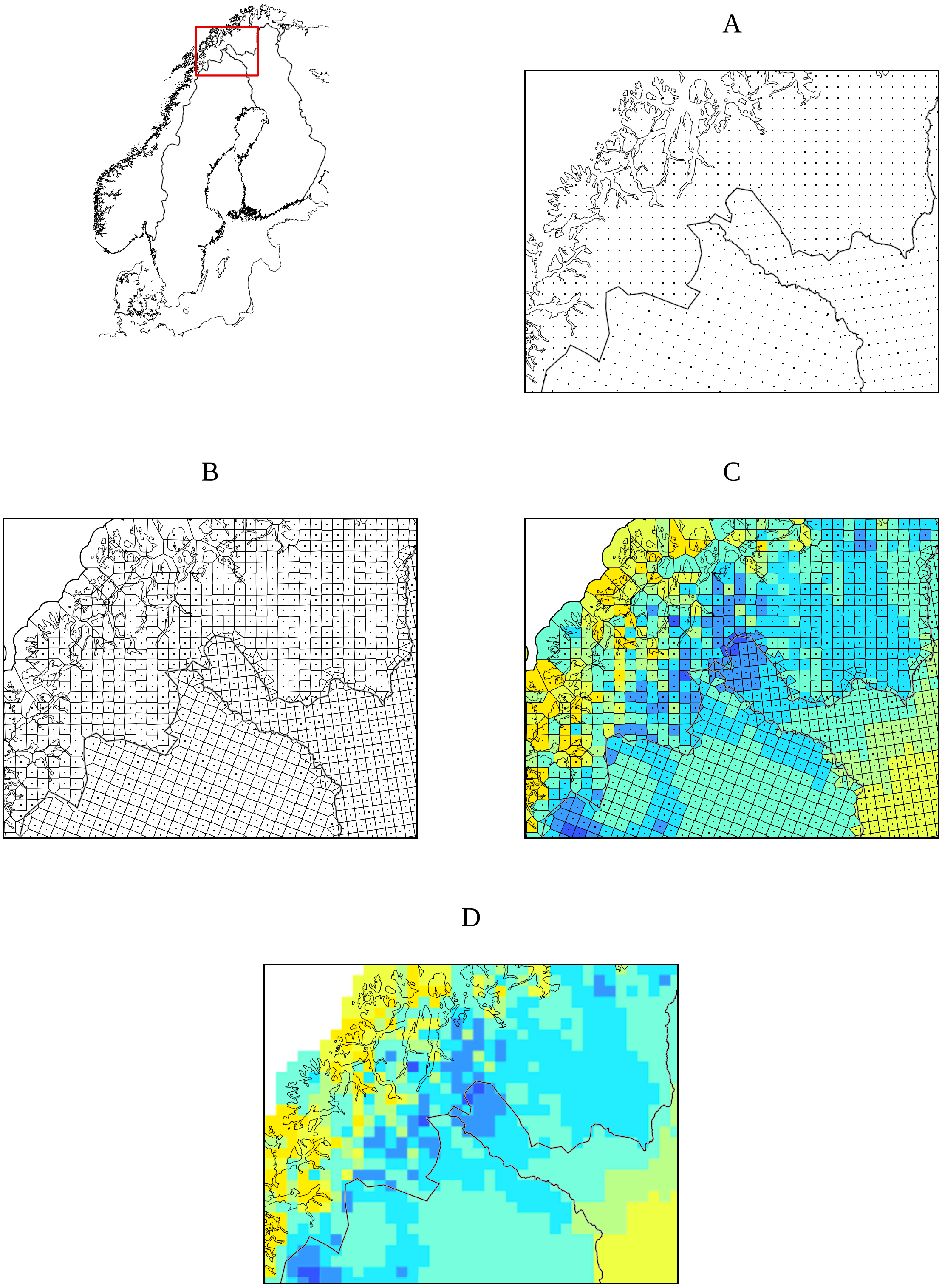


Figure 22. *The red box in the top left map shows the extent of panels A-D. Panel A shows the centre points of the national gridded datasets. Panel B the Thiessen polygons around these points. Panel C shows the temperature values of the Thiessen polygons and panel D the resulting regular grid in the chosen map projection. (Panels C and D show the daily mean temperature at the 16. April 1995.)*

