A HIGH-RESOLUTION, GRIDDED DATASET FOR MONTHLY TEMPERATURE NORMALS (1971–2000) IN SWEDEN

BY

DELIANG CHEN1 LEBING GONG2, CHONG-YU XU2,3 AND SVEN HALLDIN2

1 Earth Sciences Centre, Göteborg University, Göteborg, Sweden
2 Air and Water Science, Department of Earth Sciences, Uppsala University
3 Department of Geosciences, University of Oslo


ABSTRACT. A baseline climatology is required in evaluating climate variability and changes on regional and local scales. Gridded climate normals, i.e. averages over a 30-year period, are of special interest since they can be readily used for validation of climate models. This study is aimed at creating an updated gridded dataset for Swedish monthly temperature normals over the period 1971–2000, based on standard 2-m air temperature records at 510 stations in mainland Sweden. Spatial trends of the normal temperatures were modelled as functions of latitude, longitude and elevation by multiple linear regression. The study shows that the temperature normals are strongly correlated with latitude throughout the year and especially in cold months, while elevation was a more important factor in June and July. Longitude played a minor role and was only significant in April and May. Regression equations linking temperature to latitude, longitude and elevation were set up for each month. Monthly temperature normals were detrended by subtracting spatial trends given by the regressions. Ordinary kriging was then applied to both original data (simple method) and detrended data (composite method) to model the spatial variability and to perform spatial gridding. The multiple regressions showed that between 82% (summer) and 96% (winter) of the variance in monthly temperature normals could be explained by latitude and elevation. Unexplained variances, i.e. the residuals, were modelled with ordinary kriging with exponential semivariograms. The composite grid estimates were calculated by adding the multiple linear trends back to the interpolated residuals at each grid point. Kriged original temperature normals provided a performance benchmark. The cross-validation shows that the interpolation errors of the normals are significantly reduced if the composite method rather than the simple one was used. A gridded monthly dataset with 30-arcsecond spacing was created using the established trends, the kriging model and a digital topographic dataset.

Keywords: monthly temperature normals, 1971–2000, interpolation, kriging, Sweden

Introduction

Temperature is one of the most important climate variables because of its crucial impact on the mass and energy budgets of the earth–atmosphere system. Temperature is also a major factor controlling eco- and agricultural systems in mid- and high-latitude regions where the temperature variability is relatively high. Large seasonal variations in incoming solar radiation result in a strong seasonal cycle and varying influence of different air masses is the cause of relatively high inter-annual temperature variations in these regions. The inter-annual variations are often related to the average state of the temperature climate described by normals. The concept of normals was introduced to enable descriptions of the average state of climate elements.

The World Meteorological Organisation (WMO) defines climatological standard normals as the arithmetic mean of a climatological element computed over three consecutive decades (WMO 1989). Normals are computed every decade by individual countries to keep up with any climatic changes that may take place. The current standard normal is, according to WMO, defined by the 30-year averages of the climate elements for the period 1971–2000. The Swedish Meteorological and Hydrological Institute (SMHI), and many other meteorological institutes throughout Europe, has published a new climatological 30-year summary of monthly temperature and precipitation concerning the period of 1971–2000 (Alexandersson 2002).

Ångström (1938) presented a Swedish temperature climatology for the period 1901–30. An earlier
climatology was created by Hamberg (1908) for the period 1856–1907. Ångström (1938) compared his and Hamberg’s averages and found that the 1901–30 normals for the winter months (December, January, February) were significantly higher than those of the earlier period. Ångström (1938) concluded that latitude and elevation are the major factors controlling the temperature climate but also stressed the geographically varying maritime influence. According to Ångström (1974), the country is divided into so-called local maritime and local continental regions. Eriksson (1982) described average temperatures for the period 1951–80. He noted that the coldest stations in winter are positioned at locations affected by cold air drainage during inversions and the warmest stations are positioned in coastal areas. The lowest average summer temperatures are obtained at stations at high elevations. However, no systematic analysis of the connections between these factors and temperature was presented. Eriksson (1982) presented a comparison of the 1931–60 and the 1951–80 averages and stated that there had been a cooling of the annual temperature by 0.2–0.7°C in the later period. There was also a significant cooling in July in 1931–60. The most recent temperature normals (1961–90 and 1971–2000) at all Swedish stations have been compiled by Alexandersson (1991) and Alexandersson et al. (2003). All these studies list only station observations.

