DNMI

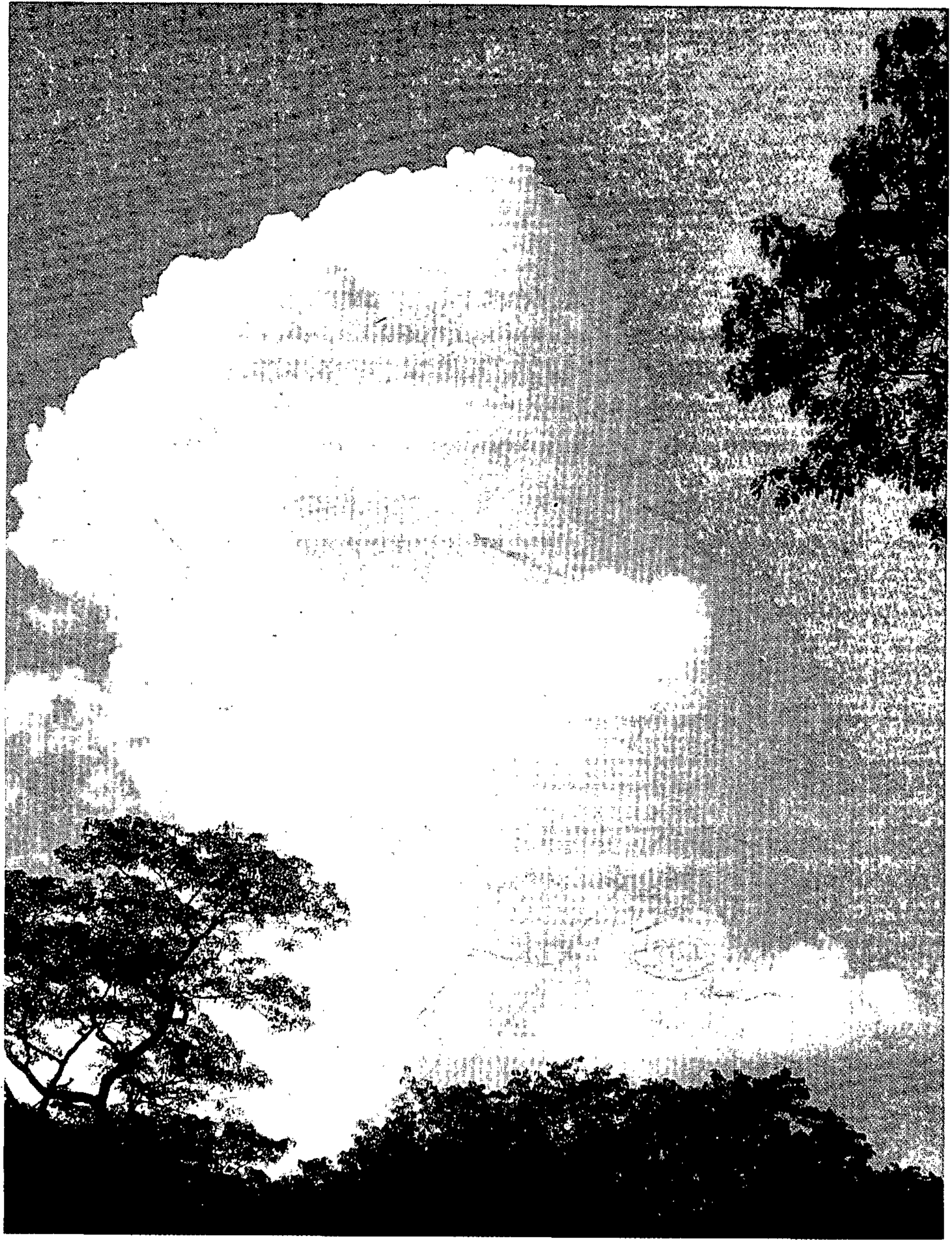
DET NORSKE METEOROLOGISKE INSTITUTT

klima

PRINCIPAL COMPONENT ANALYSIS OF THE NACD TEMPERATURE SERIES

I. HANSSEN-BAUER, P.Ø. NORDLI AND E.J. FØRLAND

REPORT NO. 1/96 KLIMA



Contents:

1. Introduction.

1.1 Background.

Most meteorological stations with long temperature series have experienced changes in instruments, location or environments which may have affected the measurements (e.g. Hanssen-Bauer and Førland 1994, Nordli et al. 1996). Except from special studies of such effects (e.g. urbanization), studies of climate variations should always be based upon homogeneous meteorological series. The North Atlantic Climatological Dataset (NACD) contains monthly values of several meteorological elements from a number of stations in Belgium, Denmark, Finland, Great Britain, Iceland, Ireland, the Netherlands, Norway and Sweden, during the period 1890-1990 (Frich et al., 1996). The series are classified with a quality code (see Appendix A), and most of the temperature series have been thoroughly tested for homogeneity. The dataset thus forms an excellent base for studies of climate variations within the 100 year period. The present report contains principal component analysis of the NACD temperature series which were available in November 1995. The main objective is to describe typical variations in the annual mean temperature in the area during the period.

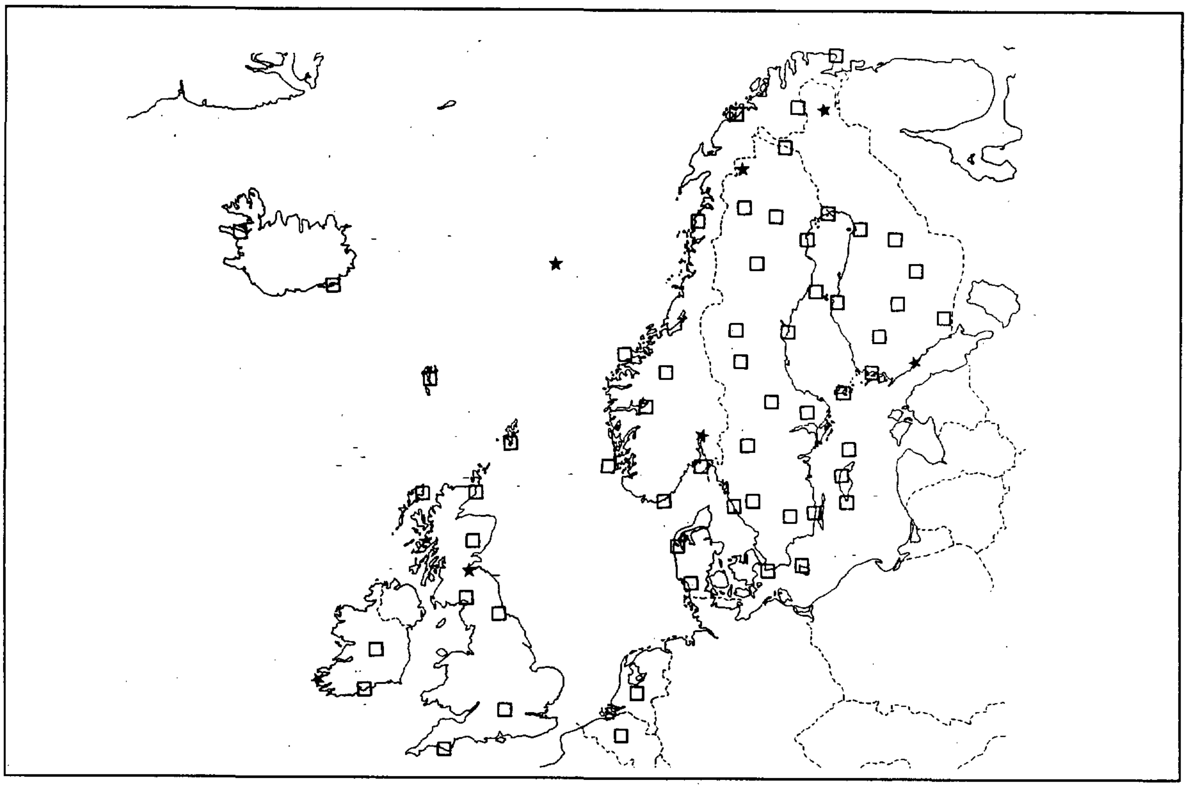


Figure 1. Map showing the positions of stations used in the principal component analysis (open squares) and stations used for verification (stars).

1.2 Data.

The NACD contains 95 temperature series. Some series were not yet available in November 1995. Series with quality code E (environmental changes) or I (inhomogeneous), and series which do not cover the entire period 1891-1990, were excluded from the present analysis. The remaining 53 NACD-series were, together with 5 additional series, used in the present analyses. In order to verify estimated temperature series on independent data, data from 7 other stations were applied. Figure 1 shows positions of the 65 stations from which test series (marked by squares) and verification series (marked by stars) originates. All stations are situated between 50 and 70°N and between 25°W and 25°E. The stations are also listed in Appendix A, where their names, positions and other relevant information are given.