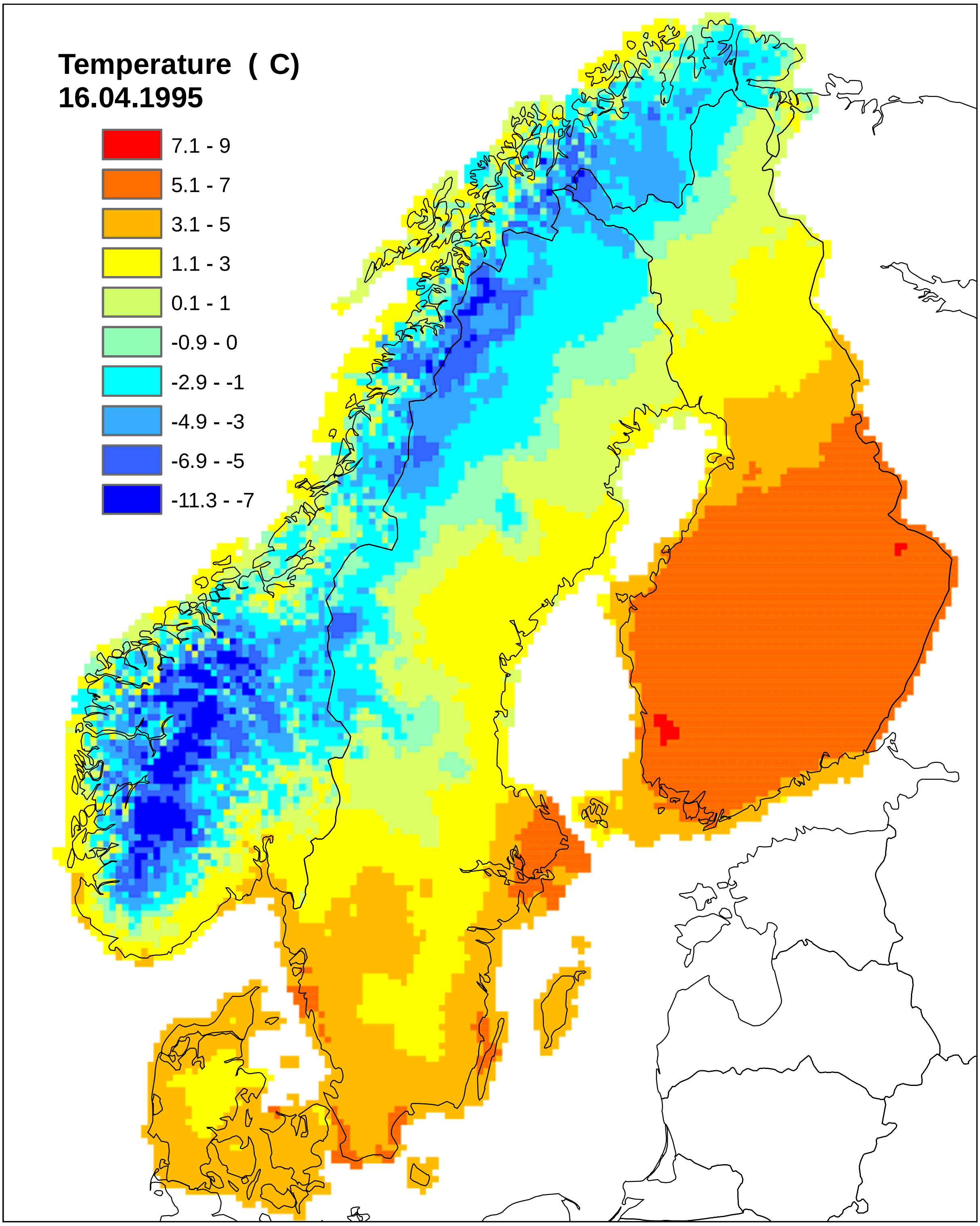


Figure 23. *Daily mean temperature at the 16. April 1995 from NORDGRID v0.*

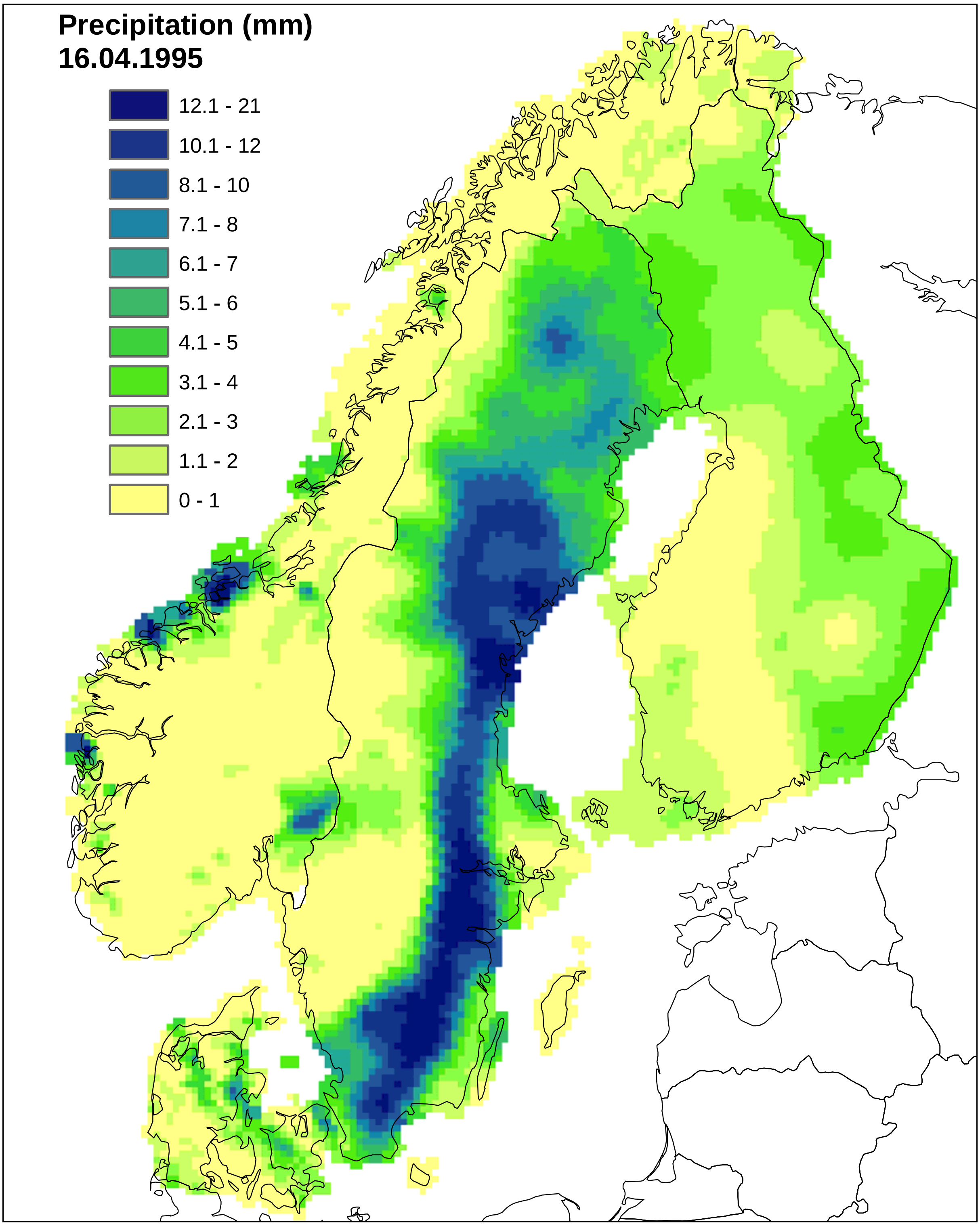


Figure 24. *Daily precipitation sum at the 16. April 1995 from NORDGRID v0.*

5 Module C: Comparison of grid point estimates along national borders

When applying different gridding methods in adjacent areas, the results along the common border should be expected to be inconsistent. This is not only a methodological problem, but also a problem in further applications of the data. In the case of the potential NORDGRID v0, the differences along the borders have been studied. This is an important issue for the possible users of this joint gridded dataset (NORDGRID v0).

