INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **23**: 1495–1521 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/joc.945

WERE SOUTHERN SWEDISH SUMMER TEMPERATURES BEFORE 1860 AS WARM AS MEASURED?

ANDERS MOBERG,[a,*] HANS ALEXANDERSSON,[b] HANS BERGSTRÖM[c] and PHILIP D. JONES[a]
[a] *Climatic Research Unit, University of East Anglia, Norwich, UK*
[b] *Swedish Meteorological and Hydrological Institute, Norrköping, Sweden*
[c] *Department of Earth Sciences, Meteorology, Uppsala University, Uppsala, Sweden*

Received 13 September 2002
Revised 10 June 2003
Accepted 10 June 2003

ABSTRACT

Temperature series from Stockholm and Uppsala in southern Sweden indicate that summers from the mid-18th century until around 1860 were, on average, warmer than the 1961–90 mean. The station histories suggest that the early observations could have been positively biased, for example because of insufficient radiation protection. We investigate if independent support for warm summers in the early period can be obtained from other climate variables. Using stepwise multiple regression analysis we investigate nine potential predictor variables: six air circulation indices, precipitation, air pressure and cloud amount. Three of these variables — cloud amount (the most important one), meridional geostrophic wind, and air pressure — together explain 65% of the June–August temperature variance in the calibration period 1873–2000. Application of the regression relationship back to 1780 shows that the model is equally successful in predicting year-to-year temperature variability before 1873 as it is in the calibration period, whereas the low-frequency component is poorly reconstructed in the early period. This reduced skill is primarily due to poorer data quality of the predictor variables in the early period, in particular the cloud amount series. The observed decadal mean temperatures during 1780–1860 are found to be above the upper limit of a 95% confidence interval that accounts for uncertainties both in the regression relationship and in the cloud amount series. We conclude that the observed temperatures before around 1860 are, therefore, most likely positively biased. The size of this bias cannot be accurately determined from the evidence used here, but seems to be about 0.7–0.8 °C for both stations. A comparison with long instrumental temperature series from central Europe suggests a slightly smaller bias (0.5–0.6 °C). For more accurate assessment of the Stockholm and Uppsala temperatures, we recommend that extensive homogeneity testing of other long northern European temperature series are undertaken.

KEY WORDS: Sweden; multiple regression; standard normal homogeneity test; early instrumental data; temperature; air circulation; cloud amount

1. INTRODUCTION

It is well established that the global mean temperature has warmed by about 0.6 °C since the late 19th century and that at least a significant part of this warming is likely due to anthropogenic emissions of greenhouse gases, although natural factors may also have contributed (Houghton *et al*., 2001). In the search for further understanding of the underlying causes there is a need to reconstruct climate before the onset of any anthropogenic effect. Only Europe has a large number of instrumental records beginning before the mid-19th century. Prior to using these early data in climate change studies, it is essential to investigate their quality and limitations thoroughly. Although there has been a renewed interest in the early European records in the last decade (e.g. Jones *et al*., 1999; Böhm *et al*., 2001; Camuffo and Jones, 2002), there still remain

* Correspondence to: Anders Moberg, Climatic Research Institute, University of East Anglia, Norwich NR4 7TJ, UK;
e-mail: a.moberg@uea.ac.uk

questions concerning the homogeneity even of some quite well investigated series. For some other series, no quality controls have been undertaken at all with modern methods.

In this paper, we investigate summer temperature data in two of the most widely used long European instrumental records, namely those for Stockholm (Moberg *et al.*, 2002) and Uppsala (Bergström and Moberg, 2002) in southern Sweden. It has been noted (Wallén, 1953; Moberg and Bergström, 1997; Moberg *et al.*, 2002) that observed summer temperatures at these sites before around 1860 were markedly warm. Decadal averages were nearly always above the 1961–90 average, with some peaks even above the 20th century warm peaks. Despite considerable efforts homogenizing the data, it can be questioned whether the pre-1860 summer temperatures really were as warm as the data suggest. There are two good reasons for this doubt. One is the fact that several changes in observing conditions took place at both sites in the 1850s to 1870s, e.g. introduction of radiation screens and changes of observing hours at both sites and site relocations at Uppsala. The second reason is that a documentary proxy-data reconstruction (based on sowing dates in farmers' diaries) of warm-season temperatures for southeastern Norway, extending back to 1749 (Nordli, 2001a), indicates that summers there were generally *cooler* before 1860 than the 1961–90 average. Similar reconstructions back to the 1820s for western Norway (Nordli, 2001b) also suggest cool summers before the 1860s. These proxy-data evidence from Norway are in conflict with the instrumental records from Stockholm and Uppsala.

To investigate if the observed warm summers in Stockholm and Uppsala before around 1860 are real or artefacts, it would be useful to develop independent long series of *estimated* temperatures for Stockholm and Uppsala, which can then be used for direct comparison and as reference series in statistical homogeneity tests. Such reference series could either be constructed from observations of other climate variables than temperatures, or they could consist of temperature observations from other sites. The main requirements, regardless of the kind of reference data, are that they should be strongly correlated with temperatures at Stockholm and Uppsala and that the reference series should be homogeneous, i.e. only show true climate and weather variability. The main focus in this paper is to determine whether it is possible to find a combination of other climate variables that fit these requirements. If this is possible, then we should be able to answer the question about whether the summer temperatures before around 1860 were positively biased. If the answer is 'yes', then we could also ask: Exactly when and by how much? We do not undertake any detailed comparison with other long European temperature series, although a brief comparison is made to make the picture more complete. The reason we do not investigate other long temperature series in detail here is that many of these series still remain to be homogenized, and this is beyond our scope.

Summer temperature changes in southern Scandinavia are influenced by changes in the atmospheric circulation over Europe (Wallén, 1953). Changes in cloudiness also have an influence on temperatures in this region (Ångström, 1946). Precipitation is another potentially important variable, as warm or cool summer temperatures are likely to be associated with dryness or wetness respectively. Long records of air circulation indices (derived from gridded sea-level pressures) and local series of cloud amount, precipitation and air pressure at both Stockholm and Uppsala are, therefore, chosen as potential reference data in this study. Our main study period is 1780–2000, although we show observed temperatures back to 1756.

Linear regression methods for studying statistical relationships between temperatures in southern Sweden and air circulation over this region have been found to be successful for winter data (Alexandersson, 1994; Chen, 2000). Here, we adopt a multiple regression method that is similar to that used by Chen (2000); but, in addition to the air circulation index series he used, we also include the local cloud amounts, precipitation and air pressure data among the potential predictor variables in our analyses of relationships in summer.

Monthly gridded pressure data, as well as station air pressure and precipitation data, were already available back to 1780, but to include cloud amount in the analysis we had to reconstruct such series directly from the original observations. We describe how this was achieved in Appendix A. In the main text, we discuss homogeneity problems for these cloud amount records and we suggest a method for homogenizing them.

Although we focus on summer data, we analysed all 12 months to compare the strength of relationships between climate variables also in the other seasons to obtain a more complete picture of data quality problems. The results presented here in figures and tables, however, are only for the summer months of June, July and August (JJA). Results for other seasons are sometimes mentioned in the text, but only when this add some complementary and useful information.

This paper is organized as follows. Section 2 provides details of our reasons to doubt the warm summers before around 1860. Section 3 presents the various data series. Section 4 analyses the cloud amount series. Section 5 contains the analysis of relationships between temperatures and other climate variables, and attempts to develop a multiple regression model for estimation of summer temperatures. Section 6 discusses the main problem of judging the veracity of the pre-1860 observed warm temperatures, including a comparison with long temperature series from central Europe. The conclusions are drawn in Section 7.

2. REASONS TO DOUBT THE VERACITY OF THE WARM SUMMERS OBSERVED BEFORE 1860

Daily mean air temperature series for Stockholm and Uppsala have recently been developed (Bergström and Moberg, 2002; Moberg *et al.*, 2002) back to 1756 and 1722 respectively. Here, we analyse monthly mean temperatures calculated from the daily means for 1756–2000. Mean summer (JJA) temperatures, averaged as $0.5 \times (T_{\mathrm{Stockholm}} + T_{\mathrm{Uppsala}})$, are shown in Figure 1(a). The smoothed thick line is obtained by applying a Gaussian filter with a standard deviation of 3, roughly corresponding to a 10 year moving average. We use the same filter in all figures.

The two stations are located only 70 km apart and their temperature series are very similar (the correlation is 0.99 for 20th century data and 0.93 before 1860). Separate plots of each station series (not shown) are, therefore, very similar to the two-station average. In particular, both series agree very well about the notably warm level before 1860 compared with afterwards. Decadal averages, illustrated by the smooth curve in Figure 1(a), are nearly always above the 1961–90 average before around 1860.

As mentioned in Section 1, there are reasons to doubt whether the warm level before around 1860 is real or an artefact. After this year, both series have been homogenized by Moberg *et al.* (2002) and Bergström and Moberg (2002) and agree well with series from rural surroundings. Unfortunately, no nearby reference series were available to Moberg *et al.* (2002) and Bergström and Moberg (2002) before 1861, so homogeneity tests based on comparisons with nearby stations could not be performed prior to this year. However, adjustments have been applied to account for factors such as changes of observation hours and thermometers with systematic errors (see extensive discussion by Moberg *et al.* (2002) and Bergström and Moberg (2002)). Below, we provide details about our arguments for suspecting a positive bias of summer temperatures observed at both Stockholm and Uppsala before around 1860.

2.1. The station history argument

Observations at Stockholm have always been made at one and the same site (see Moberg *et al.* (2002)), i.e. the old astronomical observatory. The year 1859 is particularly important in the station history for this site, because the 'modern' national meteorological observation network was initialized and observing routines in Stockholm began to follow 'new' national standards that year. In particular, the hour for morning observations changed from 06h to 08h (local time). The thermometer was placed on a north-facing wall all the time from 1756 to 1960, and this wall should have been hit by sunlight at 06h, but not at 08h, during the summer months from May to August when the sun is above the horizon at 06h. Hence, morning temperatures in these months could be positively biased before 1859. The change in observing routines, which also included the introduction of daily minimum temperature observations, caused a necessary change of equation for calculation of daily mean temperatures by Moberg *et al.* (2002), which is a further potential source of homogeneity break between 1858 and 1859.

A radiation screen was introduced in 1878. This could have led to an artificial cooling compared with the earlier measurements. There is no information indicating that radiation screens were used before 1878, although the thermometer was claimed to be 'well protected from the morning sun' (see Moberg *et al.* (2002)). Even if there was protection against direct sunlight, the observed temperatures could anyway have been biased due to reflection of sunlight and radiation from the surroundings. In particular, we suspect that the wall where the thermometer was placed was heated by sunshine at the time when the morning observations were made before 1859. Radiation from a heated wall, and also warm air rising from below the thermometer (which was placed nearly 6 m above ground), could have caused observed temperatures to be too high.