2. Principal component analysis.

2.1 Methods.

Principal component analysis (PCA) may be performed by using the correlation matrix or the covariance matrix (Preisendorfer 1988). In the present work, the analyses were mainly based upon the covariance matrix, as suggested for univariate studies by several authors (e.g. Mills 1995). More information is retained in this way, and it also simplifies the estimation of temperature series in °C (section 2.3). The drawback of using the covariance matrix is that series with large standard deviation (inland/high latitude) influence the principal components more than series with small standard deviation (coastal/low latitude). For comparative purposes, one analysis was performed using the correlation matrix.

The PCA produces loadings (weight coefficients) at each station, and time series of scores (amplitude functions) for each principal component. These are presented in section 2.2. In order to estimate temperature series representative for an arbitrary point or for an area, it is necessary to interpolate values for the loadings. By adding the weighted contributions from each amplitude function, it is possible to generate time series of temperature anomalies relative to the 100 year mean value. Estimates of local temperature series are presented in sections 2.3 and 2.4. Estimates of regional curves are presented in section 2.5.

2.2 Results from PCA of annual mean temperature during 1891-1990.

The first principal component accounts for 80% of the variance in the set of annual mean temperatures, while 5 components were needed to account for 95% of the variance (table 1). Figure 2 shows that more than 97.5% of the variance is accounted for in most parts of Finland, while only about 70% is accounted for in southwestern parts of Ireland.

Table 1. Eigenvalues and proportion of total variance accounted for by each of the first 5 principal components from PCA of 58 series during 1891-1990.

	Eigenvalue	Proportion of total variance	Cumulative proportion
PC1	36.03	0.80	0.80
PC2	3.59	0.08	0.88
PC3	1.94	0.04	0.92
PC4	0.69	0.02	0.94
PC5	0.53	0.01	0.95

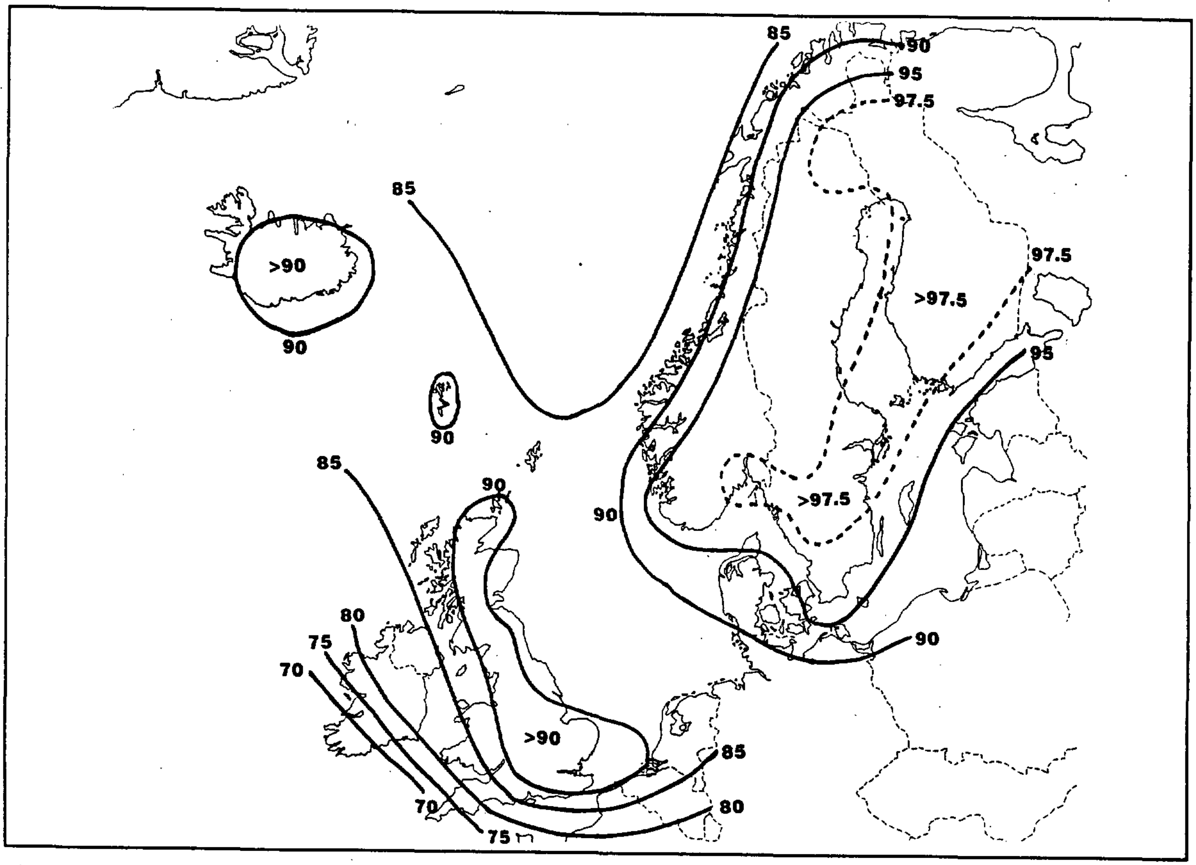


Figure 2. Percentage of variance accounted for by the first 5 principal components.

This variation is partly caused by differences in the standard deviation of annual mean temperatures, which is 1.0-1.2°C in Finland and 0.4-0.5 °C in the Irish series. For PCA based upon the correlation matrix, 80-90% of the variance in the Irish series and about 96% in the Finnish series is accounted for by using 5 components.

Figure 3 shows the loadings of the 3 first principal components (PC1-PC3), while figure 4 shows the time series of PC scores. The series of PC scores were filtered by applying Gaussian weight functions with standard deviations of 3 (filter F1) and 9 (filter F2) years, respectively. The filtered series are also presented in figure 4, in order to visualise decadal variations (F1) and long term trends (F2).

The loadings of PC1 (figure 3a) are positive all over the area, making PC1 a weighted mean of the temperature series from the entire area. The loadings are at maximum in eastern parts of the area, where PC1 accounts for more than 90% of the variance. In western Iceland on the other hand, it accounts for less than 10% of the variance. Years with high mean temperatures in Fennoscandia make high positive scores on PC1, while cold years in this area make high negative scores (figure 4a).

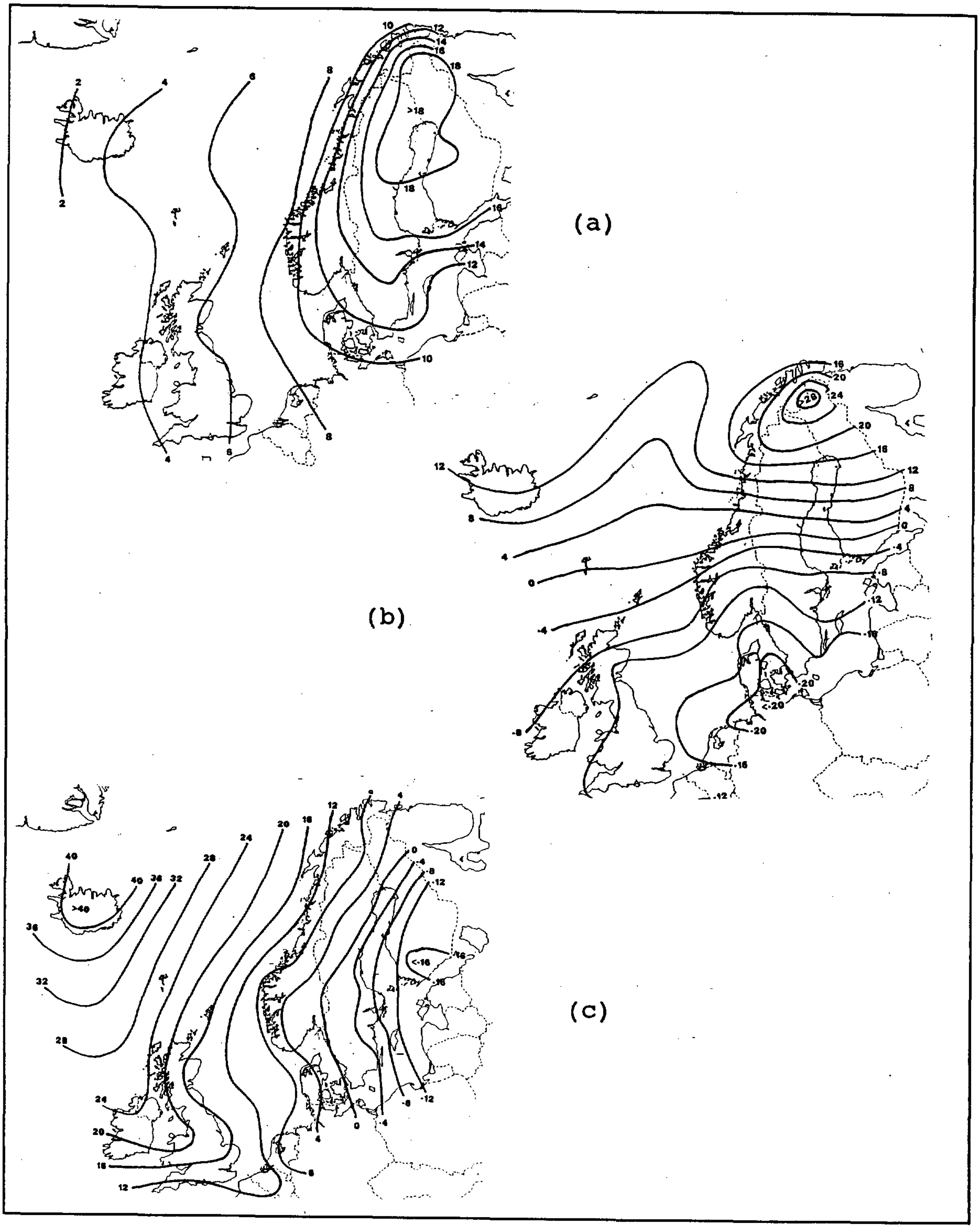


Figure 3. Loadings x 100 of a) first , b) second, c) third principal component.