Two tests has been done to possibly identify discrepancies along the borders, the first is to compare the NORDGRID v0 dataset with a consistent dataset, the ERA-40 data from ECMWF and the second is to compare the variation in the grid points along the borders.

5.1 Comparison between ERA-40 and NORDGRID v0

In this chapter, the ERA-40 dataset (Uppala et al., 2005) has been used as a reference to discover possible discrepancies along the national borders. The merged NORDGRID v0 dataset has first been converted into the same coordinate system and also divided into the grid square of the ERA-40 data (the ERA-40 data has been retrieved on 44x44 km² resolution and later on interpolated to 11x11 km² with the same coordinate system as in the Swedish model). The MBD between ERA-40 and the merged NORDGRID v0 in each grid point has then been calculated for the whole 1995. The value of the MBD is not of interest in this chapter, instead the structure and the characteristics around the national borders are of interest. In figure 25 the MBD for the daily mean temperature (left) and for the daily precipitation (right) is illustrated. Because of different coordinate systems in the ERA-40 data and NORDGRID v0 some empty grid points exist, in this case the closest grid point value has been used instead. There are only a limited numbers of empty grid values, so it does not influence on the result. In the MBD calculation for the daily mean temperature (figure 25, left) clear discrepancies occurs at the borders between Sweden and Finland and between Sweden and Norway especially in the northern parts. One reason is of course different gridding methods but for the temperature also the different temperature algorithms and the time when the daily mean temperature is valid influence the discrepancies along the borders. The MBD calculated for the daily precipitation (figure 25, right) has big discrepancies along the border between Sweden and Norway, apart from the most southern parts. For the daily precipitation sums the time range is the same for all countries, therefore the discrepancies only arises because of different gridding methods.

A common problem which probably is the main reason for the discrepancies along the borders is the different number of observations used in each national gridded dataset. Since the Norwegian and Finnish models only make gridded maps over its only land area, they only use observations from its own country. This makes the gridded maps along the borders contains in very few observations. These observations may not be representative for the area close to the borders, and might produce “false” trends which could have been avoided if observations from the neighbouring countries had been used in the calculations. The opposite arise in the Swedish model, here gridded maps are created for both, Sweden, Norway and Finland and also observations from all countries are used in the creation of the gridded data.

5.2 Along the national borders

Here the variations among grid cells in the area along the borders are studied, also the difference for cells exactly along the border with the differences for the cells in the vicinity of the border are compared. The differences are calculated as the range in values within a 3x3 grid cell rectangle. Figure 26 shows the two zones that are analysed. The black area is the zone describing the area near the border, while the red area mark the cells along the border it self. The two areas are separated in the further analysis. This is not a completely perfect approach, but sufficient to reveal systematic discontinuities across the national borders. The statistics is run for a complete year, 1995. Figure 27 shows the resulting mean daily difference range for daily precipitation. It shows no clear indications of discontinuities across the border. There are some regions with large differences, but these values only reflect the high variability in the precipitation itself in these areas. The indication of no systematic inconsistencies along the borders is confirmed by studying the frequency curves of the mean daily ranges (figure 29). Here the statistics is divided into the three separate borders Sweden-Finland, Finland-Norway and Norway-Sweden. This is done in order to reveal eventual systematic differences between the national methods individually. For precipitation the frequency curves does not show particularly large differences, indicating that inconsistency across the borders is not a big problem. The methods used for gridding precipitation are not using terrain characteristics or other external predictors that are so different that they will lead to large systematic errors along the borders. Also the precipitation network is denser than for temperature, but on the other hand the spatial variability is much more random. The latter effect can hide systematic differences due to different representation of e.g. topographical parameters.

For the daily mean temperature the situation is different. From figure 28 it can be seen that there is a larger range along the border than in the grid cells beside. This indicates a possible problem with inconsistencies across the national border, which is confirmed by investigating the frequency curves of the temperature range (figure 29).

At all the three borders the temperature range has an offset considerably larger than for the surrounding areas on each side of the border. This indicates a discontinuity across the borders.

The largest offset is found along the Norwegian-Finnish border, which starts already at 3.2 °C. The difference is less along the two other borders, but still considerably large. The reason why such inconsistencies occur for temperature is that the national methods take different physical and physiographical characteristics into consideration. Also the differences in the original daily mean temperature algorithms influence on the discrepancies. The fact that the methods applied in Finland and Norway do not include observations from the neighbouring countries lead to more inaccurate estimates in the border regions since only stations in one direction is included for the estimation. In this case the estimates are more extrapolations rather than interpolations.