Latitude mainly reflects the variations in incoming solar radiation and is the most significant large-scale climate factor, especially in Sweden with its great south–north elongation. The longitudinal extent of Sweden is small and longitudinal variations in climate represent the maritime air-mass advection from the North Atlantic. Hudson and Wackernagel (1994) present high correlation coefficients between average temperatures and longitude for January, April and October in Scotland. Elevation is a parameter with local effects and a number of studies show that its effect on temperature is large in certain mid-latitude areas (Hudson and Wackernagel 1994; Zheng and Basner 1996; Chen and Johansson 2003).

Gridded data rather than station data are preferred (e.g. Achberger et al. 2003; Johansson and Chen 2003) for applications such as model validation and spatial interpolation. Global or local fitting techniques are useful for gridding of climate data. One example of a global method is regression analysis, whereas local methods include splines, inverse distance weighting and kriging. Global methods use data from the whole study area to establish relationships and can be used in the presence of trend and large-scale features. In the case of temperature, varying latitude can produce such trends because of varying incoming solar radiation. It is also possible to include information other than geographical position. For example, Norton (1985) relates the temperature normals of New Zealand to latitude, longitude, elevation, distance to sea and distance to the west coast, using multiple regression analysis.

Local methods use information from nearby stations and can capture local anomalies without affecting the interpolation values at other points. Local interpolation techniques have been widely used in interpolation of climate elements (Hevesi et al. 1992; Ishida and Kawashima 1993; Hudson and Wackernagel 1994; Hulme et al. 1995; Zheng and Basner 1996; Pardo-Igúzquiza 1998; Goodale et al. 1998). These studies have been performed in areas of different size and hence the definitions of ‘local’ differ.

ZHENG and Basner (1996) present a spatial modelling of the temperature normals for New Zealand. In their study, the predictors used by Norton (1985) plus a forest and a regional index are incorporated in multivariate linear regression, additive and partial thin-plate smoothing spline models. Thin-plate smoothing splines with the inclusion of elevation are used in European 0.5° × 0.5° (Hulme et al. 1995) and global 0.5° × 0.5° (New et al. 1997) climatologies. Goodale et al. (1998) compare the ability of polynomial regression and the inverse-distance-squared methods to interpolate monthly long-term averages (1951–80) of precipitation, maximum and minimum daily temperatures, and daily solar radiation in Ireland. The prediction accuracy does not differ significantly between the methods but the simplicity of the polynomial regression method is concluded to be favourable.

Variogram analysis and kriging have been used to model climatological and meteorological variables (e.g. Hevesi et al. 1992; Ishida and Kawashima 1993; Hudson and Wackernagel 1994; Pardo-Igúzquiza 1998). The studies differ in the choice of variables, kriging methods and in the characteristics and size of study areas. Ishida and Kawashima (1993) predict hourly surface air temperature in central Japan using different procedures including kriging, cokriging, regression analysis and an inverse-distance-weighted method. They find that cokriging with elevation as a secondary variable...
performs best. January temperature normals (1951–1980 averages) in Scotland are mapped by Hudson and Wackernagel (1994) using kriging with and without elevation as external drift. The best results are obtained when information about elevation is integrated. Hevesi et al. (1992) use kriging and cokriging with elevation as a secondary variable to estimate the average annual precipitation in Yucca Mountain, Nevada. These methods are compared with other interpolation methods, and multivariate geostatistics provide the best cross-validation results (Hevesi et al. 1992). Pardo-Igúzquiza (1998) use the Thiessen method, ordinary kriging, cokriging, and kriging with external drift as estimators of climatological rainfall average in the Guadalhorce River basin, southern Spain. Topographical information is incorporated in the cokriging and kriging with external drift models. Kriging with external drift is reported to give the best results.