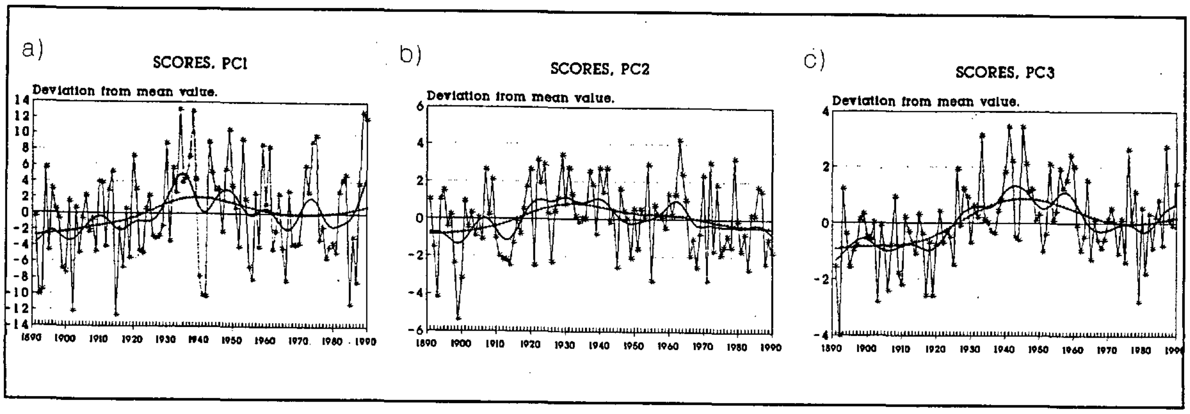


Figure 4. Amplitude functions for a) first , b) second, c) third principal component.

The loadings of PC2 (fig. 3b) are positive in northern areas, close to zero in central areas, and negative in southern areas. About 20% of the variance in the southernmost and northernmost areas is accounted for by PC2. Years with larger positive (or smaller negative) temperatures anomalies in northern than in southern areas, get positive PC2 scores. Years wich are colder in the north than in the south when compared to the longterm mean, get negative PC2 scores (figure 4b).

The loadings of PC3 (fig. 3c) are at maximum in Iceland. The values are decreasing eastwards, and small, negative loadings are found in Finland and parts of Sweden. About 55% of the variance in the Icelandic series is accounted for by this component, while less than 5% is accounted for in the Finnish and Swedish series. Years characterised by high mean temperatures in western parts of the area thus get high PC3 scores (figure 4c).

PC4 and PC5 are of little importance for the area as a whole (table 1), but are included because of their regional importance. PC4, which apparently expresses a contrast between the south-western areas and Scandinavia, accounts for 5-10% of the variance in series from the Netherlands, Great Britain and Ireland. PC5 has positive loadings in Iceland, where it accounts for about 15% of the variance, while it has negative loadings and accounts for about 5% of the variance in southern Britain and in Ireland.

2.3 Estimates of local temperature series

The loadings and scores of PC1-PC5 were used in order to estimate temperature series for some stations (marked by stars in fig.1) which were not used in the PCA. The interpolated loadings are given in appendix B. Results from Edinburgh, Great Britain, are shown in figure 5. Observed and estimated annual mean temperatures follow each other very closely (fig. 5a). The differences between observed and estimated temperatures are within +/-0.5°C, and no significant trend is seen in these differences (fig. 5b).

Temperature series filtered by F1 and F2 (see section 2.2) are also shown. The correspondance between curves based upon observations and estimates is good for both filters (figs. 5c and 5d).

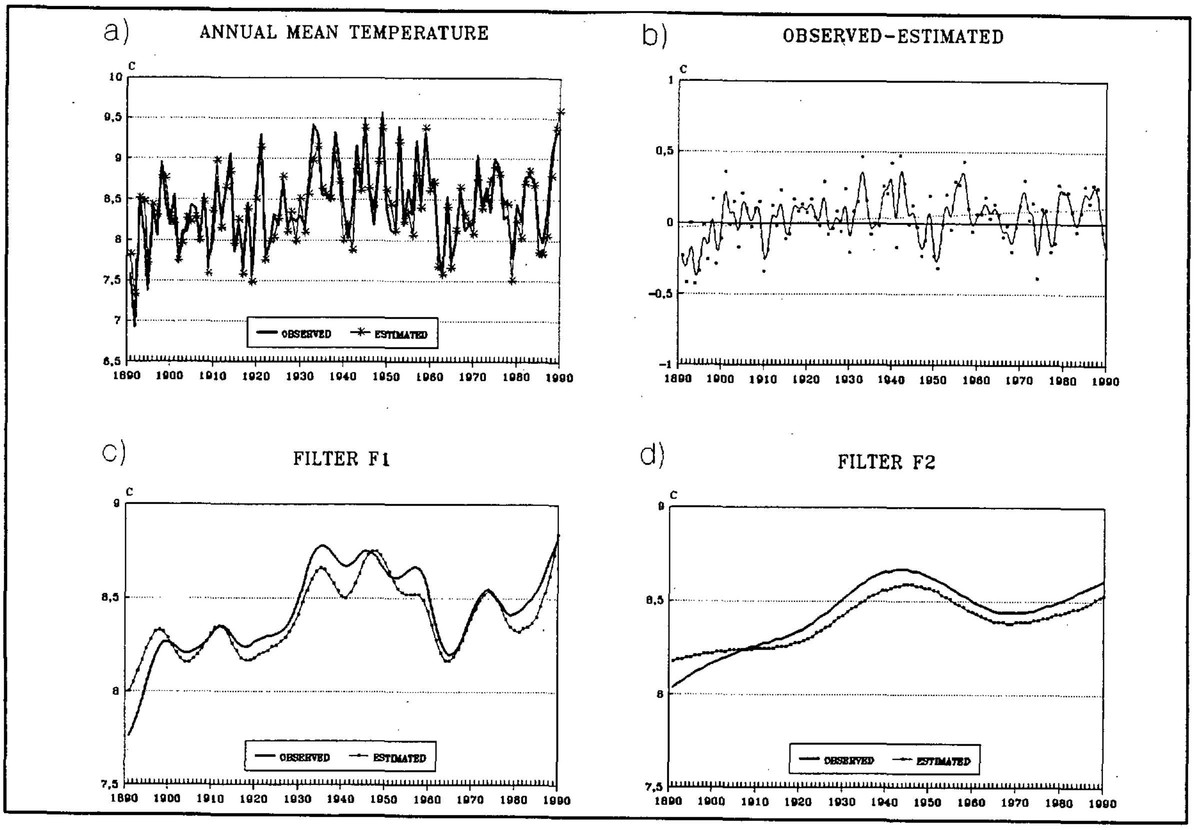


Figure 5. Annual mean temperature, Edinburgh, Great Britain. a) Observed and estimated, b) difference observed - estimated, c) observed and estimated filtered by F1, d) observed and estimated filtered by F2.

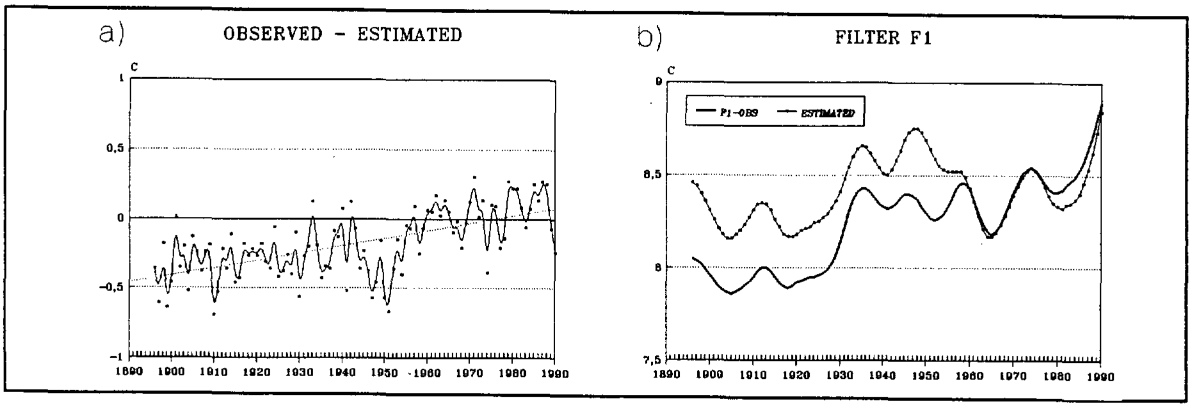


Figure 6. Inhomogeneous series of annual mean temperature, Edinburgh. a) Difference observed - estimated, b) observed and estimated series filtered by F1.