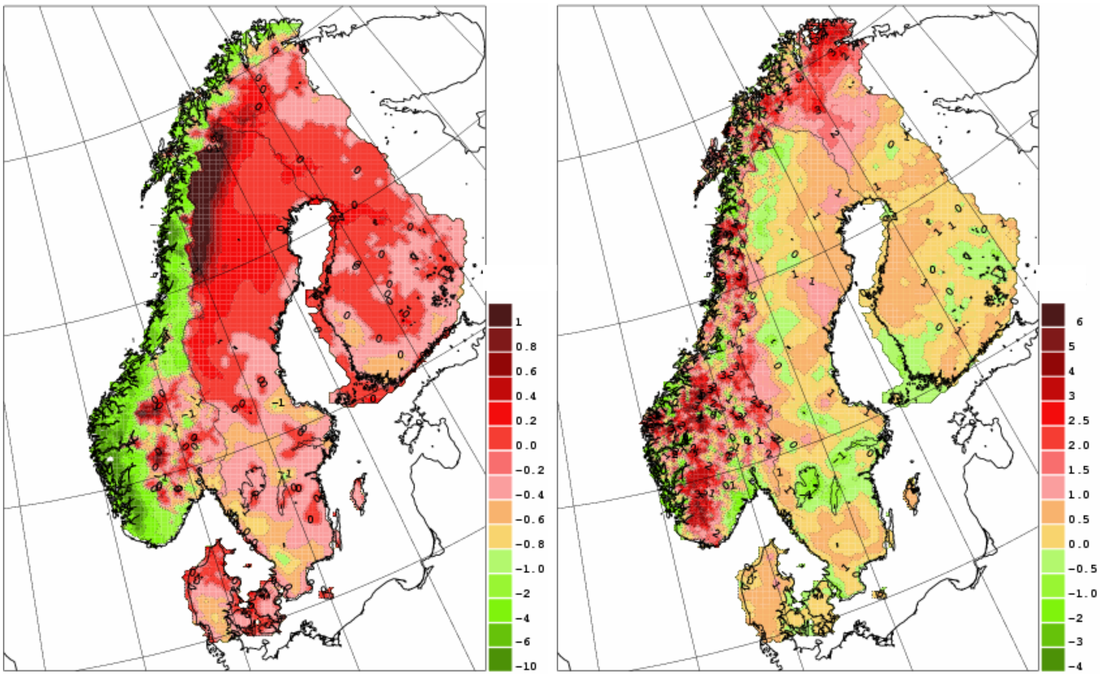


Figure 25. *To illustrate possible discrepancies along the national borders the MBD for 1995 has been calculated between ERA-40 and NORDGRID v0 for daily mean temperature (ºC) (left) and for daily precipitation sums (mm) (right).*

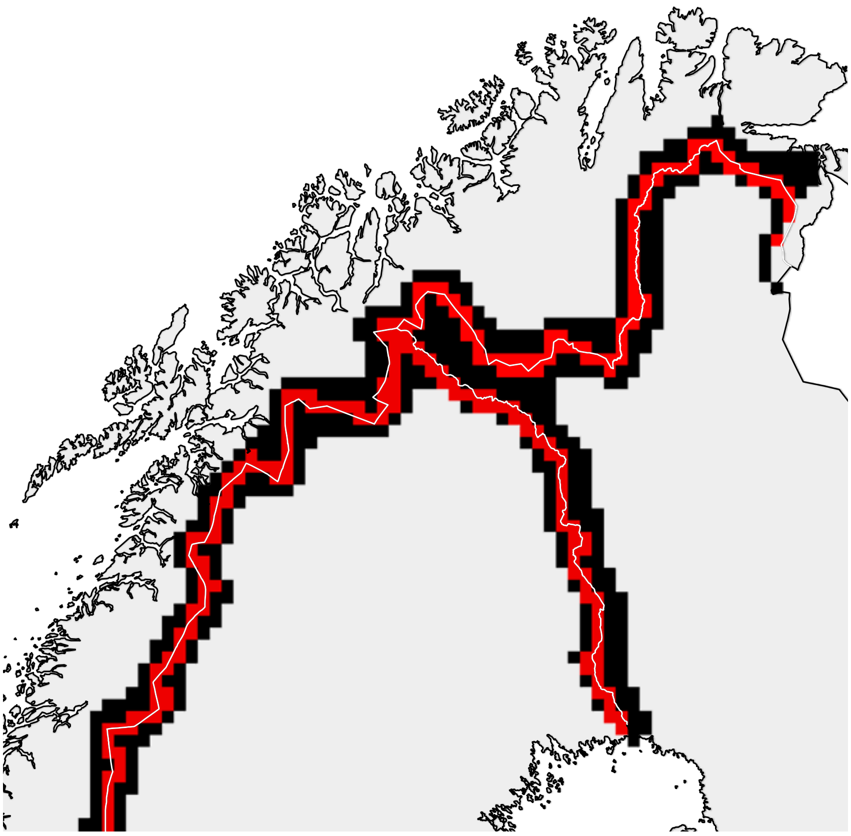


Figure 26. *The map illustrates the two different areas (red and black coloured) which are considered when discrepancies along the national borders are studied.*