Hellström and Chen (1999) studied the Swedish monthly temperature normals during 1961–1990 with ordinary kriging to create a 10 × 10-km gridded dataset. A detailed study of Nordic temperature normals for 1961–1990 with the help of kriging and GIS (Geographic Information System) is presented by Tveito et al. (2000). The present study is aimed at updating these studies and quantifying geographical factors that influenced the spatial patterns of temperature normals in Sweden during 1971 and 2000 by using trend and variogram analyses. The results formed the
basis for creating a gridded dataset for monthly temperature normals covering Sweden with a high resolution (30 × 30 arcsecond or approximately 1 km × 1 km). Such a database can be used in climate-change-scenario construction, interpolation of climate variables, and validation of high-resolution climate models.

Data and methods

Data

Alexandersson (2002) presents temperature normals for the period 1971–2000 for 595 stations covering Sweden. Data from island stations were excluded in this study, which reduced the number of stations to 510. Figure 1 shows the locations of the 510 stations used in this study. The station elevations ranged from 1 to 1280 m with an average of 166 m above mean sea level. The highest station elevations are found in northwestern and central-western Sweden. The rest of the country has a rather low relief. The inland elevation in southernmost Sweden increases to about 300 m.

Latitude and longitude coordinates of the stations were transformed to the Swedish ‘Rikets Nit’ (RT 90) coordinate system. RT 90 is based on a plane reference system and has a reference unit of a metre (Lantmäteriverket 1993). It is created to form a net with right angles on Gauss-conformal projections. The transformation to a plane reference system simplified calculation of distances between stations. The result of the calculation is used in the variogram modelling and interpolation. A digital elevation model (DEM) grid was used in the interpolation of temperature to a database with high spatial resolution. The DEM used in this study was the 30-arcsecond (approximately 1 km) GTOPO30 from the US Geological Survey’s, EROS Data Centre. The height for each temperature-data grid point was obtained from the nearest grid point of the DEM data to be comparable with the station heights used.

Spatial trend analysis

Obvious spatial trends of the temperature normals violate the homogeneity assumption of kriging (see Wackernagel (1995) for methodology). Thus, gridding was done in three steps. Spatial trends were first removed, residuals were then modelled and gridded, and trends were finally added back to the gridded residuals. The final gridded estimates of temperature normals relied on both the global multiple linear regressions and the locally kriged residuals. The method was thus termed composite gridding. Simple and multiple-regression analyses using latitude, elevation and longitude as predictors were applied to the monthly temperature normals to remove spatial trends. Regression analyses were performed on latitude, elevation and longitude, as well as on latitude, longitude and elevation. Three detrended residuals (rsl for latitude, rsle for latitude and elevation, and rslle for latitude, longitude and eleva-
A HIGH-RESOLUTION, GRIDDED DATASET FOR MONTHLY TEMPERATURE NORMALS (1971–2000) IN SWEDEN

An analysis of the temperature–elevation relationship was performed because of the importance of elevation to temperature. Low lapse rates, i.e. the decrease of temperatures with elevation, may indicate the presence of inversion that makes a temperature variation more local. Lapse rates were calculated from the slope of the linear regression line between temperature normals and elevation as a by-product of regression analyses. Lapse rates were calculated separately for stations south of 60°N, between 60°N and 64°N, and north of 64°N to minimize the latitude effect. Coefficients of determination of the regression were used as an uncertainty estimate of the lapse rate.