Eventual discrepancies between observed and estimated temperature does not necessarily mean that the estimate is poor. Figure 6 shows the estimated Edinburgh series compared to a preliminary Edinburgh series which was received while this series was still being homogenised. In the middle of the 1950's, there is a systematic change of the differences between estimated and observed values. It is satisfactory to see that no such change appeared in the final series (figure 5), which was homogenised totally independently of the present analysis.

Observed and estimated filter F1 curves for 4 other stations are presented in figure 7. The most isolated meteorological station within the area is the weathersip M (2°E, 66°N), where the 5 first PCs accounts for less than 85% of the variance. On this background, the temperature curve is reasonably well estimated (fig. 7a). The curves for Inari, Finland (fig. 7b) and Abisko, Sweden (fig. 7d) are very well estimated. This is what one should expect, as the 5 PCs account for about 95% of the variance in this area. Results from Valentia observatory, Ireland, are shown in figure 7c. Observed and estimated curves follow each other quite closely, even if this station is situated in an area where less than 70% of the variance is accounted for by PC1-PC5.

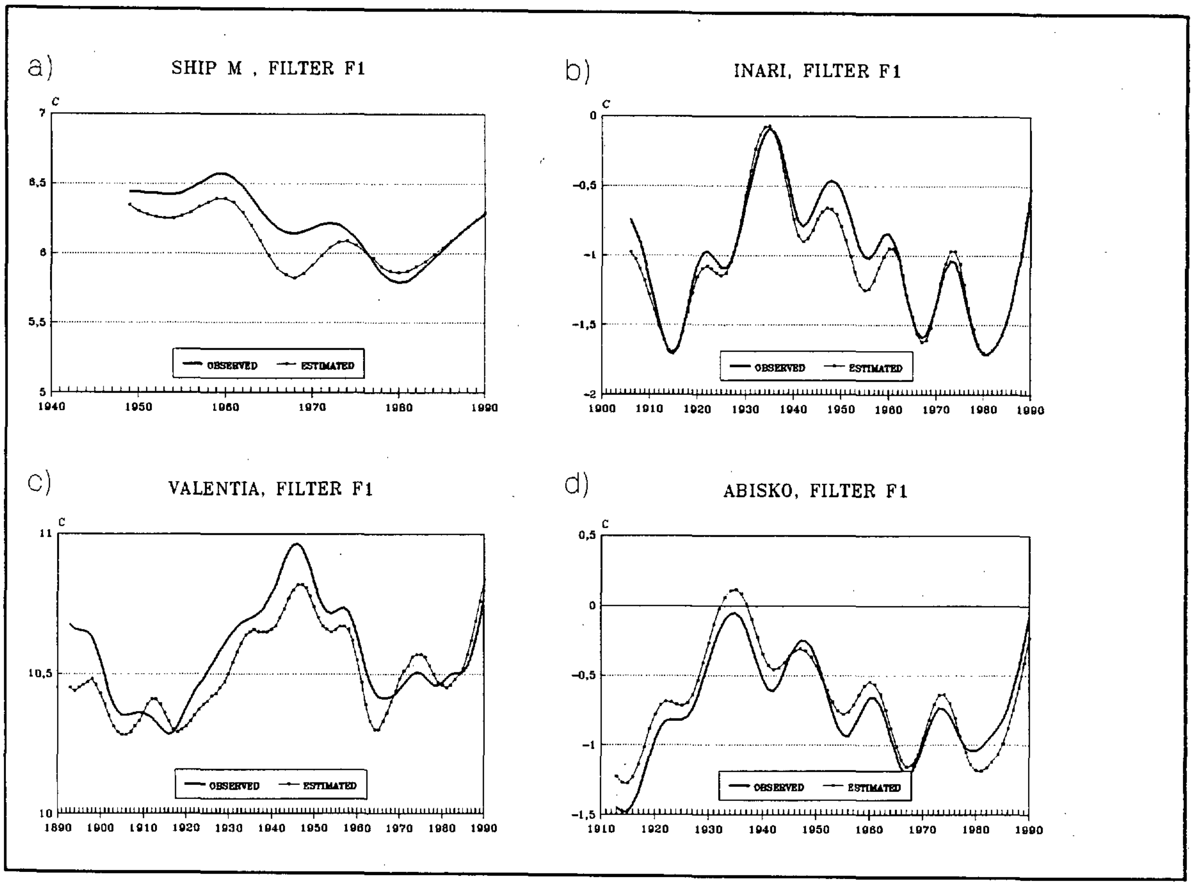


Figure 7. *Observed and estimated low pass filtered temperature series from a) Ship M, b) Inari, Finland, c) Valentia Observatory, Ireland, d) Abisko, Sweden.*

2.4 Use of PCA to detect urban varming.

Figures 5 and 6 indicate that estimates made by the principal components may be used as reference series in homogeneity studies, at least in areas with reasonable station coverage. It should thus also be possible to reveal eventual influence of urban warming by comparing observed series from cities to estimated series. The Oslo and the Helsinki NACD series are classified as E (see Appendix A), because of possible urban warming. Differences between observed and estimated temperatures in Oslo (figure 8a), as well as observed and estimated filter F1 series for the same station (figure 8d), show that the temperature in Oslo has increased by 0.1-0.2°C relatively to the estimated temperatures during the 100 year period. This may indicate that the urban warming in Oslo is small, but no conclusion should be drawn at this stage. The Oslo series was adjusted for a relocation in 1936-37. This adjustment may also have removed parts of the urban warming from the series.

Figure 8b) and e) indicates that the Helsinki series "warmed up" by almost 0.5°C during the 100 year period. The relative temperature increase happened mainly during the first 5 decades, and was at maximum during the 1920s. This is in general accordance with the analysis made by Heino (1994). Heino concludes that the urban warming in the Helsinki series amounts to 0.7-0.8°C from around 1860 to 1990, and that the warming rate was at its maximum (0.2°C/decade) during the 1920s. Heino (1994) adjusted the Helsinki series for urban warming as well as for relocations and screen changes. Figure 8c) and f) show the improved correspondance between the adjusted series and the estimated Helsinki series.

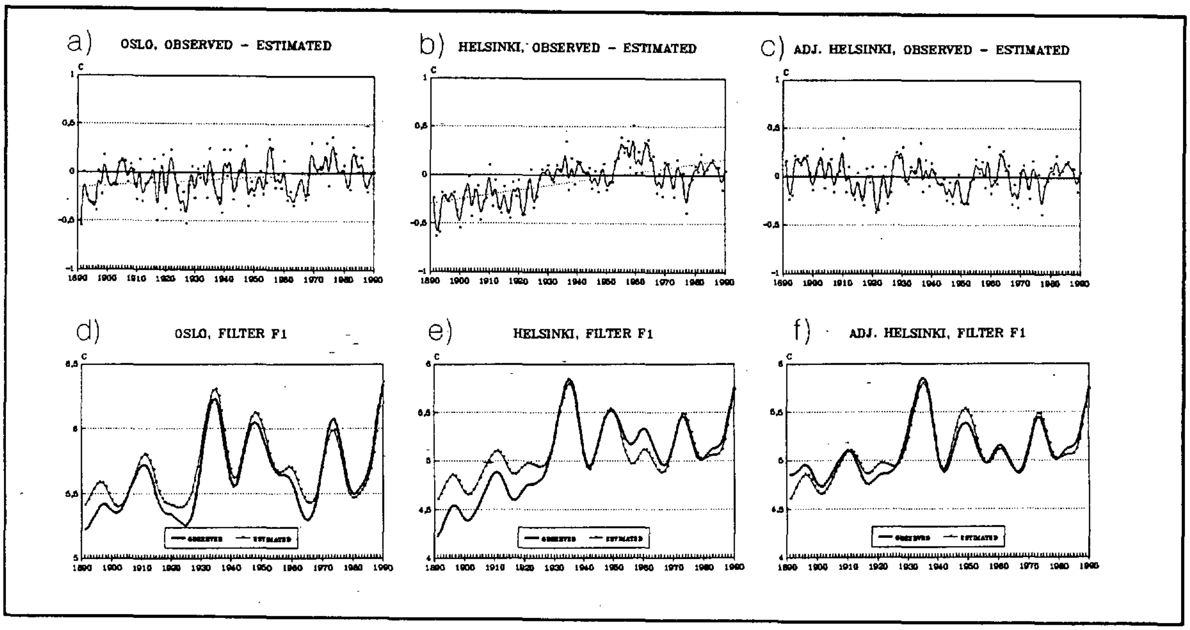


Figure 8. *Differences in annual mean temperature: a) observed - estimated, Oslo; b) observed - estimated, Helsinki; c) adjusted - estimated, Helsinki. Low pass filtered temperature series: d) observed and estimated, Oslo; e) observed and estimated, Helsinki; f) adjusted and estimated, Helsinki.*

2.5 Estimates of regional temperature series.

Temperature series for gridpoints or temperature series representative for larger regions may also be deduced from the PCA. Loadings for each PC are deduced from maps (cf. fig. 3). They may be representative for a gridpoint, a gridbox, or for any region. In the present study, the NACD area was divided into 6 temperature regions using following simple criteria concerning the influence of PC1 and PC2:

1) Regions where PC1+PC2 account for at least 75% of the variance are **eastern** (E). Other regions are **western** (W).