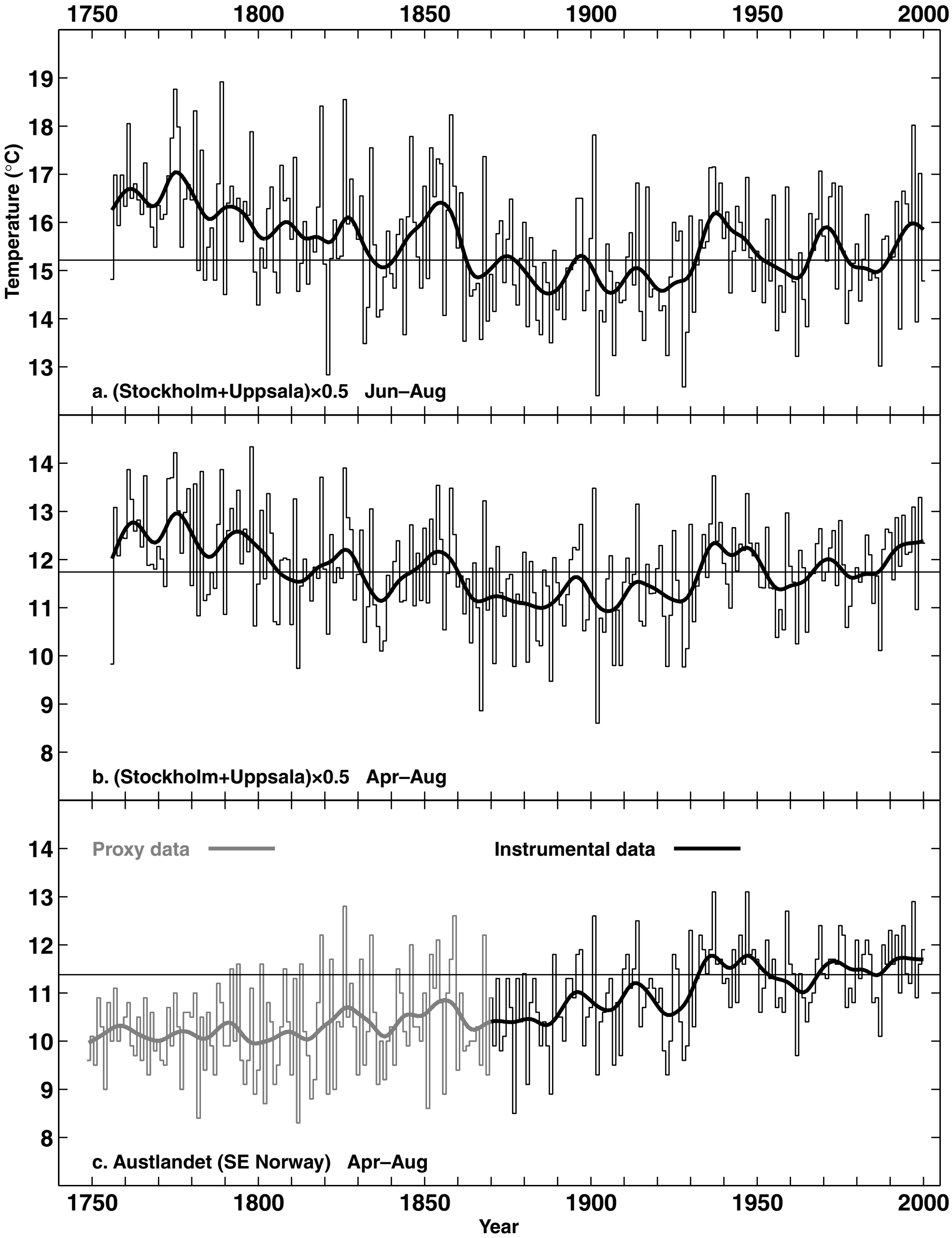

Figure 1. Time series of seasonal mean temperatures (°C). (a) Arithmetic average of Stockholm and Uppsala JJA mean temperatures 1756–2000 (from Moberg *et al*. (2002) and Bergström and Moberg (2002)). (b) As (a), but April–August average. (c) April–August mean temperature reconstruction for Austlandet, SE Norway, 1749–2000 (from Nordli (2001a)), based on harvest dates as temperature proxy data 1749–1870 (grey) and instrumental data 1871–2000 (black). Annual values are shown with bars. The thick smooth curve shows data filtered with a Gaussian filter having a standard deviation of 3, approximately corresponding to a 10 year moving average. The horizontal lines show the 1961–90 averages

At Uppsala the situation is somewhat different (see Bergström and Moberg (2002)). From 1739 to August 1853, observations were made at the old astronomical observatory in the old town. Few details are known about how thermometers were exposed in this period, except some notes that they were placed outdoors in the shadow. After a relocation in September 1853 to a site just outside the old town, a 'Lawson shelter' was introduced. In October 1865, the station was moved again, although only about 100 m. The first Stevenson-like screen (but larger than the modern standard) was introduced in August 1868. Another screen was used from August 1874. The introduction of a simple shelter in 1853, and in particular a more efficient radiation screen in 1868, may have introduced artificial cooling.

A change of observation hour in the morning from 07h to 08h took place in 1863, and since June 1865 the observations have been made every hour. These changed procedures may have caused homogeneity breaks in the mid-1860s, although they could be of either sign. More importantly, we suspect that temperatures before September 1853 could have been positively biased because an early urban warming may have existed inside the town. In fact, Bergström and Moberg (2002) found that Uppsala temperatures before September 1853 were warmer than the corresponding Stockholm temperatures in a relative sense. Hence, they adjusted all Uppsala temperature before September 1853 to force the long-term averages to be the same at both sites. The adjustments (table I in Bergström and Moberg (2002)) have an irregular annual cycle with a maximum of $-0.66\,^{\circ}$C in March and a minimum of $-0.03\,^{\circ}$C in August ($-0.15\,^{\circ}$C in June, $-0.11\,^{\circ}$C in July). As the adjustment is small in summer, it is not important here whether we analyse the adjusted or unadjusted Uppsala data. The Uppsala series agree very well with Stockholm concerning the warm summers before 1860 — both with and without the correction imposed by Bergström and Moberg (2002). We use their adjusted data in this paper.

From all details provided in this section, it follows that a number of changes in observing practices occurred at both sites in the 1850s to 1870s, and some of them may certainly have caused artificial cooling of the observed temperatures. The question of possibly warm-based temperatures was also asked by Moberg *et al*. (2002) and Bergström and Moberg (2002). They asked whether *direct* sunshine could have hit the thermometers, and argued that a bias in such a case should be larger for clear-sky conditions than for cloudy conditions. Therefore, they separated all days into three categories (clear, half-clear, cloudy) and calculated separate summer mean temperature series for each class. They found no evidence of relatively warmer early summer temperatures for clear-sky days compared with cloudy days (in fact rather the opposite), and hence concluded that there was no significant problem with *direct* sunlight. In the present study, however, we hypothesize that positive biases may be caused by diffuse light, reflection and radiation from the surroundings if the radiation protection was insufficient. We do not make any distinction here between possibly different biases under different weather conditions.

2.2. *The proxy data argument*

Nordli (2001a) has developed an April-to-August temperature reconstruction for Austlandet in southeastern Norway 1749–2000. The series is a combination of instrumental temperature data for the period 1871–2000 and reconstructed temperatures based on documentary data for 1749–1870. The proxy data are the first harvest dates each year at eight farms in southeastern Norway and two nearby farms in western Sweden. There is no period when all individual farm series overlap, but they are connected through successive overlaps. The harvest dates were calibrated against the instrumental April–August mean temperatures. We show the Austlandet series in Figure 1(c), with the proxy data part in grey and the instrumental part in black. The graph immediately above (Figure 1(b)) shows the corresponding April–August mean temperatures averaged for Stockholm and Uppsala.

After 1871 the series in Figure 1(b) and (c) are very similar, with a century-scale warming trend of about the same size ($+1.03\,^{\circ}$C/century at Austlandet, $+0.84\,^{\circ}$C/century at Stockholm/Uppsala) and a correlation coefficient of 0.89 between the two ($\sim$400 km apart). The strong correlation also holds after detrending ($r = 0.88$). Given the strong correlation after 1871, one would also expect the two series to agree before this year, but this is clearly not the case. In particular, the overall temperature levels before around 1860 differ markedly, with Austlandet being significantly below the 1961–90 average and Stockhom/Uppsala fluctuating around this level during 1800–60 and being clearly above it in the earlier decades. The correlation for the 1756–1870 period is 0.62, which is substantially less than afterwards. This reduced correlation certainly reflects a lower fidelity of year-to-year fluctuations in the proxy data used. Nevertheless, the obvious disagreement with Stockholm and Uppsala before the 1860s is puzzling and provides a strong argument for investigating the early summer temperature data further.

3. DATA

This section describes the various datasets used in studies of relationships with temperatures in Stockholm and Uppsala.

3.1. Cloud amount data

Observations of clouds have been made at Stockholm and Uppsala since 1756 and 1722 respectively. Bergström and Moberg (2002) and Moberg *et al.* (2002) developed simple daily records of cloudiness, classifying each day as being clear, half-clear or cloudy. They used these records as an aid in homogeneity checks of daily summer temperatures, but they did not study changes in cloudiness itself. Here, we analyse the original cloud observation data in an attempt to reconstruct monthly mean cloud amounts expressed as percent cloud cover at Stockholm 1756–2001 and Uppsala 1780–1995. This reconstruction includes development of a schedule for conversion of descriptive cloud observations made before the mid-19th century to numeric cloud amounts, and takes into account the diurnal cycle of cloud amounts to adjust for the various sets of observation hours used. Details of this analysis are given in Appendix A. The monthly cloud amount series developed in Appendix A are henceforth referred to as the raw cloud series. They are 'raw' in the sense that they have been developed directly from the cloud observations, without the help of any other climatic variables.

The raw cloud series for Stockholm and Uppsala are compared with a regional cloud amount record for Fennoscandia (Norway, Denmark, Sweden, Finland) 1890–1995, based on 18 to 37 stations (Tuomenvirta *et al.*, 2000) and with a weighted average for five stations in southern Sweden 1873–2000 (Falun, Göteborg, Karlstad, Växjö, Visby). The data source for the southern Sweden series is the databank of the Swedish Meteorological and Hydrological Institute (SMHI). Slightly shorter versions (1890–1999) of the same series are available in the data set of Tuomenvirta *et al.* (2001). No homogenization had been undertaken previously to any of the cloud amount records used here, apart from some Danish station records included in the Fennoscandian average.

3.2. Circulation data

Six indices of the atmospheric circulation over the region between 0–30 °E and 50–70 °N, which is centred near Stockholm and Uppsala, are derived from gridded monthly mean sea-level pressure (MSLP) data 1780–2000. The six indices, defined by Chen (2000), are the zonal (westerly) component of geostrophic wind u, the meridional (southerly) component of geostrophic wind v, the geostrophic wind speed $V(V^2 = u^2 + v^2)$ the westerly shear vorticity (zonal gradient of u) ξ_u, the southerly shear vorticity (meridional gradient of v) ξ_v and the total shear vorticity $\xi(\xi = \xi_u + \xi_v)$.

The index time series for 1873–2000 are calculated from the United Kingdom Met Office (UKMO) Northern Hemisphere gridded (5 °lat. × 10 °lon.) monthly MSLP data (http://www.cru.uea.ac.uk/cru/data/pressure.htm). The sources of the original pressure chart data used are given in Jones (1987). These grid-point data are of good quality over our study region because of the dense observation network available (Jones *et al.*, 1999). Circulation indices for the period 1780–1872 were derived from the MSLP dataset developed by Jones *et al.* (1999), who used a network of 10 to 51 monthly pressure series as predictors, with the UKMO MSLP data being the predictand, to reconstruct gridded pressure data over the region 35–70 °N, 30 °W–40 °E using orthogonal spatial regression. The explained variances in their analysis ranged from 90% in January to 70% in July around the entire European region when all 51 stations were available. Reconstruction quality reduces during years before 1821, when <17 stations were used, but remains high (∼60–95%) for our regional subset because long records were available in northern Europe (e.g. Trondheim, 1768, Edinburgh 1770, Lund, 1780, and Gdansk, 1802). To avoid artificial breaks in mean values in the index time series at the time point

(1872–73) that separates the Jones *et al*. (1999) data from the UKMO data, we adjusted all monthly mean values for 1780–1872 (derived from Jones *et al*. (1999)) by the average difference between the UKMO and Jones *et al*. (1999) data calculated for the period with overlapping data 1873–1995.

As additional measures of the local air pressure variations, we used the monthly MSLP records for Stockholm and Uppsala 1780–2000, calculated from the daily MSLPs by Moberg *et al*. (2002) and Bergström and Moberg (2002). These data were not available to Jones *et al*. (1999).

3.3. Precipitation data

Quantitative precipitation measurements have been made at Uppsala since 1722 and at Stockholm since 1785. The observed precipitation records at both sites have severe undercatch problems in their early parts; before 1836 in Uppsala (Tabony, 1980; Eriksson, 1981) and before 1893 in Stockholm (Eriksson, 1981). Despite these shortcomings, we use the records in the period 1780–2000 to make our analysis as complete as possible regarding relevant available variables. The observed early data were first adjusted to compensate for the undercatch. [*Uppsala 1780–1835, all observed monthly values are multiplied by 1.38. Stockholm 1812–70, monthly multiplication factors: 1.40 (J), 1.40 (F), 1.40 (M), 1.15 (A), 1.10 (M), 1.10 (J), 1.10 (J), 1.10 (A), 1.10 (S), 1.10 (O), 1.10 (N), 1.30 (D). Stockholm 1860–93, monthly multiplication factors: 1.50 (J), 1.50 (F), 1.40 (M), 1.15 (A), 1.10 (M), 1.05 (J), 1.10 (J), 1.10 (A), 1.15 (S), 1.20 (O), 1.20 (N), 1.30 (D). Note: double factors used 1860–70*].

The data after 1860 have also been homogeneity tested in comparison with other Swedish precipitation series and adjustments in the relatively recent period 1958–1979 were applied to the Uppsala data. [*Uppsala 1958–79, monthly multiplication factors: 1.10 (J), 1.10 (F), 1.10 (M), 1.10 (A), 1.10 (M), 1.05 (J), 1.00 (J), 1.05 (A), 1.05 (S), 1.10 (O), 1.10 (N), 1.10 (D).*] For Stockholm, the adjustments to JJA data are mostly smaller than a factor 1.10, whereas for Uppsala the adjustments prior to 1836 are 1.38. (These adjustments are, however, not that important for this study, because precipitation data before 1894 are only used for calculation of correlation coefficients with other variables, and not for calculations of trends or other changes in mean values.)

4. ANALYSIS OF CLOUD OBSERVATIONS

In this section we analyse the 'raw' JJA average cloud amount series for Stockholm (1780–2000) and Uppsala (1780–1995) developed in Appendix A. We compare these series with a regional cloud amount series for Fennoscandia (Tuomenvirta *et al*., 2000) and a southern Swedish regional series, both starting in the late 19th century. Attempts are also made to estimate cloud amounts using other climate variables to obtain independent cloud estimates back to 1780. A comparison between cloud amounts predicted from such estimates and the four observational cloud series leads to a discussion of homogeneity issues. It is concluded that the raw cloud series need to be adjusted. We propose how this can be done, and we develop 'adjusted' cloud series for the 1780–2000 period. We point out the consequences of these adjustments with respect to the main goal of this paper, i.e. to test the veracity of the observed warm summer temperatures before around 1860.