<table>
<thead>
<tr>
<th>Month</th>
<th>Lapse rate (°C per 100 m elevation increase)</th>
<th>R²</th>
<th>Lapse rate</th>
<th>R²</th>
<th>Lapse rate</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan.</td>
<td>-0.88</td>
<td>0.23</td>
<td>-0.48</td>
<td>0.27</td>
<td>-0.13</td>
<td>0.02</td>
</tr>
<tr>
<td>Feb.</td>
<td>-0.75</td>
<td>0.20</td>
<td>-0.41</td>
<td>0.32</td>
<td>-0.11</td>
<td>0.02</td>
</tr>
<tr>
<td>Mar.</td>
<td>-0.57</td>
<td>0.23</td>
<td>-0.50</td>
<td>0.63</td>
<td>-0.37</td>
<td>0.31</td>
</tr>
<tr>
<td>Apr.</td>
<td>-0.32</td>
<td>0.09</td>
<td>-0.51</td>
<td>0.66</td>
<td>-0.55</td>
<td>0.60</td>
</tr>
<tr>
<td>May</td>
<td>-0.12</td>
<td>0.01</td>
<td>-0.46</td>
<td>0.44</td>
<td>-0.62</td>
<td>0.58</td>
</tr>
<tr>
<td>Jun.</td>
<td>-0.32</td>
<td>0.15</td>
<td>-0.53</td>
<td>0.62</td>
<td>-0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>Jul.</td>
<td>-0.50</td>
<td>0.47</td>
<td>-0.59</td>
<td>0.82</td>
<td>-0.69</td>
<td>0.82</td>
</tr>
<tr>
<td>Aug.</td>
<td>-0.70</td>
<td>0.48</td>
<td>-0.62</td>
<td>0.85</td>
<td>-0.63</td>
<td>0.83</td>
</tr>
<tr>
<td>Sep.</td>
<td>-0.83</td>
<td>0.40</td>
<td>-0.60</td>
<td>0.80</td>
<td>-0.53</td>
<td>0.72</td>
</tr>
<tr>
<td>Oct.</td>
<td>-0.85</td>
<td>0.33</td>
<td>-0.57</td>
<td>0.70</td>
<td>-0.48</td>
<td>0.45</td>
</tr>
<tr>
<td>Nov.</td>
<td>-0.94</td>
<td>0.30</td>
<td>-0.61</td>
<td>0.50</td>
<td>-0.40</td>
<td>0.17</td>
</tr>
<tr>
<td>Dec.</td>
<td>-0.93</td>
<td>0.22</td>
<td>-0.53</td>
<td>0.30</td>
<td>-0.17</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Bold-italic font indicates non-significant results

Table 2. Monthly semivariogram models and parameter values in Fig. 4

<table>
<thead>
<tr>
<th>Range (km)</th>
<th>Residual 1</th>
<th>Residual 2</th>
<th>Residual 3</th>
<th>Residual 1</th>
<th>Residual 2</th>
<th>Residual 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan.</td>
<td>124.76</td>
<td>131.41</td>
<td>127.56</td>
<td>2.42</td>
<td>2.10</td>
<td>2.07</td>
</tr>
<tr>
<td>Feb.</td>
<td>104.96</td>
<td>111.78</td>
<td>114.01</td>
<td>1.36</td>
<td>1.16</td>
<td>1.17</td>
</tr>
<tr>
<td>Mar.</td>
<td>120.83</td>
<td>127.97</td>
<td>132.35</td>
<td>0.84</td>
<td>0.39</td>
<td>0.35</td>
</tr>
<tr>
<td>Apr.</td>
<td>162.67</td>
<td>94.49</td>
<td>57.96</td>
<td>1.08</td>
<td>0.39</td>
<td>0.30</td>
</tr>
<tr>
<td>May</td>
<td>153.32</td>
<td>116.32</td>
<td>99.97</td>
<td>1.71</td>
<td>1.06</td>
<td>0.89</td>
</tr>
<tr>
<td>Jun.</td>
<td>212.41</td>
<td>115.17</td>
<td>111.43</td>
<td>1.90</td>
<td>0.71</td>
<td>0.70</td>
</tr>
<tr>
<td>Jul.</td>
<td>243.48</td>
<td>89.53</td>
<td>90.19</td>
<td>1.80</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>Aug.</td>
<td>209.05</td>
<td>11.92</td>
<td>0.69</td>
<td>1.67</td>
<td>0.24</td>
<td>0.22</td>
</tr>
<tr>
<td>Sep.</td>
<td>157.21</td>
<td>35.03</td>
<td>34.65</td>
<td>1.44</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>Oct.</td>
<td>126.92</td>
<td>71.65</td>
<td>74.22</td>
<td>1.36</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>Nov.</td>
<td>128.03</td>
<td>122.97</td>
<td>123.59</td>
<td>2.11</td>
<td>1.39</td>
<td>1.33</td>
</tr>
<tr>
<td>Dec.</td>
<td>119.54</td>
<td>118.11</td>
<td>115.69</td>
<td>2.75</td>
<td>2.29</td>
<td>2.27</td>
</tr>
</tbody>
</table>

Model: Semivariance = (1 – exp(-lag / Range)) * Sill

Residual 1, Residual 2 and Residual 3 refer to temperature residuals detrended by latitude only, by latitude and elevation, and by latitude, longitude and elevation.