2) Regions where PC2 accounts for at least 5% of the variance and the PC2 loadings are positive are **northern** (N). Regions where PC2 accounts for <5% of the variance are **central** (C). Regions where PC2 accounts for at least 5% of the variance and the PC2 loadings are negative are **southern** (S).

Combining these criteria gives following regions: NW, CW, SW, NE, CE and SE. Approximate borders between them are shown in figure 9. Typical values of the loadings were interpolated for each region (Appendix B). Figure 10 shows the resulting regional temperature anomaly series, both annual values and series smoothed by the low pass filters F1 and F2.

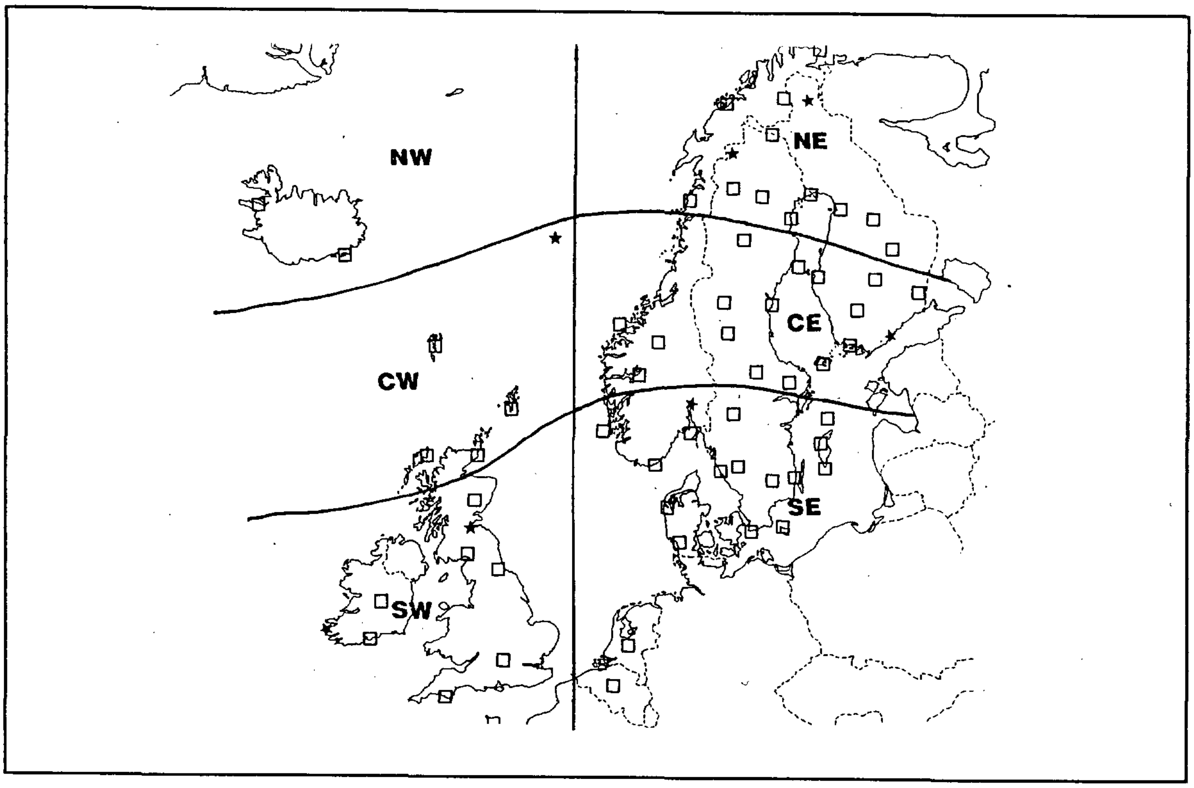


Figure 9. Temperature regions based upon the 3 first principal components.

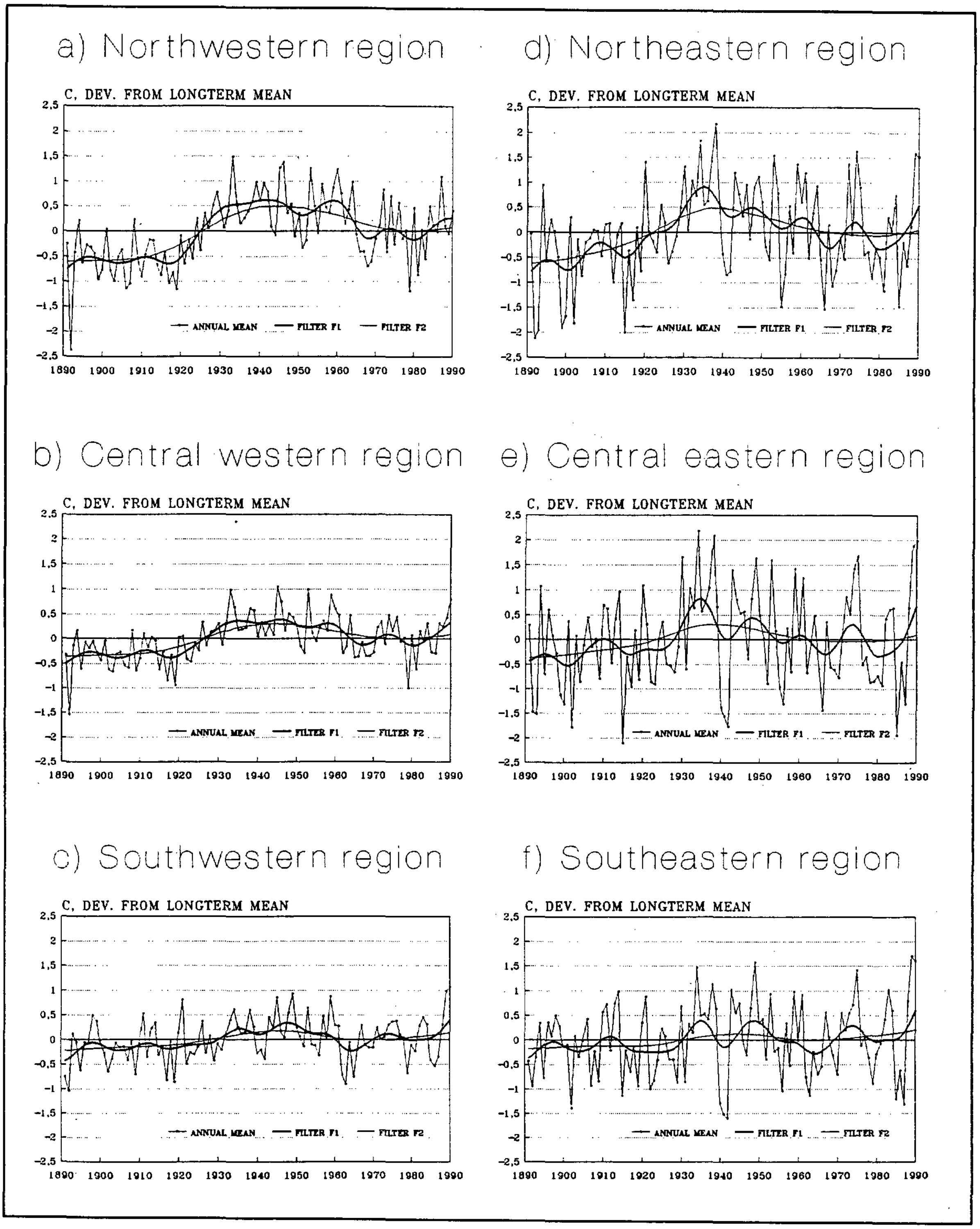


Figure 10. Temperature series representative for a) NW region, b) CW region, c) SW region, d) NE region, e) CE region and f) SE region.

3. Analyses of the regional temperature series.

3.1 Methods.

In analysing the regional time series (figure 10) three methods were used.