4.1. Raw cloud amount series for Stockholm and Uppsala and comparison with regional cloud series

Time series for the raw JJA average cloud amount series are shown in Figure 2, with Stockholm in blue and Uppsala in orange. The data are presented as low-pass filtered series in Figure 2(a) and their high-frequency components are shown in Figure 2(b). The same filter as used in Figure 1 defines the low-frequency components. The high-frequency components are obtained by subtracting the low-frequency components from the original series.

The two other observational series are plotted similarly, with the Fennoscandian series in black and the southern Sweden series in violet. The red and green curves depict two different regression models for estimating cloud amounts from other variables. The development of these is described in Section 4.2. Table I gives correlation coefficients, both for the high- and low-frequency components of the cloudiness series,

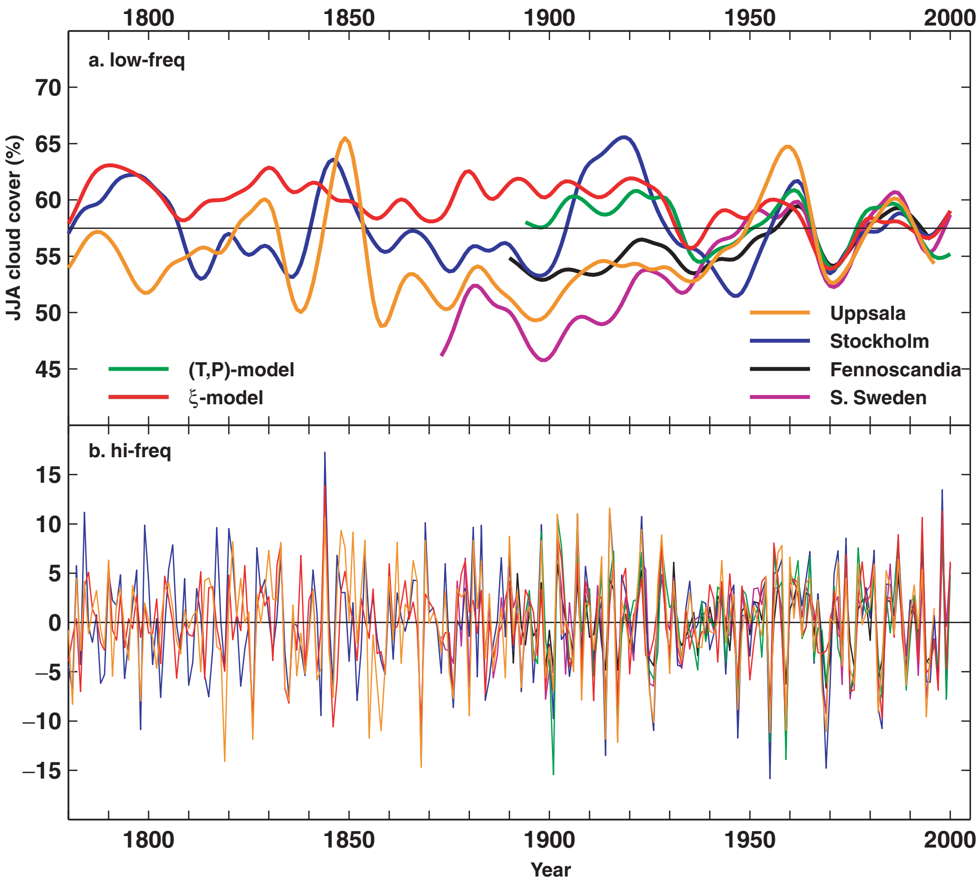

Figure 2. Observed and estimated JJA mean cloud amounts (percent cloud cover). Observational data for Uppsala 1780–1995 (orange), Stockholm 1780–2000 (blue), a Fennoscandian average 1890–1995 (black) and a southern Sweden regional average 1873–2000 (violet). Estimated cloud amounts predicted by a vorticity index series (ξ-model) 1780–2000 (red) and a combination of temperature and precipitation ((T, P)-model) 1894–2000 (green). (a) Low-frequency component obtained with the same filter as in Figure 1. (b) High-frequency component obtained by subtracting the low-frequency component from the original data. All series in (a) are adjusted to have the same 1961–90 average as Stockholm, given by the horizontal line. See main text for detailed descriptions of the series

Table I. Correlation coefficients (×100) for the JJA average raw cloud amounts (percent cloud cover) in Stockholm and Uppsala with each other and with four other series in two periods. Correlations are given separately for high-frequency (time scales <10 years) and low-frequency (>10 years) components in the time series. The six series and the high- and low-frequency decomposition are explained in the main text

	1900–1995				1780–1899			
	High frequency		Low frequency		High frequency		Low frequency	
	Stockholm	Uppsala	Stockholm	Uppsala	Stockholm	Uppsala	Stockholm	Uppsala
Stockholm	—	91	—	11	—	65	—	45
Uppsala	91	—	11	—	65	—	45	—
Fennoscandia	86	84	13	85				
Southern Sweden	88	85	−25	83				
(T, P)-model	81	79	71	32				
ξ-model	77	74	58	00	59	60	22	13

between Stockholm and Uppsala and all the other series to provide simple quantitative measurements of the degree of similarity between the series.

Two observations become immediately clear: (i) the low-frequency components in the various series have very different characters before around 1960; (ii) the high-frequency components in all six series have strong similarities in the entire common overlap (1894–1995). Focusing on Stockholm and Uppsala, it is seen that their low-frequency components are quite different in all periods before around 1960, although their high-frequency components are highly similar in the entire 20th century ($r = 0.91$) and late 19th century. The correlation between their high-frequency components decreases before around 1870. This is certainly due to difficulties in translating the descriptive cloud observations made in earlier years to cloud amounts in oktas (see Appendix A).

We expect that the 'true' variations of cloud amounts in Stockholm and Uppsala should exhibit low-frequency correlations with about the same strength as seen in their high-frequency components. The strong dissimilarity between their low-frequency components thus implies that homogenization is necessary. To accomplish this, a homogeneous reference series representative for the Stockholm–Uppsala regional cloud amount variations would be needed. The Fennoscandian and southern Sweden series, which have strong high-frequency correlations with both Stockholm and Uppsala ($r = 0.84$–0.88), are potential candidates for being reference series back to the late 19th century. But, can they be used? Little is known about homogeneity issues in Fennoscandian cloud amount series, in particular regarding their reliability for century-scale trends. Hence, there are reasons to be cautious. In the absence of well-investigated homogeneous cloud series for this region, it would be valuable if an independent estimate of cloud amounts could be developed from other data that are strongly correlated with cloud amounts.

4.2. Cloud amounts predicted from other variables

To identify climatic variables that are strongly correlated with cloud amounts, we calculated correlations between the observed cloud amount and each of the following nine elements; temperature T, precipitation P, air pressure p and the six circulation indices u, v, V, ξ_u, ξ_v and ξ described in Section 3.2. For temperature, precipitation and pressure we used the arithmetic mean of the station data from Stockholm and Uppsala to obtain one time series per variable. The time period analysed is 1951–2000. In this period we consider the Stockholm raw cloud amount series as being homogeneous, as judged from its very strong similarity with both the Fennoscandian and the southern Sweden series in this period. Uppsala is only considered homogeneous in the shorter 1961–95 period. Hence, we use only Stockholm cloud amount data in this analysis. Correlations were calculated for monthly mean values and for the JJA average, and are shown in Table II. To help in visualizing correlation patterns, we print strong correlations ($|r| > 0.7$) in bold and weak correlations ($|r| < 0.4$) in italic. The largest absolute value on each row is underscored.

Cloud amount is found to be strongly correlated ($r \approx 0.7$–0.8) with vorticity (total and zonal component), air pressure and air temperature. The correlation is only slightly weaker with precipitation, but much weaker or totally insignificant with the geostrophic wind components. The pattern is similar in all three summer months and the JJA average. Given the similar pattern in all summer months, we only consider the JJA average in the following analysis. By undertaking stepwise multiple regression experiments, we identified two independent

Table II. Correlation coefficients (×100) between JJA cloud amounts at Stockholm 1951–2000 and six air circulation indices (u, v, V, ξ_u, ξ_v, ξ), precipitation P, sea-level pressure p and temperature T. See main text for explanation of data series. Bold numbers: $|r| > 0.7$; italic numbers: $|r| < 0.4$; underscored numbers: strongest correlation in each row

	u	v	V	ξ_u	ξ_v	ξ	P	p	T
June	*22*	*4*	*26*	69	54	**<u>72</u>**	54	−62	**−71**
July	*7*	*−7*	*23*	**85**	67	**<u>86</u>**	**75**	**−79**	**−79**
August	*12*	*12*	*−12*	**<u>73</u>**	43	**72**	68	**−69**	**−72**
JJA	*20*	*3*	*10*	**79**	63	**<u>81</u>**	**73**	**−80**	**−77**

Table III. Regression statistics for two models of JJA cloud amounts in Stockholm 1951–2000. R^2 is the explained variance; F is the model F-statistic; RMSE is the root-mean-square error (percent cloud cover); p is the probability of obtaining the F-value by chance. The equations give regression coefficients for normalized predictor variables

Model	R^2	F	RMSE	p	Equation
ξ	0.66	93.4	4.2	0.0000	$c = 5.78\xi + 57.6$
(T, P)	0.73	63.7	3.8	0.0000	$c = -3.79T + 3.10P + 57.6$

possibilities to estimate summer mean cloud amounts. One model uses the vorticity ξ index series only, and the other model uses a combination of temperature T and precipitation P as predictor variables. Regression statistics for both models are given in Table III. Both models are highly significant, and in the multiple (T, P)-model both predictor variables are significant at the 0.1% level. The model based on vorticity (ξ-model) explains 66% of the variance in observed cloud amounts, whereas the (T, P)-model explains 73%. The 1951–2000 calibration period is rather short in relation to the entire two-century period of interest, but we assume that the quite strong relationships are sufficiently stable to allow a meaningful extrapolation back in time. Cloud amounts predicted by each of the two models are shown in Figure 2, with the ξ-model in red (1780–2000) and the (T, P)-model in green (1894–2000). Their correlation coefficients with Stockholm and Uppsala are shown in Table I. The first year in the (T, P)-model is determined by the first reliable year in the Stockholm precipitation record.

Both cloud estimates exhibit a strong high-frequency correlation with Stockholm and Uppsala in the period 1900–95 ($r = 0.74$–0.81), which is only slightly less than the correlation between Stockholm/Uppsala and the two regional series. As regards the low-frequency components, the two regression model series agree remarkably well with each other, but they disagree strongly with both Uppsala and the two regional series concerning the overall 20th century trends. These three latter series all suggest increasing trends, whereas the two models predict slightly decreasing trends. The low-frequency component in the Stockholm series is quite different compared with the regression models and with the other observational series in the first half of the 20th century.

Given the similar behaviour of the ξ-model and the (T, P)-model over the entire 20th century, their high degree of cloud amount variance explained in the 1951–2000 period, and also their strong high-frequency correlation with Stockholm and Uppsala in the entire 20th century, it seems likely that both models predict realistic long-term trends over the 20th century. As these trends clearly disagree with those in the two regional series, it is reasonable to conclude that Fennoscandian cloud amount series may have systematic negative biases before around 1950 in the JJA season (particularly the southern Sweden series, which shows markedly low values around 1900).

There are several reasons why observational cloud amount series (in general) may be inhomogeneous; e.g. inadequately trained observers, changes of observation times, changes of observing and reporting codes and the influence of growing cities (Henderson-Sellers, 1992). Using examples from Finnish stations, Heino (1994) argued that the subjective nature of ground-based cloud observations is a severe cause of inhomogeneities in long cloud amount series. He considered only Finnish data from airport stations after the 1940s as reliable, because these observations had been conducted by adequately trained personnel.

We conclude that it is *not* possible to test and homogenize the Stockholm and Uppsala raw cloud amount series, not even in the 20th century, using other observer-based cloud series from this region as reference data. The last resort here, for finding a reasonably homogeneous reference series, is to choose one of the two regression models. However, as our primary aim is to test summer *temperature* data, we cannot use the (T, P)-model because this would lead to circular reasoning. Hence, only the ξ-model remains useful. Consequently, we decided to employ this series in homogeneity tests of the raw cloud series. Furthermore, only the ξ-model can be calculated for all years back to 1780.

4.3. Adjustment of the raw cloud amount series

To search for homogeneity breakpoints in the raw Stockholm and Uppsala JJA average cloud amount series, the standard normal homogeneity test for break-points (Alexandersson, 1986) was applied to the series of differences between the raw cloud amounts and the cloud amounts predicted by the ξ-model. It is important to bear in mind (here and in the rest of this study) that the degree of homogeneity in the ξ-model is determined both by the quality of the gridded pressure data used to derive the vorticity index and by the extent that the regression relationships derived in the 1951–2000 period hold for the entire period back to 1780. It is difficult to quantify the uncertainties associated with these factors, but we argue that homogenization based on the ξ-model as reference is better than no homogenization at all — and no homogenization would definitely mean that the cloud series have to be rejected from the rest of this study.

