Statistical hindcast of wind climatology in the North Atlantic and northwestern European region*

Eigil Kaas**, Tian-Shi Li, Torben Schmith

Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen, Denmark

ABSTRACT: Monthly means of sea level air pressure (SLP) and sea surface temperature (SST) fields are used as predictors to reconstruct an approximate climatology of wind and cyclonic activity, valid for 10 stations, in winter over the northern North Atlantic region throughout the 20th century. The reconstructed climatology is based on statistical relationships between these predictors and local observations at 10 synoptic stations during the period 1961 to 1987 The statistical relationships are identified by canonical correlation analysis combined with a cross validation technique. We apply double cross validation to independently verify the statistical model. An interpretation of the relationship between atmospheric circulation patterns and anomalies in the wind climatology and cyclonic activity is given. The historical climatology (1903 to 1987) for each observing station is obtained from the observed large-scale conditions by using the statistical relationships. It is found that within the present century both the monthly mean windiness and cyclonic activity show variations on the decadal time scale but no significant overall trends. This is confirmed for 2 of the stations for which directly observed pressure data were available since the beginning of the century.

KEY WORDS: Wind · Cyclones · North Atlantic Hindcast · Downscaling Storminess

1. INTRODUCTION

Variations in strength and frequency of storms and ocean waves in the northern North Atlantic region during this century have attracted increasing attention from the oil industry, commercial navigation interests, fisheries, insurance companies and the public, since severe property damages caused by waves and storms have been reported from time to time (e.g. Berz & Conrad 1994). Several scientific investigations have indicated that the strength and frequency of storms (Schinke 1993, Stein & Hense 1994) and waves (Bacon & Carter 1991) have been enhanced in recent years. These investigations, however, have been criticised by, e.g., Schmidt & von Storch (1993) as being based on inhomogeneous records. Schmidt & von Storch argue that the increase in number of merchant ships and the introduction of ocean weather ships, satellites and numerical weather prognoses gradually have improved the quality of weather maps, which therefore may reflect spurious trends in, e.g., the number of lows below a certain threshold. The same criticism applies to the available gridded datasets of daily values, since they essentially are digitised weather maps. To state the problem more clearly: gridded datasets probably represent the mean features correctly, but the day-to-day variability becomes more and more underestimated the further one goes backward in time.

In order to solve the inhomogeneity problems, Schmidt & von Storch (1993) based their analysis of storminess in the German Bight on the calculation of geostrophic winds — as a proxy for real wind — from 3 nearby stations forming a triangle, and they found no overall trend in either mean or extreme wind conditions during a 100 yr period.

Another approach, which will be followed here, is to perform a downscaling of large-scale atmospheric flow. Basing our approach on certain observed large-scale climatological variables and applying a statistical

model based on canonical correlation analysis (CCA), we intend to reconstruct the climatology of wind and cyclonic activity at 10 locations in Scandinavia and the Northeast Atlantic region during this century.

The word 'downscaling' was used by von Storch et al. (1993) to mean a statistical process in which relationships between local and large-scale weather/climate variables are sought. Assuming unchanged physical conditions, the parameters of these relationships are fixed and can be used to calculate the local/large-scale variables if the other variables are given. Usually it is the large-scale variables that are known and the local ones that are being sought. The methodology used in the present work is similar to that of von Storch et al. (1993), in which a linear statistical relationship between observed large-scale atmospheric/oceanic fields (in our case the monthly means of sea level air pressure, SLP, and sea surface temperature, SST), termed 'predictors', and observed local climate variables (in our case windiness), termed 'predictands', is identified in a training period where both predictors and predictands are observed and believed to be homogeneous. Assuming the large-scale SLP and SST fields to be homogeneous prior to the training period, the derived statistical relationship can be used to extend backward in time — 'hindcast' — a homogeneous time series of predictands.

As a measure of 'windiness' we use the monthly mean of the squares of wind speeds observed twice a day (W^2). This quantity is obviously biased towards the strongest winds in the month. Furthermore, the surface drag of the wind is approximately proportional to the wind speed squared, meaning that W^2 probably is a better descriptor of damage caused by the wind in a given month than the arithmetic mean of wind speeds. A final and somewhat surprising argument for choosing W^2 is that we found that the downscaling model performs better on wind speed squared than on wind.

In addition to W^2, high-frequency variability of the mean SLP, a traditionally used measure of cyclonic activity, is also employed in the present work. One method commonly used is to apply a medium or high pass filter to the raw pressure data, and then to study the low-frequency variability of the high pass filtered data (e.g Nakamura & Wallace 1990). Here we use a very simple high pass filter, namely the pressure tendency over 12 h, and monthly means (PT) of the absolute values of these tendencies are used as a measure of cyclonic activity. The relationship between cyclonic activity and pressure tendency has previously been studied by e.g. Sanders & Gyakum (1980), Rogers & Bosart (1985), Gyakum et al. (1989) and Li (1990) In addition, synoptic experience tells us that both strong winds and large pressure changes at mid-latitudes usually are associated with developing and/or moving pressure systems and thus to a wide extent are interconnected, even though the timing and geographical positions of the two may be different.