Visualising local extreme points and trends. Time series of scattered individual values often give a rather chaotic impression. To identify local maxima and minima as well as trends, the series may be smoothed by low pass filters. The filters F1 and F2 (section 2.2) have proved to be effective tools for analysing climatic series (Hanssen-Bauer et al., 1995).

Test for trends. The Mann - Kendall test is chosen for testing the significance of trends. As a nonparametric test it can be used without knowing the exact distribution of the time series. Its test statistic t is defined by the equation

$$t = \sum_{i=1}^{n} n_i \quad (1)$$

where n is the number of elements and n_i is the number of smaller elements preceding element x_i $(i = 1,2,..\ n)$ (Sneyers, 1990). Providing that $n > 10$ the test statistic is very nearly normally distributed under the hypothesis of randomness (the null hypothesis) (Sneyers, 1995). Moreover, its expectation E(t) and variance var(t) are given by the equations

$$E(t) = \frac{n(n-1)}{4} \quad (2)$$

$$var\ (t) = \frac{n(n-1)(2n+5)}{72} \quad (3)$$

The standardised distribution u(t) of the test statistic is then

$$u(t) = \frac{t - E(t)}{\sqrt{var\ t}} \quad (4)$$

A percent table of the normal distribution function may be used to decide wether the null hypothesis should be rejected or not.

Time series may be successively tested by adding one by one year reapplying the test for each year added. Using graphical representation of the standardised test statistic the development of trends in the series may easily be traced. It has also proved to be valuable to apply the test by starting with the last year going backward in time. These tools have been nicely applied by Demarée (1990).

Test for change points. When a time series consists of two random subseries, having different distributions, the series is said to have a change point at element k, which corresponds to the last element in the first subseries (Sneyers, 1995).

A tool for detecting change points is the Pettitt test statistic X_k defined by

$$X_k = 2R_k - k(n+1), \quad R_k = \sum_{i=1}^{k} r_i \quad (5)$$

where r_i is the rank of the i-th element in the complete series. If change points are absent in the series, i.e. under the null hypothesis of randomness, $E(X_k) = 0$. An extreme value X_E is significant for an abrupt decreasing or increasing shift if

$$|X_e| \geq \sqrt{\frac{(n^3 + n^2)(\ln m - \ln \alpha)}{6}} \quad (6)$$

where m = 1 (one sided test) or m = 2 (two sided test). α is the chosen level of significance.

3.2 Results

3.2.1 Low pass filtered temperature curves.

The smoothed temperature curves in figure 10 visualise the main features concerning the variability in annual mean temperature during 1891-1990. They indicate a positive trend, especially in northern regions, from relatively low temperatures during the first decades to a warm period in the 1930s. In western areas, the temperature level remained high during most of the following two decades. In eastern areas, this generally warm period was interrupted by a cold spell around 1940. All regions were relatively cold in the 1960s and around 1980, after which a positive trend is indicated.

3.2.2 Mann-Kendall test results.

The Mann-Kendall trend test was applied in order to investigate if any of the apparent trends were statistically significant. The series were tested successively by adding one by one year, and starting from both ends of the series, as described in section 3.1.

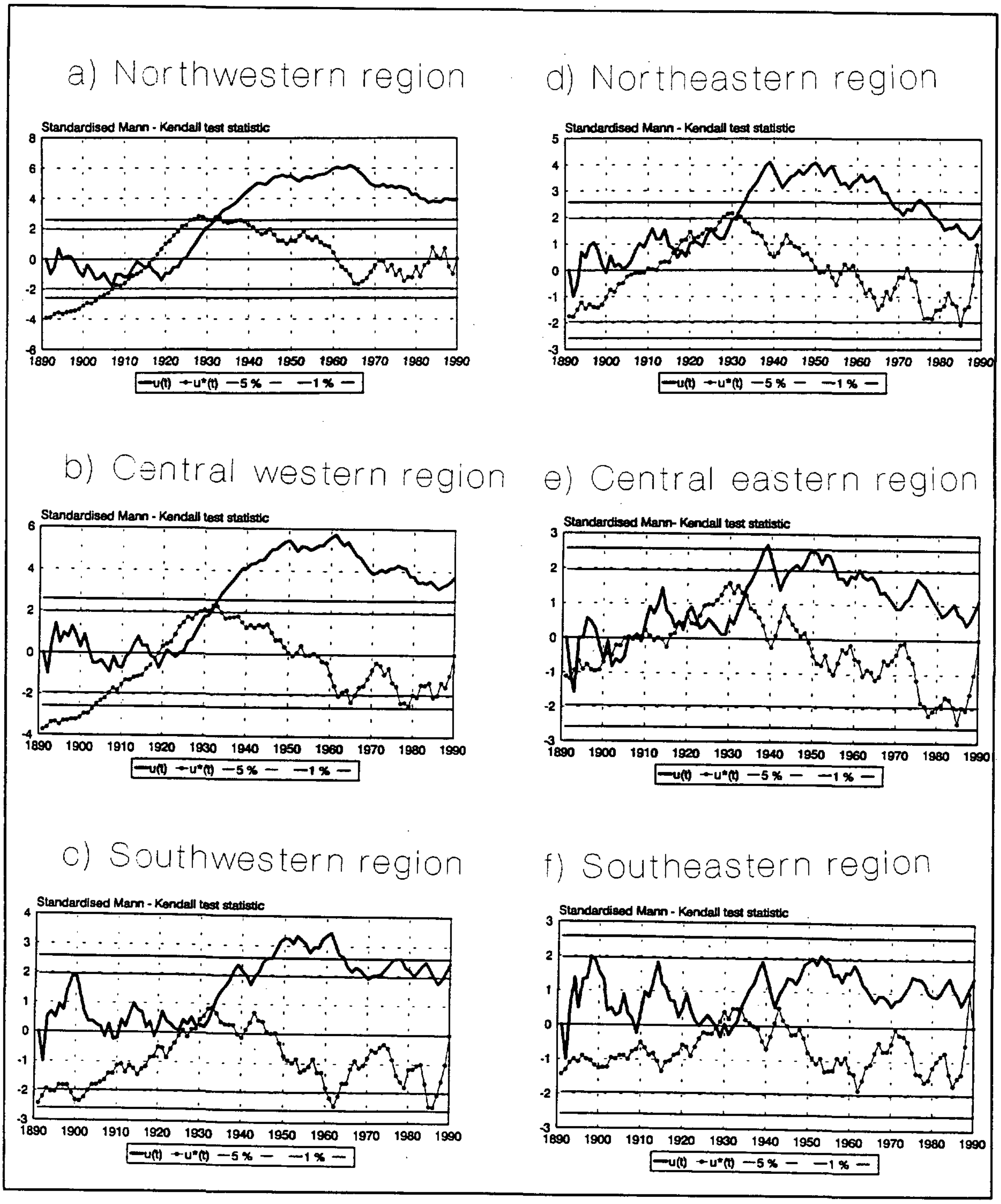


Figure 11. *Test statistics of the Mann-Kendall sequential trend test applied to the regional time series for a) NW, b) CW, c) SW, d) NE, e) CE and f) SE region. The functions $u(t)$ and $u^*(t)$ denote the forward and backward time evolution of the test statistics respectively. The indicated 1% and 5% significance levels are not valid for test periods shorter than approximately 10 years.*

Figure 11 a-f shows the standardised test statistics for the 6 regional temperature series. The significance levels shown in the figure are valid for testing periods of more than about 10 years. Weaker signals are needed to get statistically significant trends in the oceanic western regions than in the continental eastern areas. The reason is that the interannual variations are smaller over the ocean than over land, and the "noice level" thus is lower.

For the entire 100 year period, only the western regions show a significant temperature increase (fig. 11). In the SW region, the increase is significant on the 5% level. In the NW and CW regions, the increase is also significant at the 1% level. However, the full drawn curves in figure 11 show that in all regions, there has been a period when the temperature level was significantly above the level in the beginning of the series, at least at the 5% level. In the western regions this period lasts from the 1930s to the ends of the series. In the eastern regions, the period is shorter. In NE it lasts from the 1930s to the 1970s, in CE from the 1930s to the 1950s, and in SE the 5% significance level is exceeded only in one year (1953).

Starting from 1990 and going backwards in time (dotted curves in figure 11), one may similarly state that the NW, CW and NE regions experienced a significant drop in temperature from the 1930s to 1990, while this is not the case in the CE region and the southern regions. The dotted curves in figure 11 also show positive temperature trends from the 1960s to 1990 in the CW and SW regions, and from 1979 to 1990 in the CW and CE regions. Preliminary analyses including data from 1991-1995 indicate significant positive trends towards the end of all the eastern regional curves.

3.2.3 Pettitt test results.

The Pettitt test (see section 3.1) was applied on the regional temperature series in order to identify "change points" between different thermal regimes. In the present analysis, a change is said to occur if the Pettitt test statistic (eq.5) exceeds the 5% level. The change point is the year in which the test statistic achieves its extreme value. Usually, climate is changing gradually during several years. Accordingly a change point may be intepreted as the central year in a period with trend.