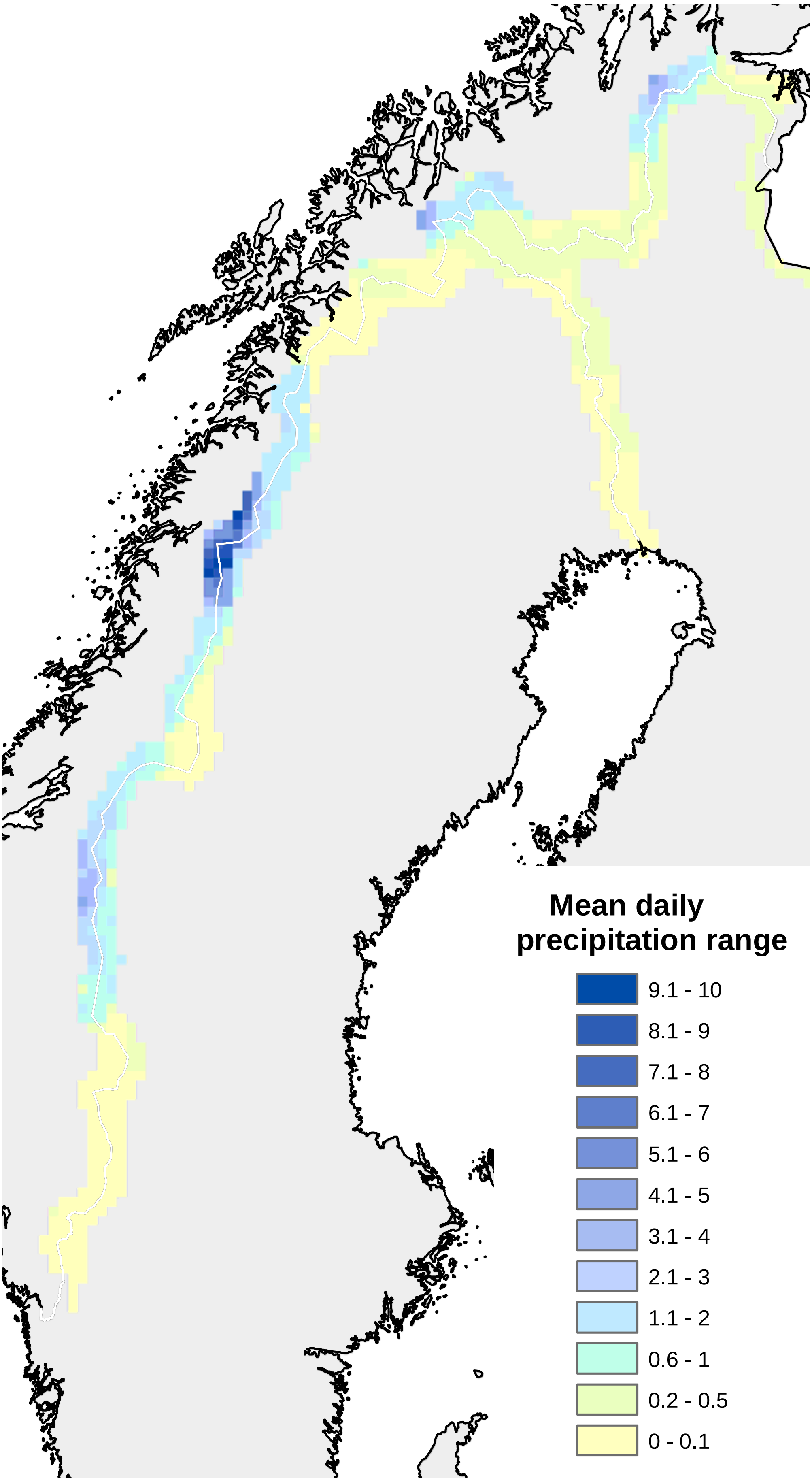


Figure 27. *Mean daily precipitation range (mm) within a moving 3x3 grid cell window for the year 1995.*

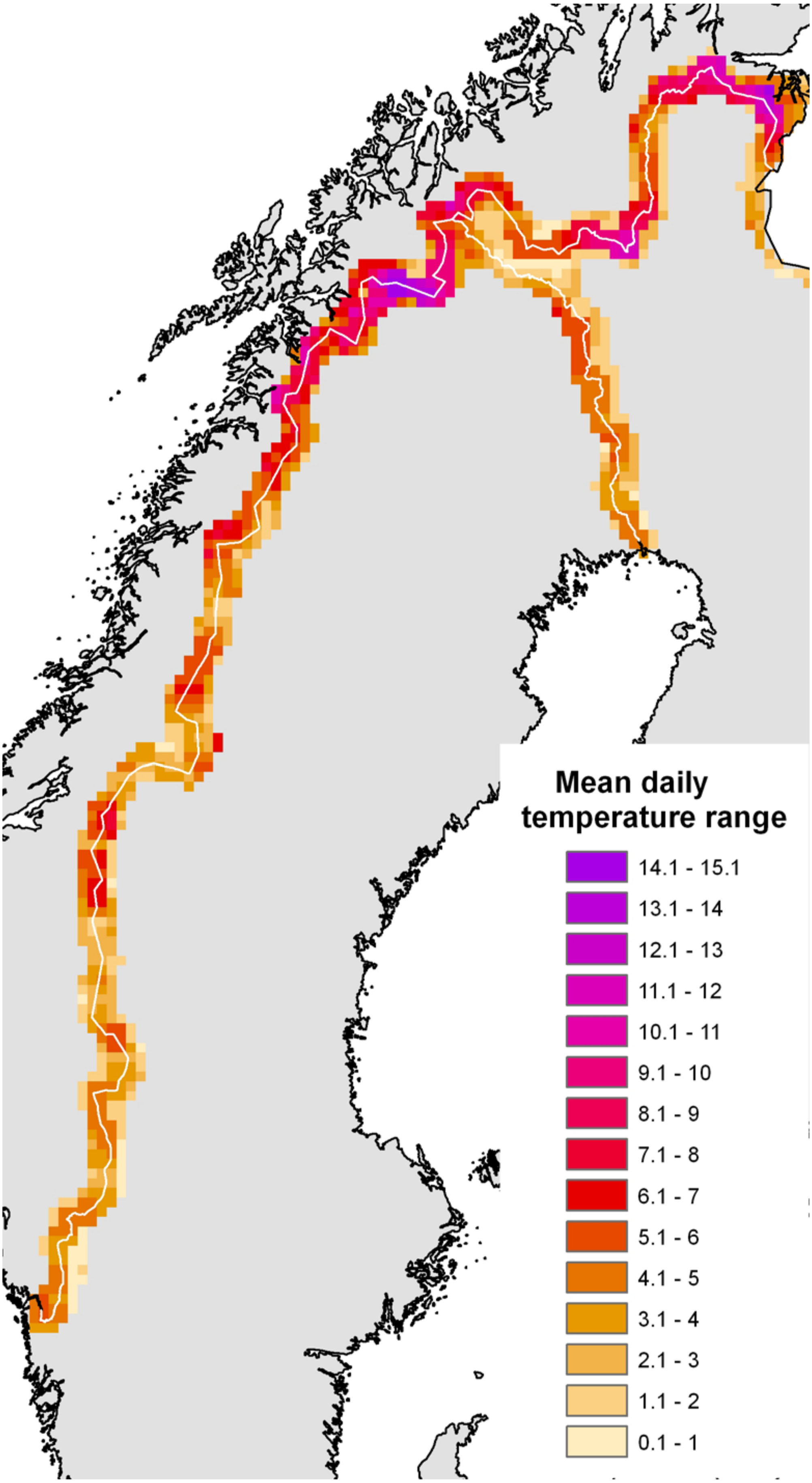


Figure 28. *Mean daily temperature range (°C) within a moving 3x3 grid cell window for the year 1995.*