Semivariogram modelling and kriging

In the present study, semivariogram analysis was used to examine the range of the spatial continuity of the normals and hence to identify the major scales (spatial variability) in the observed temperature–climate data. While this idea is widely accepted, its application in climate research is less frequent (e.g. Gunst 1995). There are various analysis techniques available for this purpose. Semivariograms are relatively widely used and are equivalent to the covariances that describe links between two variables. In the context of spatial variability, the two variables are the same entity at different locations (spatial autocorrelation). The scale is described by a spatial continuity that can be expressed by spatial correlation. Data points...
closer to each other in space tend to be more similar. A semivariogram measures the variance between data points as a function of separation distance \( h \) (the lag distance). The semivariogram \( \gamma \) is expressed as:

\[
\gamma(h) = \frac{1}{2N(h)} \sum [(z(x_i) - z(x_i + h))^2]
\]

where \( z(x_i) \) is the variable considered at location \( x_i \), \( h \) is the separation vector and \( N \) is the number of observations. The semivariogram is related to the covariance (cov) in the following way:

\[
\gamma(h) = \text{cov}(0) - \text{cov}(h)
\]

(2)

Possible changes in the averages (non-ergodic relationships) can be taken into account by using the covariance instead of \( \gamma \). Dividing covariance by the variance of the data gives the correlogram, a normalized measure traditionally used by time-series analysts (called correlation function in the meteorological objective analysis; e.g. Sen and Habib 2000). A correlogram provides a good way to compare spatial correlation structures of climate variables with different magnitudes and units.

Fig. 3. Correlograms of monthly original and detrended temperature data by latitude, latitude and elevation, and longitude, latitude and elevation
Correlograms were set up for both original temperature normals and detrended residuals ($rsl$, $rsle$ and $rslle$) for each month. The spatial structures of the detrended residuals were modelled with semivariograms. Both power and exponential variogram models were tested.

The interpolated temperature normals were validated by cross-validation, where one station observation was omitted at a time and kriged with the other observations within the correlation length. Root-mean-squared errors (RMSE) for the estimated values were then calculated against observed values:

$$RMSE = \left( \frac{1}{N} \sum (P_i - O_i)^2 \right)^{1/2}$$

where $P_i$ is the estimated value, $O_i$ is the observed value, and $N$ is the total number of values.

Finally, kriging was performed on the residuals, which are as homogeneous as possible with all known local trends removed. The composite-kriging interpolation with the best residual was used to grid the temperature normals. The residual was first interpolated to the DEM grid of the GTOPO30 dataset with a grid spacing of 30 arc-seconds (about 1 km), and the gridded temperature estimates were then obtained by adding the trends.
back to the interpolated residuals. The original temperature normals with trend were also kriged (called simple gridding in this study) to serve as a performance benchmark for the composite gridding. The differences between the simple and composite griddings shed light on the influences of geographic factors and helped determining which of them should be used in the interpolation to create the final dataset.

Results

Trend analysis

The strong latitudinal gradient of net incoming radiation in winter (e.g. Cambell and Vonder Haar 1980) could explain the very high coefficients of determination between monthly winter temperatures and latitude (Fig. 2). The temperature variation seemed mainly controlled by local elevation in the summer when the radiation gradient is small. Latitude and el-
evation could together explain around 90% of the temperature variation. Longitude had a weak negative correlation with temperature normals for all months.

As for elevation effects, it was interesting to note that southern Sweden had the minimum lapse rates in the warm part of the year and the maximum towards the cold months, while northern Sweden had an opposite seasonal cycle (Table 1). High lapse rates are often accompanied by high coefficients of determination and thus low uncertainties in the middle and northern parts of Sweden.