The test was applied for several subperiods to ensure that each test period contained, at most, one significant break. (The test was also applied to the difference between the raw Stockholm and Uppsala series and the southern Sweden series to obtain further information of possible breakpoints after 1873.) Breakpoints indicated by the tests were then compared with possible breakpoints suggested by metadata (e.g. change of cloud observation method, change of observation hours, change of observer). If metadata suggested a breakpoint just a few years from a break identified by the test, then we chose the date indicated by the metadata.

Once the breakpoints had been determined, the average difference (raw series minus reference series) in each homogeneous subperiod was compared with the corresponding difference in the most recent homogeneous subperiod (1951–2000 for Stockholm, 1961–1995 for Uppsala). If the difference in an early subperiod differed from that in the modern period at the 5% significance level according to an ordinary *t*-test, then we adjusted the raw values in the early subperiod to force the difference (adjusted series minus reference series) to be the same as in the modern period. Because there was a change in cloud observation method in Stockholm in June 1784 (see Appendix A), i.e. only 4 years after the start of the reference series, we preferred to determine adjustments for Stockholm before 1784 by comparing the raw values for 1756–83 directly with the adjusted Stockholm values for 1784–2000. The adjustment terms in different periods are given in Tables IV (Stockholm) and V (Uppsala). Approximately 40% of the Stockholm data remained unadjusted before 1950, whereas all data before 1949 were adjusted in the Uppsala series. Substantial positive adjustments (up to about 9% units) were necessary in some periods at both sites. A few periods with negative adjustments were also found.

4.4. The adjusted cloud amount series

After having applied the adjustments, we averaged the Stockholm and Uppsala series. The resulting adjusted and averaged JJA cloud series is shown (black curves) in Figure 3 together with the ξ-model time series

Table IV. Constants (in percent cloud cover) added to the raw Stockholm JJA cloud amount series to produce the adjusted cloud amount series

Period	Adjustment
1756–1783	+4.4
1784–1809	0
1810–1815	+8.8
1816–1821	0
1822–1840	+7.5
1841–1847	−4.3
1848–1872	0
1873–1905	+6.4
1906–1925	−3.4
1926–1938	0
1939–1949	+7.9
1950–2000	0

Table V. Constants (in percent cloud cover) added to the raw Uppsala JJA cloud amount series to produce the adjusted cloud amount series

Period	Adjustment
1780–1831	+5.3
1832–1845	+9.1
1846–1854	−6.9
1855–1926	+9.1
1927–1948	+4.1
1949–1995	0

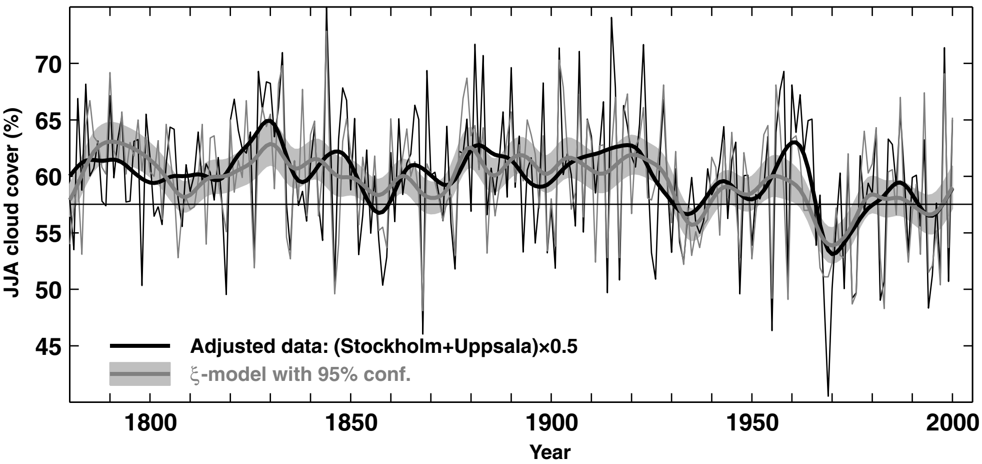

Figure 3. Arithmetic average of adjusted JJA mean cloud amounts (percent cloud cover) for Stockholm and Uppsala 1780–2000 (black) and cloud amounts predicted by the ξ-model (grey; equivalent to red series in Figure 2). The ξ-model is shown with smoothed 95% confidence intervals (light grey) for the conditional mean value calculated at each time point. Smoothing as in Figure 1. Horizontal line shows the 1961–90 average for adjusted observational data

(grey). To illustrate the uncertainty of the regression coefficients in the ξ-model, we calculated for each year in the time series a 95% confidence interval for the conditional mean of the predicted cloud amount (using equation 8.23 in von Storch and Zwiers (1999)). The upper and lower limits were smoothed with the same Gaussian filter as in Figure 2 and elsewhere in this paper, and the interval between is plotted as a light grey band centred on the smoothed ξ-model series. The smoothed adjusted cloud amount series nearly always lies within this interval, which is to be expected as the 5% significance level was chosen as the critical level for deciding whether or not to adjust the raw cloud series.

The main effect of the generally positive adjustments applied (as can indirectly be seen by comparing Figure 2(a) with Figure 3), is to raise the overall level of cloud amount before around 1920 (which were mostly below the 1961–90 average in the raw series) to be generally above the 1961–90 average in the adjusted series. This increase in cloud amount is forced by a corresponding high level in the vorticity index series before around 1920. Even the lower limit of the smoothed confidence interval is nearly always above the 1961–90 average. This latter observation has important implications: with both vorticity and cloud amount being at relatively high levels before 1860, we have obtained a first indication of non-support for the warm observed early summers, as this is in qualitative disagreement with warm summers.

We emphasize, however, that the low-frequency variability in the adjusted summer cloud amount series is almost entirely determined by the low-frequency variability in the vorticity index data used as reference in the homogenisation. Hence, the adjusted cloud amount series cannot be considered as useful on its own for assessing the veracity of the warm summer temperatures before the 1860s. Nevertheless, we decided to continue to include the adjusted cloud series in our study because the high-frequency component is thought to provide relevant information on the strength of relationships between various variables and temperatures on year-to-year time scales.

See also Appendix B, regarding analyses of data for the other seasons.

5. DEVELOPMENT OF A REGRESSION MODEL FOR ESTIMATING SUMMER TEMPERATURES

Here, we investigate statistical relationships between observed summer temperatures T in Stockholm and Uppsala and the six circulation indices u, v, V, ξ_u, ξ_v and ξ, local precipitation totals P, air pressure p and cloud amounts c. For the T, P, p and c variables, we use arithmetic averages of Stockholm and Uppsala data. The adjusted cloud amount data from Section 4.4 are used. As in Section 4, we start by analysing the pattern of correlation coefficients and continue with stepwise multiple regression. Analyses were made separately for June, July and August monthly means and for the JJA average. To obtain a view of relationships in all parts of the year, we also analysed the other 9 months, but results for non-summer data are only mentioned very briefly. At the end of this section, we present a regression model for estimating JJA temperatures for 1780–2000, provided with error bars, which we compare with the observed temperature series.

5.1. Correlations

Correlations between temperature and the other variables are listed in Table VI for the two periods 1873–2000 and 1780–1872. The 1873–2000 period is later used as calibration period in the regression analysis. This period is chosen because the year 1873 marks a transition in the construction of the gridded air pressure dataset used to calculate the circulation indices (see Section 3.2), and also because 1873 is the first year when the cloud amount series does not involve any translation of old observation styles (see Appendix A). To visualize correlation patterns, all numbers where $|r| < 0.3$ are printed in italic and correlations where $|r| > 0.50$ are printed in bold. The highest value in each row is underlined.

All three summer months display similar patterns of correlations in the 1873–2000 period. The correlation is strongest with cloud amount ($r \approx -0.7$). Slightly weaker correlations are found with air pressure ($r \approx 0.6$) and vorticity ($r \approx -0.6$; total and zonal component). Correlation with precipitation is weaker still ($r \approx -0.4$ to -0.5), and correlations with the geostrophic wind components are always below 0.4. Correlations for JJA averages tend to be slightly stronger than for the monthly data, indicating that the averaging reduces some random variability in the monthly data. Looking at the early period of 1780–1872, the pattern is qualitatively

Table VI. Correlation coefficients ($\times 100$) between averaged Stockholm and Uppsala summer temperatures and six air circulation indices ($u, v, V, \xi_u, \xi_v, \xi$), precipitation P, sea-level pressure p and adjusted cloud amounts c for two time periods. Bold numbers: $|r| > 0.50$; italic numbers: $|r| < 0.30$; underscored numbers: strongest correlation in each row

Period		u	v	V	ξ_u	ξ_v	ξ	P	p	c
1873–2000	June	*−18*	33	−35	**−58**	−44	**−62**	−36	**58**	**−71**
	July	*−25*	*29*	−40	**−58**	−45	**−60**	**−51**	**60**	**−71**
	August	*−24*	*18*	*−6*	**−64**	−35	**−61**	−39	**59**	**−69**
	JJA	−30	*27*	−32	**−65**	−49	**−67**	−44	**67**	**−74**
1780–1872	June	*0*	32	*−21*	−36	*−28*	−40	*−1*	35	**−57**
	July	*−16*	*26*	*−26*	−42	−38	−46	*−26*	**60**	**−69**
	August	*−2*	44	*1*	**−51**	−42	**−54**	−32	**60**	**−59**
	JJA	*−14*	30	*−24*	**−51**	−31	**−51**	*−15*	**61**	**−62**

similar, but the correlation strengths are generally weaker. We interpret this as an indication of generally less reliable data quality in the early period (affecting all variables analysed). Yet, the absolute correlations with both cloud amounts and air pressure are still as high as about 0.6 for JJA averages.

A qualitative physical interpretation of the correlations is rather obvious. An increase in cloud amount leads to a reduction of incoming solar radiation during summer days, which in turn leads to cooler summer temperatures. Stronger vorticity and lower air pressure means stronger cyclonic circulation, which is associated with more clouds and precipitation and hence cooler summers. The rather weak correlation with zonal wind is negative, whereas the correlation with meridional wind is positive. Hence, increased westerly wind flow in summer tends to bring cooler maritime air from the Atlantic, whereas increased southerly flow brings more warm continental air masses over southeastern Sweden. A windy summer is likely to be relatively cool, and this is reflected in the negative temperature correlations with wind speed.

A comparison with correlations for the other months (not shown) reveals that the data for May behave much like the JJA data, but there is a clear distinction between the May–August period from the rest of the year. For all other months, the strongest correlations are found with one of the geostrophic wind indices, whereas mostly weak correlations are found with cloud amount, air pressure and vorticity. Reduced correlations in the early period are also seen in the non-summer months.

5.2. *Stepwise multiple regression*

We applied forward stepwise multiple regression, attempting to identify useful linear models for estimating averaged Stockholm/Uppsala temperatures, with predictor variables chosen among the elements investigated in Section 5.1. The period 1873–2000 is used to develop the relationships. As in the correlation study above, analyses were made for the individual months of June, July and August and for the JJA average. In each case, we took the cloud amount as the first input variable to the model, because this measure displayed the strongest correlation with temperatures. Further variables were incorporated until there was no more improvement in significant regression parameters (i.e. until no further regression parameters significant at the 5% level could be found). The regression statistics are given in the last three columns in Table VII.

Table VII. Stepwise multiple regression statistics for models of averaged Stockholm and Uppsala summer temperatures 1873–2000. R^2, F, RMSE and p are defined in Table III. #1 to #3 denote the order of appearance of the predictor variables in the model. Statistics for the model applied to 1780–1872 data are given in the last three columns

	1873–2000							1780–1872		
	R^2	F	RMSE	p	#1	#2	#3	R^2	Bias	RMSE
June	0.64	74.0	0.94	0.0000	c	v	p	0.40	−0.66	1.34
July	0.63	68.9	0.93	0.0000	c	v	u	0.58	−0.67	1.38
August	0.58	55.9	0.98	0.0000	c	ξ_u	v	0.49	−0.45	1.31
JJA	0.65	75.5	0.66	0.0000	c	v	p	0.52	−0.53	1.04

Table VIII. Multiple regression models for averaged Stockholm and Uppsala summer temperatures obtained from the period 1873–2000 (same as in Table VII). Regression coefficients are given for normalized predictor variables, listed in order of decreasing absolute values from left to right

June	$T =$	$-0.89c$	$+0.49v$	$+0.34p$	$+13.9$
July	$T =$	$-1.03c$	$+0.44v$	$-0.32u$	$+16.4$
August	$T =$	$-0.71c$	$-0.47\xi_u$	$+0.34v$	$+15.3$
JJA	$T =$	$-0.58c$	$+0.30v$	$+0.30p$	$+15.2$

For each of the three monthly series, we find that three predictor variables together explain about 60% of the temperature variance. In addition to cloud amounts, all models include meridional geostrophic wind, whereas the third variable differs. For the JJA average, the model includes cloud amount, meridional wind and local air pressure, together explaining 65% of the temperature variance. All three monthly models, as well as the seasonal one, are highly significant. The relative importance of the predictor variables in each model can be seen in Table VIII, where regression coefficients are given for normalized predictor variables. The coefficients thus show how much a change of one standard deviation in each variable affects the temperature (in centigrade), provided that the other variables are held constant. Cloud amount is found to be the most important variable in all four models. For the JJA average, the coefficient for cloud amount is about twice those for meridional wind and air pressure. All four regression models were then applied to data in the 1780–1872 period. The explained variance, bias and root-mean-square error are given in the three rightmost columns in Table VII. As expected, the explained variances are lower, but still as high as 52% for the JJA average. The bias is negative, about $-0.5\,^{\circ}$C, indicating systematically $0.5\,^{\circ}$C cooler summer temperatures predicted by the model compared with those observed.