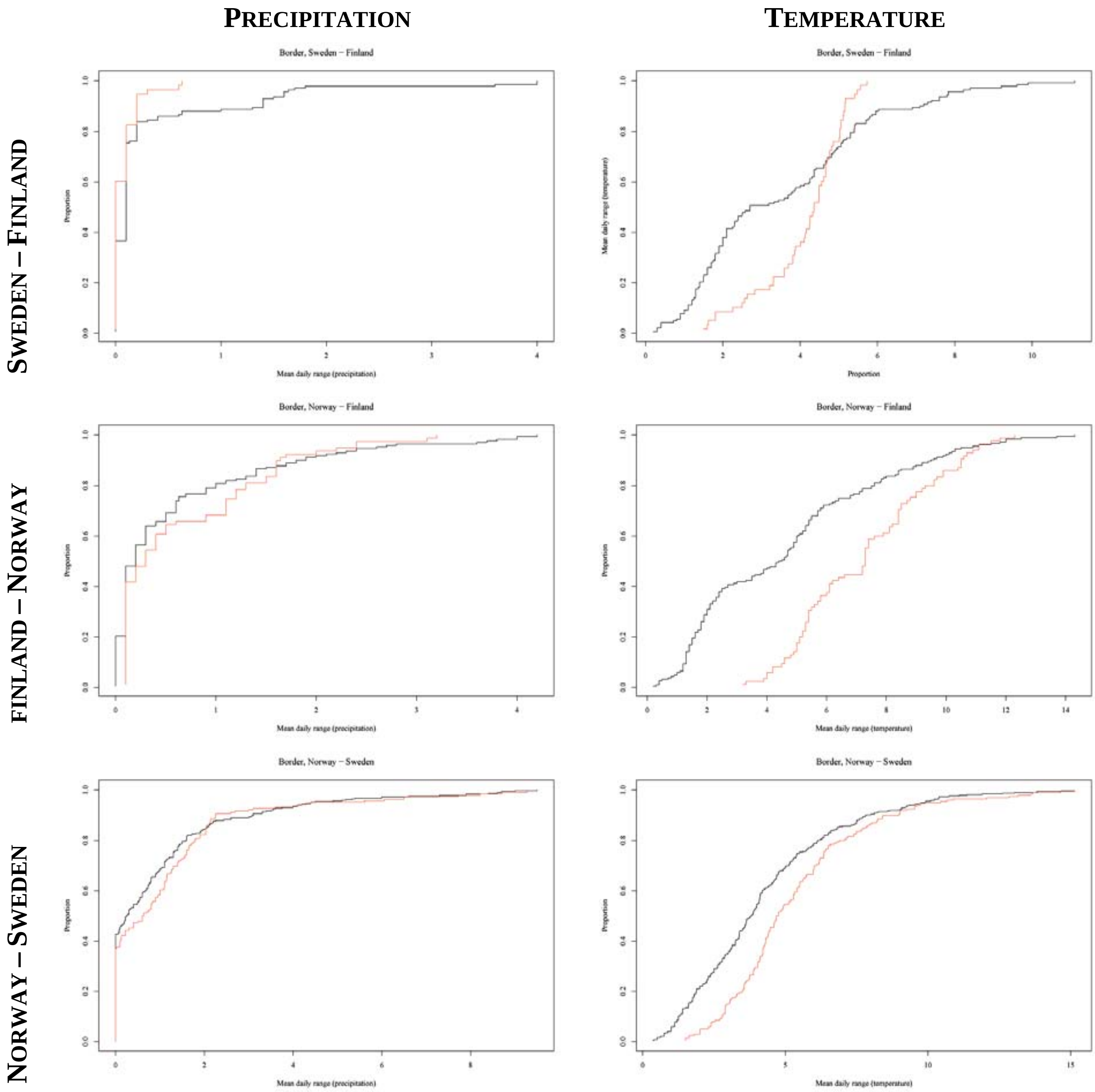


***Figure 29**. Frequency curves of daily precipitation sums (mm) and temperature (°C) ranges for a moving 3x3 grid cell window during year 1995. The red curve is for the grid cells along the borders and the black curve for the cells beside the borders (see figure 26).*

6 Discussion and Conclusions

A validation of the gridding methods used for climatological applications used at the Finnish, Norwegian and Swedish meteorological institute has been done (module A). The validation has considered daily precipitation and daily mean temperature during 1999. The validation includes Fennoscandia. The observations used in all gridding procedures originate from each institute separately and are also quality controlled at each institute. A common algorithm for daily mean temperature is used to overcome the problem of different algorithms in the respective country. For the validation, a set of reference stations has been used, the EWGLAM-stations, well used in validation of NWP-models. It is important to remember that it is not a complete validation since the EWGLAM-stations only constitutes about 9 % of the total amount of temperature stations used and about 2 % of the total amount of precipitation observations used in the analyses. Unfortunately, different procedures in neglecting the EWGLAM-stations from the total number of stations have been done differently in the participating countries. Therefore the numbers of neglected EWGLAM-stations are not the same in all countries. The differences between the Swedish and Norwegian analyses are negligible, maximum one station differs. But in the analysis from Finland, 14 stations of the EWGLAM-stations have not been neglected in the temperature analysis. In the precipitation analysis, instead 8 stations of the EWGLAM-stations have not been neglected. This leads to both fewer validated stations and also more information (more station used in the analysis) for the validated stations in the Finnish analyses. Also, the temperature analysis from Sweden might not be totally independent since the first guess field is an ERA-40 re-analysis which also includes observations.

Another misunderstanding in this project, that was discovered too late to be able to correct it, is that all available weather stations that measure temperature and daily precipitation during 1999 have been delivered from Denmark, Norway and Sweden but from Finland only 80 weather station of totally 400 weather stations were delivered and used as input data to all the national gridding applications. So the validation results for all models should probably have been better if 400 weather stations were used in Finland instead of only 80 weather stations.

For the temperature validation the Swedish model has the smallest errors, both for MBD, MAD and RMSD, in respect to the yearly (figure 15), monthly (figure 16) and extreme values (figure 17) validation. In the precipitation validation the errors are more equally in the participant models. In the systematic error (MBD) aspect the smallest errors originates mostly from the Norwegian model when the whole year is validated and the other errors (MAD and RMSD) are smallest in the Swedish model, but the differences are very small. In the validation for extreme precipitation amounts the remarkably best results are found in the Norwegian model. In the monthly validation it is difficult to say which model has the smallest errors, the Norwegian model has most months with the smallest magnitude of the MBD while the Swedish model has most months of lowest MAD and RMSD. Further discussions on how to carry on considering these validation results will be given in chapter 7; Future plans.