**Semivariogram models**
The only major global correlation is found between temperature and latitude. This global correlation structure controls the long spatial correlation range
in winter and the short range in summer. Removal of the global trend with respect to latitude results in a major drop in the correlation, indicating the much more local spatial pattern of $r_{sl}$ (Fig. 3). $r_{sl}$ shows a consistent 200-km correlation range throughout the year, mainly affected by elevation. $r_{sle}$, obtained by removing elevation trends from $r_{sl}$, lowers the correlation only in the warm months (June to September) when the correlation between elevation and temperature is the highest. It is difficult to distinguish between $r_{sle}$ and $r_{sle}$ at any time of the year, which confirms the negligible longitude influence on Swedish temperature normals. The power semivariogram model is not very successful, but the exponential semivariogram model (Table 2) fits the measured data well (Fig. 4). In the semivariogram models the nugget effect has been neglected although some of the experimental variograms indicate such a discontinuity at the origin. Removing strong spatial trends can largely reduce the sill and, at the same time, shorten the range.  

**Kriging and cross-validation**

Cross-validation results (RMSE) for simple kriging of the original temperature normals, and for composite kriging of detrended residuals ($r_{sl}$, $r_{sle}$ and $r_{sle}$), show that the errors are distinctly higher in winter for all methods (Fig. 5). The composite method gave much smaller errors in the remaining months of the year (March–October). It is notable that removal of the latitude trend gave no improvement compared to simple kriging. It is the removal of local elevation trends that changes the amount and distribution of the cross-validation error and improves the overall interpolation quality. The largest improvements occur in those months when the elevation–temperature relationships are strong.

**Creation of the gridded dataset**

Since only latitude and elevation are significantly related to temperature, these trends were removed
before kriging, i.e. the composite-kriging interpo-
lation with the \textit{rsle} residual was used to grid the
temperature normals to a 30 × 30 arcsecond (ap-
proximately 1 km × 1 km) resolution over all of
mainland Sweden. As an example, Fig. 6 shows the
spatial distribution of the seasonal temperature nor-
ma ls over Sweden.

\begin{image}
\centering
\includegraphics[width=\textwidth]{fig6}
\caption{Swedish seasonal temperature normals for 1971–2000 s at a 30 × 30 arcs econd (approximately 1 km × 1 km) resolution derived by kriging on original data detrended from latitude and elevation influences.}
\end{image}

\section*{Conclusions}
Most spatial variations of the average Swedish tem-
perature, in particular during the cold season, could
be explained by latitude, while elevation was a ma-
jor factor determining summer temperatures.
Throughout the year, 82–96% of the spatial vari-
ance in monthly normal temperatures could be ex-
Deliang Chen, Lebing Gong, Chong-Yu Xu and Sven Halldin

plained by latitude and elevation. The influence of longitude was minor and could only be seen in spring. Therefore, only latitude and elevation are considered important in determining spatial trends in Swedish temperatures. It was demonstrated that kriging with data, detrended with respect to latitude and elevation, had an overall better performance than kriging with the original temperature data. The overall performance of the kriging of the normal temperature residuals was fairly good (average RMSE = 0.5°C). The performance during March–October was much better than for the remaining months. It was concluded that interpolation with detrended data could be used for a gridded temperature climatology for Sweden. Such a monthly climatology with a grid spacing of 30 arcseconds (about 1 km) has been created and made available on the Internet.

Acknowledgement

Hans Alexandersson from SMHI is thanked for kindly providing the temperature data at all Swedish stations. This work is supported by a VR grant to Deliang Chen. Regine Hock’s comments improved the manuscript. This is contribution No. 3 from TELLUS, the centre for Earth System Science at Göteborg University.

Deliang Chen, Earth Sciences Centre, Göteborg University, Box 460, SE-405 30, Göteborg, Sweden E-mail: deliang@gvc.gu.se

Lebing Gong, Air and Water Science, Department of Earth Sciences, Uppsala University, Sweden

Chong-Yu Xu, Department of Geosciences, University of Oslo, Norway

Sven Halldin, Air and Water Science, Department of Earth Sciences, Uppsala University, Sweden

References


A HIGH-RESOLUTION, GRIDDED DATASET FOR MONTHLY TEMPERATURE NORMALS (1971–2000) IN SWEDEN


Manuscript received Nov. 2006, revised and accepted May 2007.