5.3. *JJA average temperatures predicted from a multiple regression model*

The similarity of relationships with other variables in all three summer months suggests that JJA mean temperatures can be meaningfully estimated directly by using the regression model obtained for the JJA average. Application of this model to the entire predictor time series gives the sequence of predicted temperatures for 1780–2000 plotted in Figure 4(a). Annual values are shown with bars and low-frequency variability is shown with the black curve. The statistical uncertainty in the determination of the regression coefficients is illustrated with the dark grey band centred on the black curve. This band is defined as the interval between the smoothed sequences of the upper and lower limits of a 95% confidence interval for the conditional mean of the response variable (equation 8.33 in von Storch and Zwiers (1999)), calculated at each point in the time series.

There is also an additional uncertainty due to the fact that the cloud amount series used in the model was adjusted using another regression model (the ξ-model in Section 4.2) as reference. This latter uncertainty cannot be easily calculated because the Stockholm and Uppsala cloud series were adjusted separately, with a number of piece-wise adjustments in different periods (see Section 4.3). However, as a rough estimate of this uncertainty, we use the maximum width of the smoothed 95% confidence interval for the ξ-model in Figure 3, which nearly always envelops the smoothed adjusted cloud series. This interval has a maximum half-width of about 1.9%-units cloud cover. Taking this as an approximation of the uncertainty in the mean value of the adjusted cloud series, we repeated the calculation of a 95% confidence interval for the predicted temperatures, after having added 1.9%-units to the adjusted cloud series *in the pre-calibration 1780–1872 period*. The same calculation was then repeated after subtracting 1.9%-units. Together, these calculations give a wider range of uncertainty in predicted 1780–1872 temperatures, which accounts also for the uncertainty in the adjusted cloud series. This extended confidence interval is illustrated with the light-grey band surrounding the dark-grey band in Figure 4(a).

5.4. *Comparison of observed and predicted summer temperatures*

For an easy direct comparison between the modelled and the observed JJA temperatures, we show in Figure 4(b) the average of the observed Stockholm and Uppsala temperatures (same as in Figure 1(a)) together with the smoothed confidence intervals for temperatures predicted by the regression model (same as in Figure 4(a)). The behaviour of the smoothed predicted JJA temperatures has clear similarities with the smoothed observed temperatures in the calibration period 1873–2000. Both series have values generally below the 1961–90 average during 1880–1930, followed by a rather marked decadal variability with warm peaks in the 1930s, around 1970 and in the 1990s. The confidence interval for the predicted temperatures mostly envelops (or nearly so) the smoothed observed temperatures in the calibration period, except in the 1880s (the cold trough is too weak) and the 1930s and 1990s (warm peaks are underestimated). Apart from these underestimated decadal peaks and troughs, the model reproduces the observed temperature variability

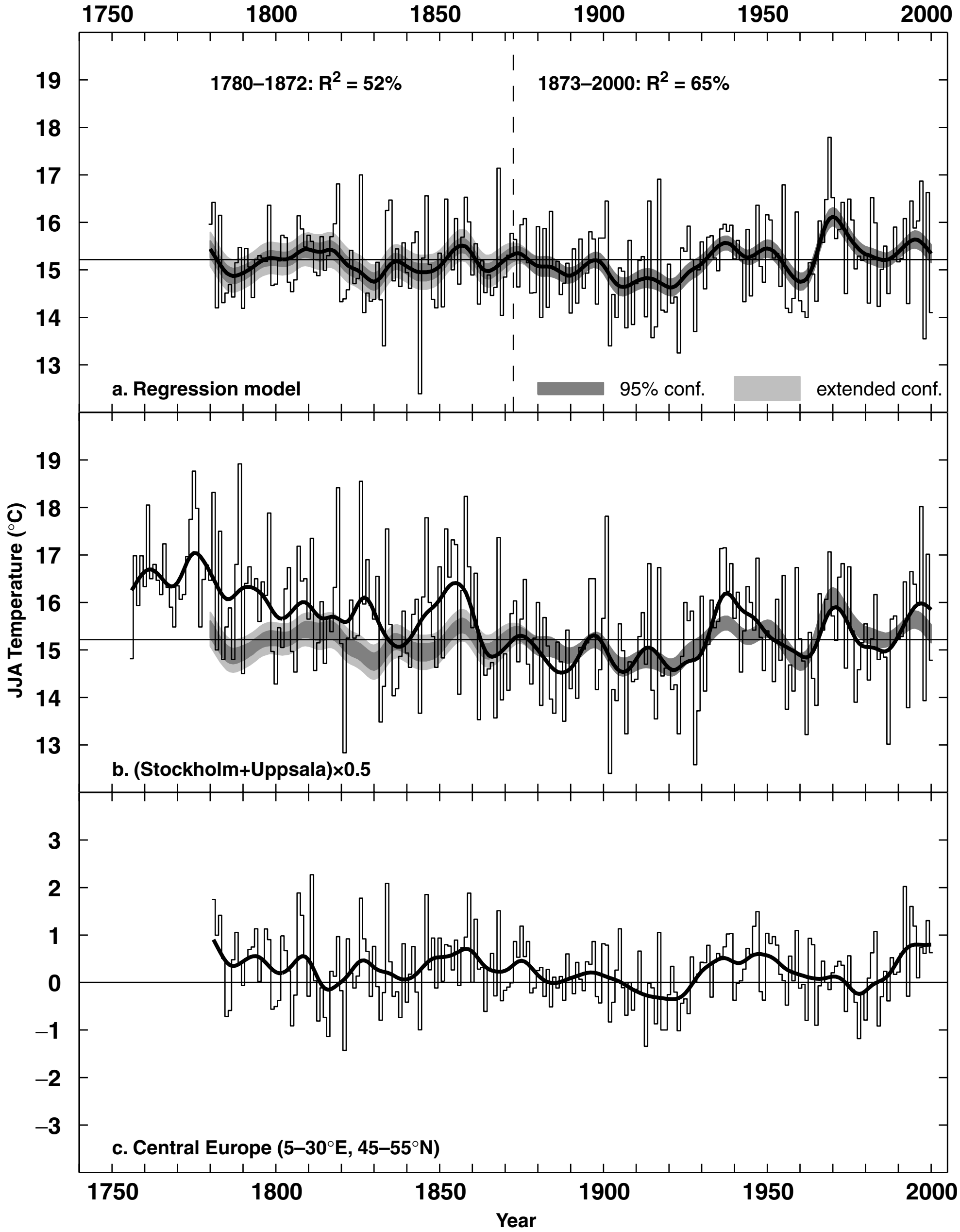

Figure 4. Observed and estimated JJA mean temperatures. (a) Average of Stockholm and Uppsala temperatures 1780–2000 predicted from a multiple linear regression model (predictor variables: cloud amounts, meridional geostrophic wind, sea-level pressure). The series is shown with smoothed 95% confidence intervals (dark grey) for the conditional mean value calculated at each time point. Additional uncertainty due to uncertainty in the cloud amount series is illustrated with the wider (light grey) interval before 1873. (b) Average of observed Stockholm and Uppsala temperatures 1756–2000 (same as in Figure 1(a)), plotted together with the confidence intervals for the regression model in (a). Horizontal lines in both (a) and (b) show the 1961–90 average for the observed temperatures. (c) Central European temperatures expressed as anomalies from the 1961–90 mean (from Jones *et al*. (in press)). Smoothing as in Figure 1

rather well, both concerning the high-frequency and the low-frequency components. This is further illustrated by a strong, and nearly equal, correlation both in the high-frequency ($r = 0.80$) and low-frequency ($r = 0.82$) components.

Looking at the early period 1780–1872, however, there is a clear discrepancy between the low-frequency components in predicted and observed temperatures ($r = 0.14$), despite the fact that the high-frequency correlation ($r = 0.81$) is as strong as in the calibration period. In particular, the extended (light grey) confidence interval envelops the 1961–90 average level virtually all years in the early period, whereas the smoothed observed series lies above its upper bound most of the time before around 1860. Hence, taking into account the range of uncertainties both in the determination of regression coefficients in the temperature model and also uncertainties in the development of cloud amount adjustments, we find no support from this analysis for the warm summer temperatures observed before around 1860.

6. DISCUSSION

We have investigated whether it is possible to find a combination of climate variables that are both strongly correlated with temperatures at Stockholm and Uppsala, and at the same time have homogeneous time series extending back before 1860. We found in Section 5 that it is indeed possible to find a combination of different climate variables that together explain a large fraction of variance in summer temperatures. Local cloud amount, meridional geostrophic wind and local air pressure together explain 65% of the JJA temperature variance in the period 1873–2000, with nearly identical skill for both high-frequency and low-frequency variability. An interpretation of this linear multiple regression model developed is intuitively very logical: cloudy summers are generally cool, southerly winds bring warm air from the continent and high air pressures are associated with 'good weather' and warm summers.

The relative importance of the three predictor variables in the model may, of course, have changed with time, as time-varying strengths of relationships between atmospheric circulation and temperatures in Europe have previously been observed to occur in the period after the late 18th century (Jacobeit *et al.*, 2001; Slonosky *et al.*, 2001; Jones *et al.*, 2003). The *sign* of the relationships in our model, however, could hardly have reversed. The fact that the high-frequency correlation between the observed summer temperatures and those predicted by the regression model in the pre-calibration period 1780–1872 have the same high values ($r \approx 0.8$) as in the calibration period actually suggests that the relationships have, on average, remained remarkably similar. Our regression model should, therefore, *in principle*, be able to identify periods that were systematically unusually cold or warm.

So, from a climatic viewpoint, we find it possible to derive a useful estimate of summer temperatures. The main obstacle, however, is the quality of the time series used as predictors. There is little similarity between the low-frequency components of observed and estimated temperatures before the mid-19th century, despite a strong correlation between their high-frequency components in the same period. This is unfortunate for our purposes, as reliable low-frequency components in the estimated temperatures would be necessary for an accurate answer to the question of possibly biased observed early summer temperatures.

6.1. More considerations on the temperature series reconstructed from cloud amounts, meridional geostrophic wind and local air pressures

Cloud amount is found to be the most important variable in our model (Section 5.2). Unfortunately, this is also the most problematic variable from a data quality viewpoint. We developed cloud amount series from the original observations (in Appendix A), and found (Section 4) that these were highly inhomogeneous, and hence had to be adjusted. To derive adjustments we used an estimate of cloud amount based on a vorticity index series as reference. This procedure implies that the adjusted cloud amount series has its low-frequency variability determined by the gridded sea-level pressure dataset (Jones *et al.*, 1999) from which we derived the vorticity index series.

The second variable in our model, the meridional geostrophic wind component, is directly derived from the Jones *et al.* (1999) dataset. In addition, even the third predictor variable, the averaged Stockholm and Uppsala sea-level pressure series, is also dependent on the same gridded air pressure data. This is because the Stockholm and Uppsala pressure series have been homogenized (by Moberg *et al.* (2002) and Bergström and Moberg (2002)) using data from the gridded pressure dataset as reference. Recall also (from Section 3.2) that this gridded air pressure dataset was developed (by Jones *et al.* (1999)) using regression techniques, with the UKMO hemispheric MSLP dataset as predictand. It follows from the above that our summer temperature reconstruction is heavily dependent on regression techniques applied at several stages, not only here but also by Jones *et al.* (1999). Furthermore, uncertainties in the gridded pressure dataset are not only caused by the regression used for its development, but also relate to the quality of the individual station pressure series from which the gridded data were constructed. It is clear from this discussion that, even if we have used three different climate variables as predictors in the regression model, the low-frequency behaviour of all three depends on the same gridded air pressure data, and only their high-frequency components are entirely independent (in terms of how they were originally observed).

It is not an easy task to quantify confidence intervals that account for all the uncertainties that affect our summer temperature model. It is obvious, though, that an interval covering this total uncertainty must be wider than the extended confidence interval shown in Figure 4(a) and (b). The average half-width of the latter is 0.32 °C, so a reasonable guess is that the true uncertainty in the mean level is at least about ±0.5 °C. This is approximately the same amount as the average difference between the observed and predicted temperatures before the 1860s. Hence, the regression model cannot be used for an accurate quantification of biases in the observed temperatures, but we can still draw a qualitative conclusion: there is no support from the data used in our temperature model for an overall level of summer temperatures before 1860 being above the 1961–90 average, as suggested by the observed temperatures. Hence, the latter are most likely positively biased.