In module B a creation of a common gridded dataset for 1995-1999 has been performed. The dataset is constructed of already existing maps from the participant countries for daily mean temperature and daily precipitation. This means that the grids will not be consistent over the whole Fennoscandia as the national gridded datasets are based on different assumptions and methodological principles. The dataset also consist of four different map projections, an example how to deal with this problem with a minimum loss of information is also presented. To address the inconsistency, and the possible discontinuity in the area along the national borders, an analysis of the variability along the borders and the areas near the borders is carried out (module C). The results of this analysis indicate inconsistencies along the borders for the temperature grids, while this was not a large problem concerning precipitation.

7 Future Plans

There is an objective for the Nordic NMS's to develop a methodology to establish a joint gridded dataset. The model should cover all Nordic countries and utilize high quality controlled observations from all the Nordic NMS's. The methodology, the analyses and the derived grids shall all be available for the participating countries. The model will be used to create long term gridded datasets of climatological parameters, such as temperature and precipitation, on a daily basis for all Nordic countries. The work presented in this report is the first step in this direction. The aim have been to establish a basis for deciding which model concept to use for the joint dataset, if it should be one of the existing models, a combination of several models, or something new.

It is important to consider the use of a long time series of climatological parameters:

1) If the dataset will be used as input to other models, such as radiation and chemistry atmospheric models, then it might be of importance that the database covers the sea areas where no observations are available. In that case the Swedish model, probably with further developments to fit the Nordic NMS´s needs, has to be used since it is the only model that makes analyses over sea areas.
2) If the dataset instead should be used to visualize past and present climate perhaps a more substantial validation should be done, for example a cross validation, and/or to develop a new model that takes each models strengthens into consideration.

Another issue is also which parameters that should be included in a gridded climatological database for the Nordic countries. If the database contains too few meteorological parameters it might be possible that the future users might choose another data source that includes all needed meteorological parameters for their purposes. However, the use of an extensive number of high quality controlled data series will make the future joint NORDGRID dataset unique and outstanding compared to other available datasets. The spatial resolution of the dataset will also be an advantage compared to other accessible datasets.

All issues and considerations described earlier are all a matter of the resources from each Nordic NMS's, it probably has to weight together in some way. These issues and considerations will hopefully be discussed and decided after the next NORDKLIM-workshop in late autumn 2006. During this meeting the focus and usage of the future NORDGRID work will be discussed. After the meeting a detailed work plan for the next 3 years where European wide re-analysis projects (EURRA and EUROGRID) are taken into consideration will be created, and sent to NOSC for approval.

NORDGRID in Europe

Below, a short description of the background and purpose of long time gridded datasets in our community will be given. Also a description about the European wide re-analysis project that may influence the Nordic cooperation will be given.

Gridded datasets of meteorological parameters that are time consistent, has high spatial and temporal resolution and covers many years back in time is a hot topic in Europe today. Re-analyses have stepwise become a basis for gridded historical datasets. The most used re-analysis today is the ERA-40 dataset created at the ECMWF. The ERA-40 dataset covers the years 1957-2002, the analyses are time consistent but the spatial resolution are quite coarse (~125 km). There are many discussions and plans for creating re-analyses with a much finer spatial resolution in different countries. Especially 2 planned projects will be of great importance for the future of NORDGRID.

1) The EURRA project
2) The Showcase EUROGRID

The plan for the EURRA (European regional re-analysis) project is to create a high resolution re-analyses over Europe. This project was initialized by the European Environment Agency (EEA) at a meeting at ECMWF in late 2005. The EURRA project is supposed to fulfil a special requirement on meteorological data for EEA. Today no decision is made on when and on which institute the project will be performed at.

The Showcase EUROGRID was first presented by Bengt Dahlström, SMHI, and conceded for the European Climate Support Network (ECSN) Advisory Committee in Copenhagen, June 2005. At that stage the project plans were much larger than today and only called EUROGRID. The aim of the EUROGRID project was to create a database of daily values for temperature and precipitation and other meteorological parameters over Europe with a high spatial resolution and for a relatively long time period that subsequently should be updated. The project proposal also included a product generating system mainly web based, that should make it easy for the users to collect data from the database and generate products.

When the EUROGRID project plan was presented at the EUMETNET Council in Helsinki in October 2005, the outcome of the meeting was that the project was to expensive, therefore the EUROGRID project was cut down to the “Showcase EUROGRID” and later on passed at the latest EUMETNET Council meeting. The Showcase EUROGRID only covers one year and is instead focused on illustrating how a meteorological database effectively can be a base for different kinds of meteorological products and applications. The Showcase EUROGRID was approved by EUMETNET Council (April 2006) and is just in the starting blocks.

The way to look at these two projects is that the EURRA project is the first step in a re-analyse cycle, to create high resolution re-analyses over Europe, and that the Showcase EUROGRID is a parallel step, to create products and applications from existing datasets (NORDGRID and a couple of other datasets). The two activities have to be worked out with a mutual adjustment so that there will be a later marriage – the two enters the intended EUROGRID (or how the new unified couple will named).

8 Acknowledgements

The authors would like to express their thanks to Hans Alexandersson (SMHI), Tomas Landelius (SMHI), Bengt Dahlström (SMHI), Erik Liljas (SMHI), Eirik Førland (met.no), Ari Venäläinen (FMI), Claus Kern-Hansen (DMI) and Torill Engen-Skaugen (met.no) for fruitful discussions and help with data retrieval.

9 References

Bjørge, D., Haugen, J E., Homleid, M., Vignes, O. and Ødegaard, V. (2003) Updating the HIRLAM numerical weather prediction system at *met.no* 2000-2002, NOTE Report nr 145, met.no, Postboks 43 Blindern, 0313 OSLO, Norway, 25 pp.

ESRI (2001) ArcInfo – geographic information system v.8.1, Environmental Systems Research Institute, Redlands, USA.

Førland, E.J. and Tveito, O.E. (1997) Temperatur og snødata for flomberegning, DNMI Report 28/97 KLIMA, met.no, Postboks 43 Blindern, 0313 OSLO, Norway, 50 pp.

Häggmark, L., Ivarsson, K-I., Gollvik, S. and Olofsson, P.-O. (2000) Mesan, an operational mesoscale analysis system, Tellus, 52, pp 2-20.