Figure 12 shows the test statistics for testing the whole series, and table 2a summarises these results. The Pettitt test statistics for all regions have their extreme values in connection with the temperature increase prior to the warm period starting in the 1930s. The NW and CW curves peak in 1925. The NE curve has its extreme value in 1919, but also peaks in 1928, almost simultanously with the CE curve, while the test statistics for the southern regions peak in 1931. Changepoints are identified in the western regions. The CE and SE extreme values on the other hand, are not significant at the 5% level, and the NE extreme value just barely passes this level. This is mainly in agreement with the Mann-Kendall test results, which imply significant trends from start to end in the western regions only.

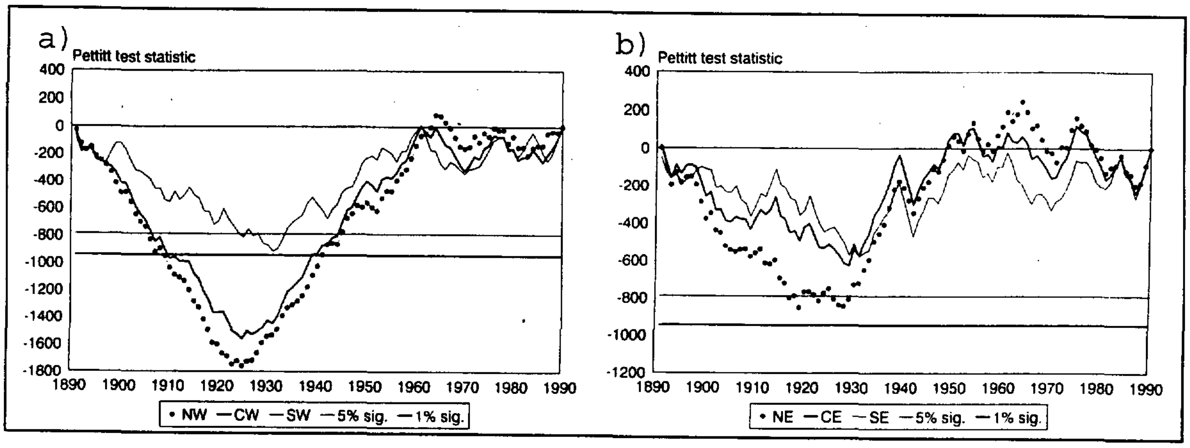


Figure 12. Pettitt test statistics for a) western and b) eastern regions for the period 1891-1990.

Table 2a. Year of extreme value for the Pettitt test statistics when testing the regional temperature series during the period 1891-1990. In case the extreme value is significant, at least at the 5% level, this year is a changepoint.

REGION	NW	CW	SW	NE	CE	SE
YEAR OF EXTREME VALUE	1925	1925	1931	1919	1929	1931
SIGNIFICANT? (EVT. LEVEL)	1%	1%	5%	5%	no	no

Table 2b. Changepoints according to Pettitt's test of parts of the regional temperature series.

REGION	NW	NW	CW	CW	SW	SW	NE	NE
2. TEST: PERIOD TESTED	1891-1925	1926-1990	1891-1925	1926-1990	1891-1931	1932-1990	1891-1919	1920-1990
CHANGEPOINT	no	1964	no	1961	no	no	no	no
3. TEST: PERIOD TESTED	1926-1964	1965-1990	1926-1961	1962-1990	X	X	X	X
CHANGEPOINT	no	no	no	no	X	X	X	X

Whenever a change point was identified, new tests were run on the periods before and after this point, respectively. Results from these tests are presented in table 2b. In the NW and CW regions, secondary change points were found in connection with the cooling from the warm period 1930-1960 to the cold period later in the 1960s. Again, the test results are mainly in accordance with the results from the Mann-Kendall tests, which imply that there is a negative trend from around 1930 to 1990 in both these series.

4. Concluding remarks.

Principal component analysis was applied on annual mean temperatures from 58 NACD stations during the period 1891-1990, primarily to extract the main features in the NACD temperature series. Statistical analyses were used in order to identify statistically significant changes. Following conclusions may be drawn.

* Principal component analysis is a useful tool for describing the temperature variations in the area. Five components are needed to account for 95% of the variance in the dataset. This percentage is at minimum in southwestern parts of Ireland, where about 70% of the variance is accounted for, and at maximum in mid-Finland, where 98% is accounted for.

* PCA based estimates of temperature series for single points are generally in good accordance with observed temperature curves, even in southwestern Ireland. An estimated series for the isolated weathership M is not quite satisfactory, but even here, the main features of the observed temperature trends are recognized in the estimated curve. Temperature series estimated from PCA may be used as reference series when investigating environmental effects like urban warming.

* Regions of common temperature variation patterns may be defined based upon the percentage of the variance accounted for by the first PCs. In the present analysis, 6 regions were defined.

* Estimated temperature series representative for the 6 regions show the following main features:

- The annual mean temperatures during the period 1891-1920 fluctuated around a relatively stable average.

- During the period 1920-1935, the average temperature increased by at least 1°C in the northern regions, above 0.5°C in central regions and less than 0.5°C in southern regions. The increase was statistically significant in western regions, where the "noise" from interannual variation is at minimum, and in northern regions, where the "signal" was at maximum. In the southeastern region, the question of statistical significance of the trend is highly dependent on the choices of start and end points of the period investigated.

- In western areas, the annual mean temperatures fluctuated around this higher average until 1950 in the southern region, and even to 1960 further north. In eastern areas, the generally warm period was interrupted by the very cold years 1940, 1941 and 1942. In the southeastern region, the temperature level from the 1930s was regained later in the 1940s. Further north, the temperatures did not quite reach this level again.

- In all regions, there is apparently a negative trend in the annual mean temperatures from this high level to cold periods in the 1960s and/or around 1980. The average temperature level for the period 1960-1980 is increased by some warm years during the 1970s. Again the trend is statistically significant in northern and western regions, while its significance further south depends on the period in consideration.

- All regions tend to show positive temperature trends from the 1960s and/or the late 1970s to the ends of the series. However, it is not statistically significant in the northern regions up to 1990. In the southern regions, the temperature level towards the ends of the series is about the same as it was during the warm period in the 1930s. In central regions, this maximum level was not quite reached in 1990, but preliminary analyses indicates that it is reached when 1991-1995 data are included. In the northern regions, the temperature level is still below the level of the 1930s.

- The temperature level towards the ends of the series is about 0.5°C above the level in the beginning of the series for all regions, however this temperature increase is statistically significant in the western regions only.

- The regional temperature series show generally the same periods of warming and cooling as the series of mean temperatures for the northern hemisphere (Jones, 1994), though the interannual variation generally is larger in the regional curves. The main discrepancies is that, in the hemispheric curve, the warming during the last decades and the warming during the 1920s are of the same size. The temperature level around 1990 for the hemisphere is thus clearly the highest during the 100 year period, while the temperature levels in the analysed regions are still below (northern regions) or about equal to (southern regions) the level in the middle of the series.

Acknowledgements

This study was partly founded by the Norwegian Research Council through the Environmental Programme of the European Commission (Contr. EV5V-CT93-0277) and by the Nordic Environmental Research Programme NMR (Contr.FS/ULF/93001). Data used in the analyses were kindly delivered by H. Alexandersson, SMHI, J. Ashcroft, UKMO, A. van Engelen, KNMI, P. Frich, DMI, L. Keegan, IRMET, T. Jónsson, VI, and H. Tuomenvirta, FMI.

References

Demarée, G.R. 1990: Did an abrupt global climate warming occur in the 1920 ? *Bijdragen tot de studie van de klimaatveranderingen. Publ. Série*, **124**, Royal Meteorological Institute, Brussels, Belgium. 32-37.

Frich, P., Alexandersson, H., J.Ashcroft, B.Dahlström, G.R.Demarée, A.Drebs, A.F.V. van Engelen, E.J.Førland, I.Hanssen-Bauer, R.Heino, T.Jónsson, K.Jonasson, L. Keegan, P.Ø.Nordli, T.Schmidt, P.Steffensen, H.Tuomenvirta, O.E.Tveito, 1996: "North Atlantic Climatological Dataset (NACD Version 1) -Final Report", Danish Meteorological Institute, *Scientific Report 96-1*, 47 p.

Hanssen-Bauer, I. and E.J. Førland, 1994: Homogenizing long Norwegian precipitation series. *J. Climate,* Vol.7, No.6, 1001-1013.

Hanssen-Bauer, I., E.J. Førland and O.E. Tveito, 1995: Trends and variability in annual precipitation in Norway. *DNMI klima,* 27/95, 25 p.

Heino, R., 1994: Climate in Finland during the period of meteorological observations. *Finnish Meteorological Institute contributions*, **12**, Finnish Meteorol. Inst., Helsinki, Finland, 209 p.