To quantify this bias, it seems to us that the simplest way to go would be to develop a homogeneous reference series based on temperature series from other stations with records starting well before the 1860s. Such stations would have to be located relatively near Stockholm and Uppsala and have temperatures that are strongly correlated with these sites. There are no Swedish series that fit these requirements, so it remains only to use data from other countries in Europe.

6.2. Comparison with temperature series for central Europe

There exist other temperature series from northern Europe (e.g. Trondheim, Tallinn, Vilnius, Copenhagen, St Petersburg) that start in the 18th century. Unfortunately, only the St Petersburg series has been thoroughly investigated (Jones and Lister, 2002), and this series may have an urban warming trend affecting data before 1835, so the long northern European series cannot be used without first undertaking homogeneity tests. Such an exercise is beyond the scope of this paper. However, some relevant information can be obtained from the rather numerous central European temperature series. Some of them, in particular the Austrian ones, have undergone extensive homogenization recently (Böhm *et al.*, 2001).

Jones *et al.* (in press) constructed a temperature series for central Europe for the period 1781–2000, which includes the homogenized Austrian series and several others from neighbouring countries. This central European series is an average of grid-box temperature anomaly series within the region 5–30 °E and 45–55 °N, and expresses temperatures as departures from the 1961–90 average. The number of stations involved varies with time. In total, 34 stations have data starting before 1851, of which 18 have data extending back to 1781. The large number of stations in the average should effectively cancel any random inhomogeneities in individual station series.

The central European JJA temperature series is shown in Figure 4(c). Before around 1880 the smoothed series is generally above the 1961–90 normal. The central European temperatures thus actually support the Stockholm and Uppsala observations of summer temperatures before 1860 being marginally above the 1961–90 level. However, the early central European temperatures appear less markedly warm compared with the Stockholm and Uppsala observations. The relatively large distance between the two regions makes it difficult to interpret the discrepancies between the series, and we consider this direct comparison as rather inconclusive.

6.3. Homogeneity tests of the observed summer temperatures

It has already been stated that we cannot use our regression model for an accurate quantification of a warm bias in the early Stockholm and Uppsala data, mainly because of uncertainties that are too large in the

Table IX. Results of standard normal homogeneity tests (Alexandersson, 1986) applied to JJA mean temperatures for Stockholm and Uppsala. T_{max} is the test statistic, where a T_{max} of 11.1 is the critical 97.5% level for a record length of $n = 220$

Series tested	Reference series	Test period	Break year	Change (°C)	T_{max}
Stockholm	Regression model	1780–2000	1861	−0.77	42.0
Uppsala			1861	−0.72	40.6
Stockholm	Central Europe	1781–2000	1858	−0.60	18.6
Uppsala			1861	−0.53	16.5

predictor variables. Likewise, we consider the central European temperature series to represent a too distant region. However, even if the *size* of a bias cannot be determined, a statistical test may anyway provide a reliable estimate of the *timing* of a homogeneity break in the observed temperature series. Therefore, we apply the standard normal homogeneity test for abrupt changes (Alexandersson, 1986) to the observed JJA mean temperatures, separately for Stockholm and Uppsala. The test was applied twice to each series, first with the regression model as reference series and then with the central European temperatures as reference. The test results are given in Table IX.

All tests reveal a highly significant break, with a jump from higher to lower values close to 1860. For Stockholm, the break is determined to occur between 1858 and 1859 with the central European temperatures as reference and between 1861 and 1862 with the regression model data as reference. The earlier of these indicated breaks coincides exactly with the time point when the observations in Stockholm began to follow the 'new' national standard routines (see Section 2.1). In particular, this included a change of hour for the morning observation, so that the wall where the thermometer was placed could no longer be heated by the sun at the time of observation, which it would likely have been before 1859. The combination of evidence from the test and metadata thus strongly suggest that Stockholm summer temperatures in the period 1780–1858 are positively biased. For Uppsala, the indicated break is between 1861 and 1862 in both tests, which does not correspond exactly to any particular known change of observation routines. Rather, this is about in the middle of a sequence of years from 1853 to 1874 when several changes occurred, including a relocation from an urban to a rural site and introduction of radiation screens, which both could have induced artificial cooling of the observations.

The estimated size of the bias (although uncertain as discussed above) is found to be nearly the same for Stockholm and Uppsala, but slightly larger when the regression model data are used as reference series (∼0.7–0.8 °C) compared with when the central European temperatures are used (∼0.5–0.6 °C). It is not surprising that Stockholm and Uppsala have the same indicated biases, as the Uppsala data before 1854 were actually adjusted by Bergström and Moberg (2002) to agree with Stockholm (see Section 2.1). However, this adjustment is only about −0.1 °C for the JJA average and, therefore, any bias in the unadjusted Uppsala summer temperatures is only 0.1 °C larger than in the adjusted series analysed here.

We emphasize that our analyses only provide an *approximate* estimate of *average* biases in summer temperatures between 1780 and the years of 1858 (Stockholm) and 1861 (Uppsala). The 'true' biases may have been different in June, July and August and could also affect May, which has not been analysed here. Furthermore, biases may differ between various subperiods, and may even be absent in some periods. All such details may also differ between Stockholm and Uppsala. In addition, nothing can be concluded from this study concerning possible biases in data before 1780, when the observed temperatures are at an even higher level than during 1780–1860. Only further homogeneity tests (with much more reliable reference data) can provide more accurate information.

The fact that tests with central European temperature series as reference data indicate about 0.15 to 0.2 °C smaller biases at Stockholm and Uppsala, compared with tests with the regression model data as reference, may be interpreted as an indication that some of the early station records in the central European average could also be positively biased. However, the difference in test results is within the range of uncertainty in temperatures predicted by the regression model and, furthermore, century-scale summer temperature trends in southern Sweden and central Europe may have been slightly different. More research would be needed to assess the extent to which positive biases in early temperature records due to inefficient radiation screening is a widespread problem. In this context, we recommend digitizing the original observations for other European sites back to the start of their records, both of temperature observations and also other variables (such as cloudiness, precipitation and pressure).

Finally, we mention that there is definitely *no evidence* from this study that summers in Stockholm and Uppsala before 1860 could have been on average *cooler* than the 1961–90 mean, as suggested by the Austlandet proxy series for southeastern Norway (Nordli, 2001a; see Section 2.2). Rather, it would be useful to put further efforts into investigating whether the cool reconstructed temperatures reflect real climate conditions in southeastern Norway, or whether the proxy data part of the Austlandet series is *negatively* biased.

7. SUMMARY AND CONCLUSIONS

Temperature series for Stockholm and Uppsala (Bergström and Moberg, 2002; Moberg *et al*., 2002) suggest that summers from the mid-18th century until 1860 were generally relatively warm, with decadal means nearly always above the 1961–90 average and with some decadal peaks above the warmest periods in the 20th century. We question whether the warm summers indicated by the observational data are real or whether the data are positively biased. Station metadata suggest that the latter cannot be excluded. There are several possible reasons for this: poor protection against radiation (at both sites), morning sunshine heating the wall where the thermometer was placed (hence biasing morning observations, in the case of Stockholm), and urban bias before a relocation to a site outside the old town (in the case of Uppsala).

Our approach to the study of this problem has been to analyse statistical relationships between summer temperatures and several other climate variables, and to investigate whether it is possible to develop an estimate of summer (JJA) temperatures that can be directly compared with the observed temperatures. As cloud amount was initially considered to be a potentially useful variable, and because no long cloud series was available prior to this study, we began with developing cloud amount series from the original observations at Stockholm and Uppsala. The cloud amount series from the two sites were found to have very different low-frequency behaviour (even in the 20th century), despite a strong similarity of their high-frequency components. Hence, homogenization was necessary. To develop homogeneous cloud series back to 1780, we observed that a vorticity index series derived from gridded pressure data could be used as reference. A drawback of a homogenization based on this series is that the low-frequency behaviour of the final adjusted cloud amount series is dependent on the gridded air pressure data used to derive the vorticity series.

We find, in principle, that it is possible to estimate summer temperatures with a combination of local cloud amount, the meridional geostrophic wind component and local air pressure as predictor variables in a multiple linear regression model, where cloud amount is the most important. Such a model explains 65% of the temperature variance in summers 1873–2000. However, owing to homogeneity problems, in particular for cloud amount, the low-frequency component of estimated temperatures before the mid-19th century is poorly correlated with the observed temperatures, despite a strong high-frequency correlation. Hence, the regression model has limited use for a quantitative estimation of a possible bias in the early summer temperature observations.

Taking account of uncertainties in the determination of regression coefficients in the temperature with a 95% confidence interval, and also uncertainties in the development of adjustments applied to the cloud series, we find that decadal means of observed summer temperatures in Stockholm and Uppsala before the 1860s are nearly always above the upper limit of the uncertainties. Hence, the data used in the regression model give no support for the warm observed temperatures, but instead strongly suggest that these are positively biased. Additional comparisons were also made with a regional temperature series for central Europe. These data indicate that summers before 1860 were often moderately warmer than the 1961–90 average, and hence partly support the warm early summers observed in Stockholm and Uppsala. The warmth indicated by the central European data before 1860, however, is less pronounced than at Stockholm and Uppsala.

Neither of the two reference series discussed, the regression model series nor the central European temperature series, are considered accurate enough for quantifying a bias in early summer temperatures with good precision, but they are considered useful for testing the *timing* of a homogeneity break. Separate tests of Stockholm and Uppsala temperatures indicate significant breaks near 1860 (either between 1858 and 1859 or between 1861 and 1862). For Stockholm, this corresponds very closely to the year 1859, when important changes were made in the observational procedures. This result is regarded as strong evidence that summer temperatures before 1859 are positively biased. For Uppsala, the indicated year of break does not correspond to any particular year suggested by metadata, but is rather in the middle of a two-decade long period with several changes of observational procedures. The size of the estimated bias before the breaks varies between about 0.5 and 0.8 °C, depending more on the reference series being used than on the station tested. A more accurate quantification cannot be obtained from this study. In summary, the main conclusions drawn from this study are:

1. It is very likely that the summer temperature series for both Stockholm and Uppsala developed by Moberg *et al.* (2002) and Bergström and Moberg (2002) are positively biased before around 1860. For Stockholm, the last year with biased temperatures is concluded to be 1858, whereas for Uppsala it is not possible to determine an exact year.
2. The sizes of the bias cannot be accurately determined from the evidence used here, but they are probably about 0.5–0.8 °C on average for the period from 1780 to the years near 1860 when the biased periods end.

Further efforts would be needed to investigate the problem more closely and to derive more accurate estimation of summer temperature biases. This could probably best be achieved if a homogeneous reference series can be developed from northern European stations with long temperature records, which would first require extensive homogeneity testing of these series, and also an assessment of how widespread the problem is with inefficient radiation screening in early temperature series. We recommend such a study be undertaken, not only for the purpose of preparing for further assessment of Stockholm and Uppsala temperatures, but because this would be necessary for more reliable reconstruction of climate change in the whole of Europe back to the 18th century.

APPENDIX A: DEVELOPMENT OF MONTHLY CLOUD AMOUNT SERIES FOR STOCKHOLM AND UPPSALA

Here, we provide details of how observations of clouds have been reported in different time periods and we describe how the 'raw' monthly cloud amount series were developed for Stockholm 1756–2001 and Uppsala 1780–1995.

A.1. Stockholm

Pehr Wargentin made the meteorological observations in Stockholm from 1754 to 1783. He described with a sequence of words, abbreviations or symbols, the evolution of 'the air's condition' during each day (including night) divided in four parts. The most common terms for daytime observations were 'kl.' (= *klart*, clear), 'str.' (= *strömoln*, scattered clouds) and 'm.' (= *mulet*, cloudy). For example, the sequence 'kl.kl.m.' in the column 'from dawn to noon' probably indicates that the sky was clear for most of the day, but changed to cloudy during its later part. We used the first symbol in each of the three columns 'dawn to noon', 'noon to dusk' and 'dusk to midnight'. The symbols selected thus represent the cloud conditions at approximately the times when temperature observations were made (see Moberg *et al.* (2002)).