Henttonen, H. (1991) Kriging in interpolating July mean temperatures and precipitation sums, University of Jyväskylä, Department of Statistical Science, Publication No. 12, Jyväskylä, 41 pp.

Jansson, A., Persson, C., Strandberg, G. (2007) 2-D mesoscale reanalysis of precipitation, temperature and wind over Europe – ERAMESAN, SMHI, S-60176 Norrköping, Sweden, In preparation.

Kuttinen, R., Ahokas, E., Granqvist, M., Heikinheimo, M., Ikäheimo E., Jänne, S., Keskisarja, V., Parmes, E. and Venäläinen, A. (1999) An Early Crop Yield Estimation Method for Finnish Conditions: The Crop Growth Monitoring System of the Joint Research Centre with and without remotely Sensed and other Additional Input Data, Space Application Institute, Agricultural and Regional Information System Unit, 120 pp.

Nordli, P.Ø., Alexandersson, H., Frich, P., Førland, E., Heino, R., Jónsson, T., Steffensen, P., Tuomenvira, H., and Tveito, O.E. (1996) The effect of radiation screens on Nordic temperature measurements, DNMI Report 4/96 KLIMA, met.no, Postboks 43 Blindern, 0313 OSLO, Norway, 56 pp.

Ripley, B. (1981) Spatial Statistics, Wiley, New York, 272 pp.

Shaw, E. (1983) Hydrology in Practise, Van Nostrand Reinhold, Wokingham, UK, 569 pp.

Schyberg, H., Landelius, T., Thorsteinsson, S., Tveter, F. T., Vignes, O., Amstrup, B., Gustavsson, N., Järvinen, H. and Lindskog, M. (2003) Assimilation of ATOVS data in the HIRLAM 3D-VAR System, HIRLAM Technical Report 60, HIRLAM Report, SMHI, S-60176 Norrköping, Sweden, 67 pp.

Tveito, O.E. (2007) Spatial distribution of winter temperatures in Norway related to topography and large-scale atmospheric circulation, PUB Kick off meeting, Brasilia, Brazil

20-22. (http://cost719.met.no/papers/PUB%20paper%20Tveito.pdf), IAHS Publ. 309, in press.

Tveito, O.E. and Førland, E. (1999) Mapping temperatures in Norway applying terrain information, geostatistics and GIS, Norsk geogr. Tidsskr. 53, pp 202-212.

Tveito, O.E., Førland, E.J., Heino, R., Hanssen-Bauer, I., Alexandersson, H., Dahlström, B., Drebs, A., Kern-Hansen, C., Jónsson, T., Vaarby-Laursen, E. and Westman, Y (2000) Nordic Temperature Maps, DNMI Klima 9/00 KLIMA, met.no, Postboks 43 Blindern, 0313 OSLO, Norway, 55 pp.

Vajda, A. and Venäläinen, A. (2003) The influence of natural conditions on the spatial variation of climate in Lapland, northern Finland, International Journal of Climatology, 23, pp 1011-1022.

Venäläinen, A. and Heikinheimo, M. (1997) The spatial variation of long-term mean global radiation in Finland, Int., J. Climatol., 17, pp 415-426.

Venäläinen, A. and Heikinheimo, M. (2002) Meteorological data for agricultural applications, Physics and Chemistry of the Earth. vol 27/23-24, pp 1045-1050.

Venäläinen, A., Tuomenvirta, H., Pirinen, P. and Drebs, A. (2005) A Basic Finnish climate data set 1961-2000- description and illustrations, Reports 2005:5, P.O. Box 503, FI-00101 Helsinki, Finland, 27 pp.

Undén, P., Rontu, L., Järvinen H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A., Jones, C., Lenderlink, G., McDonald, A., McGrath, R., Navascues, B., Nielsen, N. W., Ødegaard, V., Rodriguez, E., Rummukainen, M., Rõõm, R., Sattler, K., Hansen Sass, B., Savijärvi, H., Wichers Schreur, B., Sigg, R., The, H. and Tijm, A. (2002) HIRLAM-5 Scientific Documentation, HIRLAM Report, SMHI, S-60176 Norrköping, Sweden, 144 pp.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A, Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., Mcnally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P. and Woollen, J. (2005) The ERA-40 re-analysis, Q.J.R. Meteorol. Soc., 131, pp 2961-3012.

USGS (1996) GTOPO30 Global Digitial Elevation model, US Geological Survey, Sioux Falls, USA (http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html).

Appendix

List of all EWGLAM-stations in Fennoscandia with name, country and WMO-number:

Name	Country	WMO-number
Tromsø/Langnes	Norway	1025
Alta lufthavn	Norway	1049
Vardø	Norway	1098
Sklinna Fyr	Norway	1102
Rana-Basmoen	Norway	1149
Bodø VI	Norway	1152
Svinøy fyr	Norway	1205
Trondheim/Vaernes	Norway	1271
Bergen/Florida	Norway	1317
Oslo/Gardemoen	Norway	1384
Stavanger/Sola	Norway	1415
Kristiansand/Kjevik	Norway	1452
Rygge	Sweden	1494
Karesuando	Sweden	2080
Gunnarn	Sweden	2128
Luleå/Kallax	Sweden	2186
Holmögadd	Sweden	2288
Sveg	Sweden	2324
Söderhamn	Sweden	2376
Karlstad Airport	Sweden	2418
Stockholm/Bromma	Sweden	2464
Jönköpings Airport	Sweden	2550
Gotska Sandön	Sweden	2584
Osby	Sweden	2626
Hoburg	Sweden	2680
Ivalo	Finland	2807
Sodankylä	Finland	2836
Pudasjärvi	Finland	2867
Kajaani	Finland	2897
Kuopio	Finland	2917
Joensuu	Finland	2929
Jyväskylä	Finland	2935
Tampere/Pirkkala	Finland	2944
Pori	Finland	2952
Turku	Finland	2972
Helsinki-Vantaa	Finland	2974
Thorshavn	Denmark	6011
Aalborg	Denmark	6030
Thyborøn	Denmark	6052
Skrydstrup	Denmark	6110
København/Kastrup	Denmark	6180