Jones, P.D.,1994: Hemispheric surface air temperature variations: A reanalysis and update to 1993. *Journal of Climate*, 7, 1794-1802.

Mills, G.F., 1995: Principal Component Analysis of Precipitation and Rainfall Regionalization in Spain. *Theor.Appl. Climatol.*, **50**, 169-183.

Nordli, P.Ø., H. Alexandersson, P. Frich, E. Førland, R. Heino, T. Jónsson, P. Steffensen, H. Tuomenvirta, O.E. Tveito, 1996: The effect of radiation screens on Nordic temperature measurements. *DNMI klima,* in prep.

Preisendorfer, R.W, 1988: *Principal Component Analysis in Meteorology and Oceanography*, Elsevier, 425 p.

Sneyers, R. 1990. On statistical analysis of series of observation. *Technical note* No. **143**, WMO No. 415. WMO, Geneva, Switzerland. 192 p.

Sneyers, R. 1995. Climate instability determination. Discussion of methods and examples. *Proc. from: 6th International Meeting on Statistical Climatology*. 19-23 June, 1995, Galway, Ireland. 547-550.

APPENDIX A - 1

Relevant information about the dataset.

Coloumn 1: National station number
Coloumn 2: Country code
Coloumn 3: Official name of station
Coloumn 4: Latitude, degrees and minutes N
Coloumn 5: Longitude, degrees and minutes E or W
Coloumn 6: Height above mean sea level in meters
Coloumn 7: Quality code; H = Homogenous, rigourously tested and maybe adjusted
T = Tested, maybe adjusted but not perfectly homogenous
N = Not tested, but not necessarily inhomogenous
E = Environmental changes prevents climatic change studies
I = Inhomogeneous
- = Not NACD station.
Coloumn 8: Application in present analysis.
A = Used in the PCA
* = Estimated from the PCA

01026	N	TROMSØ	69 39 N	18 56 E	100	T	A
01065	N	KARASJOK	69 28 N	25 31 E	129	H	A
01098	N	VARDØ	70 22 N	31 05 E	14	H	A
01100	N	SHIP "M"	66 00 N	2 00 E	0	N	*
01152	N	GLOMFJORD	66 49 N	13 59 E	39	T	A
01212	N	ONA	62 52 N	6 32 E	13	T	A
01235	N	KJØREMSGRENDI	62 06 N	9 03 E	626	H	A
01355	N	LÆRDAL	61 04 N	7 31 E	36	T	A
01403	N	UTSIRA FYR	59 18 N	4 53 E	55	H	A
01448	N	OKSØY FYR	58 04 N	8 03 E	9	H	A
01482	N	FERDER FYR	59 02 N	10 32 E	6	H	A
01492	N	OSLO-BLINDERN	59 57 N	10 43 E	94	E	*
05223	S	FALSTERBO	55 23 N	12 49 E	5	H	A
06452	S	VÄXJÖ	56 52 N	14 48 E	166	T	A
06641	S	KALMAR	56 43 N	16 17 E	15	T	A
06855	S	HOBURG	56 55 N	18 08 E	38	T	A
07243	S	GØTEBORG	57 46 N	11 53 E	20	T	A
07245	S	BORÅS	57 46 N	12 56 E	135	T	A
07839	S	VISBY	57 40 N	18 20 E	42	H	A
08924	S	GOTSKA SANDØN	58 23 N	19 11 E	12	T	A
09322	S	KARLSTAD	59 21 N	13 28 E	46	T	A
09752	S	UPPSALA	59 51 N	17 37 E	13	T	A
10537	S	FALUN	60 37 N	15 37 E	160	T	A
12402	S	SVEG	62 01 N	14 21 E	360	T	A
12738	S	HÄRNÖSAND	62 37 N	17 56 E	8	T	A
13411	S	ÖSTERSUND	63 11 N	14 29 E	376	T	A
14036	S	HOLMÖGADD	63 35 N	20 45 E	6	H	A
15772	S	STENSELE	65 04 N	17 09 E	325	H	A
16179	S	PITEÅ	65 19 N	21 29 E	6	T	A
16395	S	HAPARANDA	65 49 N	24 08 E	5	T	A
16798	S	KVIKKJOKK	66 57 N	17 44 E	337	T	A
16988	S	JOKKMOKK	66 37 N	19 38 E	260	T	A

APPENDIX A - 2

Relevant information about the dataset - continued.

Coloumn 1: National station number
Coloumn 2: Country code
Coloumn 3: Official name of station
Coloumn 4: Latitude, degrees and minutes N
Coloumn 5: Longitude, degrees and minutes E or W
Coloumn 6: Height above mean sea level in meters
Coloumn 7: Quality code; H = Homogenous, rigourously tested and maybe adjusted
T = Tested, maybe adjusted but not perfectly homogenous
N = Not tested, but not necessarily inhomogenous
E = Environmental changes prevents climatic change studies
I = Inhomogeneous
\- = Not NACD station.
Coloumn 8: Application in present analysis.
A = Used in the PCA
* = Estimated from the PCA

18880	S	ABISKO	68 21 N	18 49 E	388	H	*
19283	S	KARESUANDO	68 26 N	22 29 E	327	T	A
00304	FIN	HELSINKI	60 10 N	24 57 E	4	E	*
01101	FIN	TURKU	60 31 N	22 16 E	51	H	A
01202	FIN	TAMPERE	61 28 N	23 44 E	85	T	A
01701	FIN	LAPPEENRANTA	61 05 N	28 09 E	105	T	A
02425	FIN	JYVASKYLA	62 12 N	25 43 E	137	H	A
03602	FIN	KUOPIO	62 54 N	27 41 E	119	T	A
04601	FIN	KAJAANI	64 17 N	27 40 E	132	H	A
05404	FIN	OULU	65 02 N	25 29 E	13	T	A
00044	GB	LERWICK	60 08 N	01 11 W	82	H	A
00293	GB	WICK	58 27 N	03 05 W	36	H	A
00425	GB	STORNOWAY	58 13 N	06 19 W	15	T	A
01215	GB	BRAEMAR	57 00 N	03 24 W	339	T	A
01646	GB	EDINBURGH	55 55 N	03 11 W	134	H	*
06641	GB	DUMFRIES	55 04 N	03 36 W	49	T	A
03952	IRL	ROCHES POINT	51 48 N	08 15 W	40	N	A
03953	IRL	VALENTIA OBS.	51 56 N	10 15 W	9	N	*
03965	IRL	BIRR	53 05 N	07 53 W	70	N	A
04013	IS	STYKKISHOLMUR	65 05 N	22 44 W	8	H	A
04092	IS	TEIGARHORN	64 18 N	15 12 W	14	H	A
06011	FR	TORSHAVN	62 01 N	06 46 W	43	H	A
06193	DK	HAMMERODDE FYR	55 18 N	14 47 E	11	H	A
21100	DK	VESTERVIG	56 46 N	08 19 E	18	H	A
25140	DK	NORDBY	55 26 N	08 24 E	5	H	A
06260	NL	DE BILT	52 06 N	05 11 E	2	H	A
06447	B	UCCLE	50 48 N	04 21 E	100	H	A
00001	FIN	MARIEHAMN	60 07 N	19 54 E	4	-	A
03001	FIN	VAASA	63 03 N	21 46 E	4	-	A
09602	FIN	INARI	69 04 N	27 07 E	152	-	*
02165	GB	DURHAM	54 46 N	1 35 W	102	-	A
04522	GB	OXFORD	51 46 N	1 16 W	63	-	A
08811	GB	PLYMOUTH	50 21 N	4 07 W	27	-	A

APPENDIX B

Interpolated loadings (weight functions) at the verification stations.

STATION NUMBER AND NAME	PC1	PC2	PC3	PC4	PC5
01100 SHIP "M"	0.068	0.06	0.21	0.00	0.15
01492 OSLO-BLINDERN	0.135	-0.13	0.01	-0.11	0.00
18880 ABISKO	0.142	0.22	0.11	0.00	-0.20
00304 HELSINKI	0.165	-0.03	-0.16	0.10	0.16
09602 INARI	0.180	0.28	0.05	-0.12	0.20
01646 EDINBURGH	0.058	-0.12	0.16	0.08	-0.07
03953 VALENTIA OBSERVATORY	0.030	-0.09	0.18	0.10	-0.12

Interpolated loadings (weight functions) for temperature regions.

REGION	PC1	PC2	PC3	PC4	PC5
NORTHWESTERN REGION	0.040	0.12	0.36	0.00	0.35
CENTRAL WESTERN REGION	0.045	0.00	0.25	0.00	0.10
SOUTHWESTERN REGION	0.050	-0.10	0.15	0.15	-0.12
NORTHEASTERN REGION	0.140	0.18	0.10	0.00	0.00
CENTRAL EASTERN REGION	0.160	0.00	0.00	-0.15	0.00
SOUTHEASTERN REGION	0.110	-0.18	0.00	0.10	0.00