The observation routines changed on 2 June 1784, when thrice-daily cloud observations began to be made at fixed hours (Moberg *et al.*, 2002). Until 31 December 1858, the cloud amount, or rather the appearance of the sky, was denoted with various symbols. There seems to have been three different sets of symbols (June 1784–December 1815, January 1816–June 1841 and July 1841–December 1858), although there are some ambiguities concerning the exact time when symbols changed. In particular, the third set seems to have been introduced gradually during a period of several years rather than abruptly. While coding the symbols, we had to set exact dates for practical purposes. The meaning of the symbols (see Table X) are deduced

Table X. Relative frequencies and cumulative relative frequencies for each class of cloud observations in Stockholm in various periods. For data before 1873, the approximate range of oktas for each class, and the corresponding average in tenths, are indicated

Period	Class		AMJJAS				ONDJFM			
	Swedish term	English translation	Frequency	Cum. freq.	Oktas	Tenths	Frequency	Cum. freq.	Oktas	Tenths
Jan 1756–May 1784	Klart	clear	0.40	0.40	0–4	2.65	0.27	0.27	0–3	1.82
	Strömoln, dunkelt, töknigt, rök	scattered clouds, gloomy, misty, haze	0.27	0.67	5–6	7.05	0.14	0.42	4–6	6.51
	Mulet, dimba, regn, urväder, snö	cloudy, fog, rain, stormy weather, snow	0.33	1.00	7–8	9.50	0.58	1.00	7–8	9.71
Jun 1784–Dec 1815	Klar himmel	clear sky	0.15	0.15	0–1	1.01	0.12	0.12	0–1	0.95
	Ganska glesa moln, eller himmelen nästan klar	rather sparse clouds, or almost clear sky	0.15	0.31	2–3	3.09	0.13	0.25	2–3	3.02
	Glesa moln	sparse clouds	0.15	0.46	4–5	5.57	0.06	0.31	4–5	5.61
	Himmelen half mulen	half cloudy sky	0.14	0.60	6	7.50	0.07	0.38	6	7.50
	Himmelen nästan helt mulen	almost completely cloudy sky	0.19	0.78	7	8.75	0.14	0.52	7	8.75
	Himmelen öfver allt mulen, dimba, rägn, snö	completely cloudy sky, fog, rain, snow	0.22	1.00	8	10.00	0.48	1.00	8	10.00
Jan 1816–Jun 1841	Klart	clear	0.33	0.33	0–3	2.05	0.25	0.24	0–3	1.82
	Halvklart	half clear	0.06	0.39	4	5.00	0.05	0.30	4	5.00
	Strömoln	scattered clouds	0.27	0.66	5–6	7.05	0.10	0.40	5–6	7.06
	Mulet, dimma	cloudy, fog	0.34	1.00	7–8	9.50	0.60	1.00	7–8	9.71
Jul 1841–Dec 1858	Klart	clear	0.22	0.22	0–2	1.53	0.18	0.18	0–1	0.95
	Mest klart	mostly clear	0.07	0.29	3	3.75	0.05	0.23	2	2.50
	Halvklart	half clear	0.10	0.39	4	5.00	0.08	0.31	3–4	4.34
	Strömoln	scattered clouds	0.28	0.67	5–7	7.85	0.07	0.39	5–6	7.06
	Mest mulet	mostly cloudy	0.06	0.73	5–7	7.85	0.04	0.43	5–6	7.06
	Mulet	cloudy	0.27	1.00	8	10.00	0.57	1.00	7–8	9.71
1859–1872	0	0 (fourths)	0.17	0.17	0–1	1.01	0.15	0.15	0–1	0.95
	1	1	0.09	0.27	2	2.50	0.06	0.21	2	2.50
	2	2	0.39	0.65	3–7	6.80	0.16	0.37	3–5	4.92
	3	3	0.10	0.75	3–7	6.80	0.08	0.45	6	7.50
	4	4	0.25	1.00	8	10.00	0.55	1.00	7–8	9.71
1873–1960	0	0 (tenths)	0.07	0.07			0.07	0.07		
	1	1	0.08	0.15			0.06	0.13		
	2	2	0.06	0.20			0.04	0.17		
	3	3	0.08	0.29			0.05	0.22		
	4	4	0.05	0.34			0.02	0.24		
	5	5	0.06	0.40			0.03	0.27		
	6	6	0.07	0.46			0.03	0.30		

Table X. (*Continued*)

Period	Class		AMJJAS				ONDJFM			
	Swedish term	English translation	Frequency	Cum. freq.	Oktas	Tenths	Frequency	Cum. freq.	Oktas	Tenths
	7	7	0.07	0.53			0.03	0.34		
	8	8	0.09	0.62			0.05	0.39		
	9	9	0.11	0.73			0.08	0.46		
	10	10	0.27	1.00			0.54	1.00		
1961–2001	0	0 (oktas)	0.03	0.03			0.04	0.04		
	1	1	0.13	0.16			0.11	0.15		
	2	2	0.08	0.24			0.06	0.21		
	3	3	0.07	0.32			0.04	0.25		
	4	4	0.08	0.40			0.04	0.29		
	5	5	0.07	0.46			0.04	0.33		
	6	6	0.12	0.58			0.07	0.40		
	7	7	0.17	0.75			0.14	0.54		
	8	8	0.25	1.00			0.46	1.00		

from contemporary legends. There is no legend in the Stockholm observation diaries for the first of the three periods, but the symbols are similar to those used by the Meteorological Society in Mannheim (see Kington (1974)). We used a legend found in an observation diary from Umeå, northern Sweden, in 1796 (see Moberg (1998)) for the first period. Cloud legends in the Stockholm observation diary are found in the years 1823, 1835 and 1850. The first two are identical and define our interpretation in the second period. The third legend is different and defines our interpretation for the last period.

From 1 January 1859 onwards, the cloud amounts have been estimated by the observers and denoted with digits. During 1859–1872, the cloud-covered fraction of sky was reported in fourths (0–4), whereas tenths (0–10) were used for 1873–1960 and oktas since 1961. In addition to the changes of cloud-reporting systems, there has also been a number of changes of observation times (see Moberg *et al*. (2002) for details).

In the development of monthly series of percent cloud cover there were two main problems: (i) to translate the descriptive observations used during 1756–1858 to oktas (or tenths), and (ii) to account for the various sets of observation times. Our goal was to develop a monthly record where the entire series for 1756–2001 is directly comparable to the most recent part of the series for 1961–2001, since when the observations have been made in oktas at 07, 13 and 19h (all observation hours refer to local times before 1879 and Central European Time afterwards).

The translation of descriptive cloud observations to percent cloud-covered sky was made by comparing the frequency distribution of each cloud class in each of the four subperiods having different sets of symbols during 1756–1858 with the distribution of oktas during 1961–2001. We used only observations made at noon in this comparison, as the noon observations have always been made at 13h or 14h. The diurnal cycle of cloud amounts is near its maximum at this time of the day, and hence the distribution of cloud classes changes only a little between 13h and 14h. We did not use morning or evening observations because the diurnal cycle has a steep slope at this time of the day, and observations made at different hours may thus have different cloud amount distributions. Furthermore, we divided the year into two halves (April–September and October–March), because there is a strong annual cycle in the cloud amount distribution. In particular, the 8 oktas class is much more frequent in the October–March season than in April–September (see Table X).

We compared histograms of the relative frequency and the cumulative frequency of cloud symbols in each early subperiod with the corresponding histograms for oktas for 1961–2001, in an effort to define the range of oktas that best corresponded to each descriptive cloud term. It was also necessary to perform this kind of comparison for the period 1859–73, when observations were made in fourths (only five classes). The interpretation was sometimes easy, when the class limits clearly coincided, but was more difficult in other cases. For example, the term 'clear' was used in all subperiods, but its translation to a range of oktas varies from 0–1 to 0–4 in the various subperiods. It is thus obvious that 'clear' does not mean that the cloud amount was strictly zero. The term 'cloudy' covers the range 7–8 oktas in some periods, but corresponds to 8 oktas in other periods. The final choice of translation to a range of oktas, given in Table X, is therefore partly subjective.

Once the range of oktas had been defined for each cloud symbol, we then assigned a unique number to each symbol by calculating the average cloud amount in the respective range of oktas during 1961–2001. We also applied the corresponding numbers to the morning and evening observations. A comparison of the average differences (noon–morning and noon–evening) in different periods revealed that the numbers were not applicable to the morning and evening observations. On deciding whether either to reject the morning and evening observations or keep them with some kind of adjustment, we chose the latter. Hence, all morning and evening observations were adjusted by an amount so that the average differences (noon–morning and noon–evening) are the same in all subperiods of different observation times. Finally, monthly cloud amount averages could be calculated, where all data are adjusted corresponding to observations made at 07, 13 and 19h in a statistical sense (*i.e.* using the 1961–2001 diurnal cycle).

A.2. Uppsala

Cloud observations have been made in Uppsala since 1722, but here we only consider data from 1780. The earlier data were not considered because the distribution of observation hours differed from that in the period after 1780 and because our main investigation period begins in 1780. Between January 1780 and May 1832 the observations were generally taken at sunrise and in the afternoon. Hence, they include both night-time and daytime cloud observations in the summer half year. From May 1832 to May 1865 the observations were made three times per day (07, 14, 21h). From June 1865 to July 1868 hourly observations were made, and from August 1868 to 1958 cloud observations were made eight times per day (06, 08, 10, 12, 14, 17, 19, 21h). During 1959–1982 there were three cloud observations per day (07, 13, 19h), and only one daily observation (at 07h) was made for 1983–84. After 1984 there were no cloud observations made in Uppsala, but observations are available from the Uppsala airport, located about 4 km north-northwest of the city site. Here, we use observations from the airport for 1961–1995, made at three-hourly intervals (01, 04, 07, 10, 13, 16, 19, 22h). We compared the city and airport observations in the overlapping period 1961–82 and found that monthly averages for observations at 07, 13 and 19h differed (at the 5% significance level) only in four months (March, April, September, December), where the airport had 1.2–1.5% units higher cloud amounts. The distribution of oktas was very similar at the two sites. We consider the airport site to be the better one from 1961 onwards, as it contains eight daily observations. All monthly averages for the period 1961–95 in the series we develop here are taken from the airport. All data for 1780–1960 are taken from the city site.

Cloud observations during 1780–1854 were made with descriptive terms. Twelve different terms were used, ranging from 'clear' to 'cloudy'. They are listed in Table XI. To translate the descriptive terms to numeric cloud amounts we used a similar approach to that for Stockholm, i.e. we compared the distribution of cloud terms in early data with the distribution of oktas in modern data. The airport data for 1961–95 were used as modern data in this comparison. Hence, the monthly cloud amounts obtained for the early period 1780–1854 become directly comparable to the airport site, and no further adjustment was needed once the translation had been done.

In contrast to the Stockholm case, we had access to information of the diurnal cycle of cloud amounts for Uppsala from the airport data. Therefore, we were able to derive translations of the descriptive terms both for the sunrise and noon observations from 1780 to April 1832 and for all three daily observations from May 1832 to December 1854. As the morning observation time in the earlier period varied during the year (following the time of sunrise), we identified five groups of months with approximately the same observation hour (see Table XI). The distribution of cloud terms in morning observations within these five groups was compared with the distribution of oktas at approximately the same time of the day for the airport data. The afternoon observation, generally being made at 14–15h in the early period was compared with the 13h observation in the cold half year (October–March) and with the combined 13h and 16h observations in the warm half

Table XI. Relative frequencies and cumulative relative frequencies for each class of cloud observations in Uppsala in various periods. The approximate range of oktas for each class and the corresponding average in tenths are indicated

Class: Swedish term	Class: English translation	Freq.	Cum. freq.	Oktas	Tenths	Freq.	Cum. freq.	Oktas	Tenths	Freq.	Cum. freq.	Oktas	Tenths
		Jan 1780–Apr 1832 Jun–Jul, Morning				Jan 1780–Apr 1832 May, Aug, Morning				Jan 1780–Apr 1832 Mar, Apr, Sep, Morning			
Klart	clear	0.30	0.30	0–1	1.03	0.38	0.38	0–2	1.41	0.35	0.35	0–3	1.47
Nästan klart	almost clear	0.02	0.32	2	2.50	0.02	0.40	3	3.75	0.01	0.37	0–3	1.47
Mest klart	mostly clear	0.06	0.38	2	2.50	0.05	0.44	3	3.75	0.04	0.41	4	5.00
Halvklart	half clear	0.05	0.43	3	3.75	0.06	0.50	4	5.00	0.05	0.46	5	6.25
Strömulet	scatter cloudy	0.18	0.62	4–5	5.62	0.05	0.55	5	6.25	0.03	0.49	6	7.50
Strömoln	scattered clouds	0.00	0.62	4–5	5.62	0.00	0.55	5	6.25	0.00	0.49	6	7.50
Glesmulet	sparse cloudy	0.02	0.63	6	7.50	0.01	0.56	5	6.25	0.01	0.49	6	7.50
Halvmulet	half cloudy	0.05	0.68	6	7.50	0.03	0.60	6	7.50	0.02	0.52	6	7.50
Mest mulet	mostly cloudy	0.05	0.73	7–8	9.41	0.05	0.64	6	7.50	0.03	0.55	7–8	9.63
Nästan mulet	almost cloudy	0.00	0.73	7–8	9.41	0.01	0.65	7–8	9.49	0.00	0.56	7–8	9.63
Molnigt	cloudy	—	—	—	—	—	—	—	—	—	—	—	—
Mulet	cloudy	0.27	1.00	7–8	9.41	0.35	1.00	7–8	9.49	0.44	1.00	7–8	9.63
		Jan 1780–Apr 1832 Feb, Oct, Morning				Jan 1780–Apr 1832 Jan, Nov, Dec, Morning				Jan 1780–Apr 1832 AMJJAS, Afternoon			
Klart	clear	0.24	0.24	0–3	1.88	0.20	0.20	0–2	1.27	0.21	0.21	0–2	1.42
Nästan klart	almost clear	0.01	0.25	0–3	1.88	0.01	0.21	0–2	1.27	0.01	0.22	3	3.75
Mest klart	mostly clear	0.03	0.28	4	5.00	0.03	0.24	3	3.75	0.05	0.28	3	3.75
Halvklart	half clear	0.05	0.33	5	6.25	0.05	0.29	4–5	5.66	0.06	0.34	4	5.00
Strömulet	scatter cloudy	0.02	0.35	6	7.50	0.01	0.30	4–5	5.66	0.26	0.60	5–6	6.98
Strömoln	scattered clouds	0.00	0.35	6	7.50	0.00	0.30	4–5	5.66	0.00	0.60	5–6	6.98
Glesmulet	sparse cloudy	0.01	0.36	6	7.50	0.01	0.31	6	7.50	0.02	0.62	5–6	6.98
Halvmulet	half cloudy	0.02	0.37	6	7.50	0.02	0.33	6	7.50	0.04	0.65	7–8	9.44
Mest mulet	mostly cloudy	0.02	0.40	7–8	9.69	0.02	0.34	6	7.50	0.05	0.70	7–8	9.44
Nästan mulet	almost cloudy	0.00	0.40	7–8	9.69	0.00	0.35	7–8	9.67	0.01	0.71	7–8	9.44
Molnigt	cloudy	—	—	—	—	—	—	—	—	—	—	—	—
Mulet	cloudy	0.60	1.00	7–8	9.69	0.65	1.00	7–8	9.67	0.29	1.00	7–8	9.44
		Jan 1780–Apr 1832 ONDJFM, Afternoon				May 1832–Dec 1854 AMJJAS, All obs.				May 1832–Dec 1854 ONDJFM, All obs.			
Klart	clear	0.22	0.22	0–3	1.77	0.29	0.29	0–2	1.38	0.22	0.22	0–2	1.02
Nästan klart	almost clear	0.01	0.23	0–3	1.77	0.10	0.40	3	3.75	0.07	0.29	3	3.75
Mest klart	mostly clear	0.03	0.26	4	5.00	—	—	—	—	—	—	—	—
Halvklart	half clear	0.06	0.32	5	6.25	0.03	0.43	4	5.00	0.04	0.32	4	5.00
Strömulet	scatter cloudy	0.04	0.36	6	7.50	0.10	0.53	5–6	6.99	0.04	0.37	5–6	7.00
Strömoln	scattered clouds	0.00	0.36	6	7.50	0.10	0.63	5–6	6.99	0.04	0.40	5–6	7.00
Glesmulet	sparse cloudy	0.02	0.38	6	7.50	—	—	—	—	—	—	—	—
Halvmulet	half cloudy	0.02	0.40	6	7.50	0.02	0.65	7	8.75	0.02	0.42	5–6	7.00
Mest mulet	mostly cloudy	0.03	0.43	7–8	9.66	0.00	0.65	7	8.75	0.00	0.42	7–8	9.66
Nästan mulet	almost cloudy	0.00	0.43	7–8	9.66	0.07	0.72	7	8.75	0.07	0.49	7–8	9.66
Molnigt	cloudy	—	—	—	—	0.03	0.75	7	8.75	0.01	0.50	7–8	9.66
Mulet	cloudy	0.57	1.00	7–8	9.66	0.25	1.00	8	10.00	0.50	1.00	7–8	9.66

(April–September). Data for the period May 1832 to December 1854 were also divided into warm and cold half years, and all observations were compared with the combined airport observations at 07, 13 and 22h.

From 1855 onwards, the actual cloud amounts were estimated directly and given in the observation registers (in tenths for 1855–1957; in oktas for 1958–84). Before calculation of monthly averages, the observations at each particular time were adjusted by the average difference between cloud amount at that time of the day and the 'true' diurnal mean. These adjustments were obtained from the airport data for 1961–95. This procedure thus adjusts all data to 'true' monthly averages. The principle is the same as that used for calculation of daily mean temperatures in Uppsala (Bergström and Moberg, 2002), and is as important for cloud amounts as for temperature series. Finally, the March, April, September and December data for 1855–1960 were adjusted to the airport level by adding +1.5, +1.5, +1.4 and +1.2% units respectively.

APPENDIX B: ANALYSES OF DATA FOR THE OTHER SEASONS

We undertook similar analyses to the data discussed in Section 4 for all parts of the year in order to obtain a complete picture of seasonal behaviour of data properties, correlation patterns and homogeneity problems. Some of these results are briefly summarized here.

Concerning the possibility of estimating cloud amounts from other data: we found two useful estimates for the November–January (NDJ) season and one for the March–May (MAM) season. For NDJ, one estimate was based on a combination of westerly wind and temperature ($R^2 = 59\%$), whereas the other was based on a combination of westerly and southerly wind ($R^2 = 69\%$). Both models were found to agree rather well with the observed regional winter cloud amount series (including the low-frequency behaviour). This suggests that the general homogeneity problems for Fennoscandian cloud data mostly affect summer data, whereas winter data seem much more reliable. Our model for MAM (using vorticity and southerly wind, $R^2 = 58\%$) disagrees with the sign of observed century-scale trends (decreasing in model, increasing in the regional series), but less markedly than in JJA.

We also performed homogeneity tests of the raw cloud series back to 1780 for the MAM and NDJ seasons. For NDJ, both Stockholm and Uppsala were found to be homogeneous with respect to the reference series used (u, v), apart from a few subperiods. Stockholm was also considered nearly homogeneous in the MAM season, whereas the corresponding Uppsala data were found to be quite inhomogeneous. These results further point out summer cloud data as being particularly prone to homogeneity problems, and winter cloud data as being less problematic.

The seasonal difference between estimated and observed cloud amounts is likely related to seasonally varying problems with the subjective classification of cloud amounts, as there are strong seasonal contrasts both in mean cloud amounts and dominant cloud types. There is a substantially higher average cloud amount in winter than summer. In summer cumulus clouds are relatively more important, whereas in winter the clouds are more often related to fronts (Raab and Vedin, 1995). Furthermore, the diurnal cycle of cloud amounts is larger in summer than in winter in the Stockholm and Uppsala region. The distribution among the okta classes is also more evenly spread in summer than in winter. For example, in November to February, ~50% of the 13h cloud observations for 1961–2001 at Stockholm fell in the 8 okta class, whereas in the JJA months only about 20% fell in this class. Hence, the sky is much more often overcast in winter compared with summer, which implies a smaller risk for different interpretations made by different observers in winter. This risk should be larger in summer, because of the more even distribution among classes and the relatively more important cumulus clouds.

ACKNOWLEDGEMENTS

AM and PDJ were supported by the Office of Science, US Department of Energy, under grant no. DE-FG02-98ER62601. We thank Øyvind Nordli for providing the Austlandet series and for discussions concerning the comparisons with Stockholm and Uppsala temperatures. Two anonymous reviewers are also thanked for their suggestions.

REFERENCES

Alexandersson H. 1986. A homogeneity test applied to precipitation data. *International Journal of Climatology* **6**: 661–675.

Alexandersson H. 1994. Some comments on recent climate fluctuations — from a Swedish horizon. In *Climate Variations in Europe, Proceedings of the European Workshop on Climate Variations held in Kirkkonummi (Majvik), Finland, 15–18 May 1994*, Heino R (ed.). Publications of the Academy of Finland 3/94. Academy of Finland: Helsinki; 293–305.

Ångström A. 1946. *Sveriges Klimat*. Generalstabens Litografiska Anstalt: Stockholm; 105 pp.

Bergström H, Moberg A. 2002. Daily air temperature and pressure series for Uppsala (1722–1998). *Climatic Change* **53**: 213–252.

Böhm R, Auer I, Brunetti M, Maugeri M, Nanni T, Schöner W. 2001. Regional temperature variability in the European Alps: 1760–1998 from homogenized instrumental time series. *International Journal of Climatology* **21**: 1779–1801.

Camuffo D, Jones P (eds). 2002. Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources. *Climatic Change* **53**(1–3).

Chen D. 2000. A monthly circulation climatology for Sweden and its application to a winter temperature case study. *International Journal of Climatology* **20**: 1067–1076.

Eriksson B. 1981. Statistisk analys av nederbördsdata, Del III, 200-åriga nederbördsserier. SMHI rapporter. Meteorologi och klimatologi, Sveriges Meteorologiska och Hydrologiska Institut, Norrköping.

Heino R. 1994. *Climate in Finland During the Period of Meteorological Observations, Academic Dissertation*. Finnish Meteorological Institute Contributions No. 12. Finnish Meteorological Institute: Helsinki.

Henderson-Sellers A. 1992. Continental cloudiness changes this century. *Geo Journal* **27**: 255–262.

Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds). 2001. *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press: Cambridge and New York.

Jacobeit J, Jönsson P, Bärring L, Beck C, Ekström M. 2001. Zonal indices for Europe 1780–1995 and running correlations with temperature. *Climatic Change* **48**: 219–241.

Jones PD. 1987. The early twentieth century Arctic High — fact or fiction? *Climate Dynamics* **1**: 63–75.

Jones PD. 1994. Hemispheric surface air temperature variations: a reanalysis and an update to 1993. *Journal of Climate* **7**: 1794–1802.

Jones PD, Lister DH. 2002. The daily temperature record for St. Petersburg (1743–1996). In *Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Records*, Camuffo D, Jones P (eds). *Climatic Change* **53**: 253–267.

Jones PD, Davies TD, Lister DH, Slonosky V, Jónsson T, Bärring L, Jönsson P, Maheras P, Kolyva-Machera F, Barriendos M, Martin-Vide J, Rodriguez R, Alcoforado MJ, Wanner H, Pfister C, Luterbacher J, Rickli R, Schuepbach E, Kaas E, Schmith T, Jacobeit J, Beck C. 1999. Monthly mean pressure reconstructions for Europe for the 1780–1995 period. *International Journal of Climatology* **19**: 347–364.

Jones PD, Osborn TJ, Briffa KR. 2003. Pressure-based measures of the NAO: a comparison and an assessment of changes in the strength of the NAO and in its influence on surface climate parameters. In *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, Hurrell JW, Visbeck M, Kushnir Y, Ottersen G (eds). Geophysical Monograph Series Vol. 134. American Geophysical Union: Washington, DC; 51–62 (DOI:10.1029/134GM03).

Jones PD, Moberg A, Osborn TJ, Briffa KR. In press. Surface climate responses to explosive volcanic eruptions seen in long European temperature records and mid-to-high latitude tree-ring density around the Northern Hemisphere. In *Volcanism and the Earth's Atmosphere*, Robock A, Oppenheimer C (eds). American Geophysical Union: Washington, DC.

Kington JA. 1974. The Societas Meteorologica Palatina: an eighteenth-century meteorological society. *Weather* **29**(November): 416–426.

Moberg A. 1998. Meteorological observations in Sweden made before A.D. 1860. *Paläoklimaforschung/Palaeoclimate Research* **23**: 99–119.

Moberg A, Bergström H. 1997. Homogenization of Swedish temperature data. Part III: the long temperature records from Stockholm and Uppsala. *International Journal of Climatology* **17**: 667–699.

Moberg A, Bergström H, Ruiz Krigsman J, Svanered O. 2002. Daily air temperature and pressure series for Stockholm (1756–1998). In *Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Records*, Camuffo D, Jones P (eds). *Climatic Change* **53**: 171–212.

Nordli PØ. 2001a. *Spring and summer temperatures in south eastern Norway (1749–2000)*. DNMI report No. 01/01 Klima. Norwegian Meteorological Institute, Oslo.

Nordli PØ. 2001b. Reconstruction of nineteenth century summer temperatures in Norway by proxy data from farmers' diaries. *Climatic Change* **48**: 201–218.

Raab B, Vedin H (eds). 1995. *National Atlas of Sweden: Climate, Lakes and Rivers*. SNA Publ. (Sveriges nationalatlas), Almqvist & Wiksell International: Stockholm.

Slonosky V, Jones PD, Davies TD. 2001. Atmospheric circulation and surface temperature in Europe from the 18th century to 1995. *International Journal of Climatology* **21**: 63–75.

Tabony RC. 1980. A set of homogeneous European rainfall series. Met. O. 13 Branch Memorandum No. 104. Meteorological Office, Bracknell.

Tuomenvirta H, Alexandersson H, Drebs A, Frich P, Nordli PØ. 2000. Trends in Nordic and Arctic temperature extremes and ranges. *Journal of Climate* **13**: 977–990.

Tuomenvirta H, Drebs A, Førland E, Tveito OE, Alexandersson H, Vaarby Laursen E, Jónsson T. 2001. Nordklim data set 1.0 — description and illustrations. DNMI Report No. 08/01 Klima. Norwegian Meteorological Institute, Oslo.

Von Storch H, Zwiers FW. 1999. *Statistical Analysis in Climate Research*. Cambridge University Press: Cambridge.

Wallén CC. 1953. The variability of summer temperature in Sweden and its connection with changes in the general circulation. *Tellus* **2**: 157–178.