Our basic assumption in the present study is that the average windiness in a given month at a given location is determined to a certain degree by large-scale atmospheric conditions. In the actual application we represent these conditions by gridded monthly means of SLP and SST By including both SLP and SST (which we take as a crude approximation of the air temperature over the sea in the lower troposphere) as predictors, we hope to represent both the large-scale barotropic and large-scale baroclinic conditions of the atmosphere. See also Section 6 below for a discussion of pressure tendencies and windiness.

Only winter months (DJF) are considered in the present work since this is the season when most storms occur.

The paper is organised as follows: Section 2 describes the data used while Section 3 mainly addresses the methodology used in model development, validation and assessment. In Section 4 we take a look at the identified flow patterns associated with high W^2 and PT values. In Section 5, hindcasted time series of W^2 and PT are presented and discussed. Section 6 examines the relation between W^2 and PT Finally, Section 7 discusses several key aspects of the methodology.

2. DATA

The predictands used in this study, W^2 and PT, are derived from observations at 10 synoptic stations in Denmark, The Faeroe Islands, Finland, Greenland, Norway and Sweden (Fig. 1, Table 1). These quantities cover the 81 winter months (DJF) from 1961 to 1987

The predictors were obtained from available datasets. The SLP data are gridded monthly averages of mean sea level air pressure fields obtained from the National Center of Atmospheric Research. These data are available on a $5° \times 5°$ grid and are used here on the geographical domain 65° W–45° E and 45° N–75° N (Fig. 1). The SST data are from the GISST dataset, version 2.2 (Rayner et al. 1995). This dataset is used here on a $5° \times 5°$ grid (although it is available at a higher resolution) and the domain is 65.5° W–44.5° E and 44.5° N–74.5° N, which makes it fairly compatible with the SLP dataset. Obviously points on land are of no value in this dataset; in addition, points covered by sea ice for more than one tenth of the winter period are disregarded. Both predictor fields are available over the period 1903 to 1987
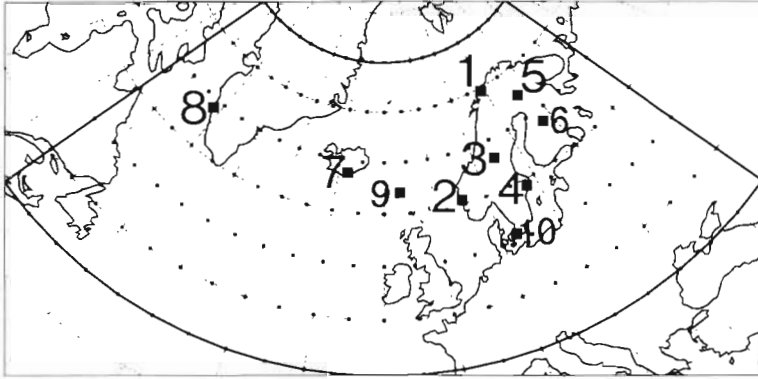
Fig. 1. Location and number of the 10 synoptic stations used in this study (large squares) and all the grid points representing the large-scale flow fields (small dots). See Table 1 for station names

3. METHODOLOGY AND DEVELOPMENT OF THE DOWNSCALING MODEL

Before any statistical analysis is performed, the yearly cycle is removed from both the predictor and predictand fields. Thus only anomalies from the long-term average annual cycle are studied.

3.1. Prefiltering of data

Generally, meteorological data contain high inter-correlations, which often cause trouble when estimating the inverse matrices needed in statistical analyses, since the matrices may degenerate. To avoid such problems and to minimise noise, we performed a pre-filtering that is often used (e.g. Barnett & Preisendorfer 1987), by compressing both the predictor and predictand data into a limited number of principal components (PCs), i.e. a limited set of time series obtained as projections onto the dominant empirical orthogonal functions (EOFs).

Table 1 Station number, station name, WMO (World Meteorological Organisation) number and geographical location for the 10 stations used in the study

Station		WMO no.	Lat.	Long.
1	Tromsø	01025	69° 41' N	18° 55' E
2	Bergen	01317	60° 23' N	5° 20' E
3	Östersund	02226	63° 11' N	14° 30' E
4	Stockholm	02464	59° 21' N	17° 57' E
5	Sodankylä	02863	67° 22' N	26° 39' E
6	Kuopio	02917	63° 47' N	27° 48' E
7	Kirkjubajarklaustur	04064	63° 47' N	18° 04' W
8	Nuuk/Godthaab	04250	64° 10' N	51° 45' W
9	Torshavn	06011	62° 01' N	6° 46' W
10	Kastrup	06180	55° 37' N	12° 39' E

Choosing the number of PCs to retain is always a matter of compromise: with many PCs retained, much noise is introduced; with few PCs retained, there is a risk of disregarding valuable information. For the present work a very simple truncation based on the variance explained by the individual PCs was chosen. It was found that 5 PCs explain more than 90 % of the variance for both predictands and for the SLP predictor. For the SST predictor only 70 % of the variance is explained by 5 PCs, but adding more PCs increases the explained variance only very slowly, so in this case as well 5 PCs are retained. Table 2 shows the sum of the variance explained (in %) by PCs 1 to 5 for both the predictor and predictand data. Note that a total of 10 predictors are used, i.e. 5 SLP PCs plus 5 SST PCs.

3.2. Canonical correlation analysis

The downscaling model is based on optimised relationships between the predictor and predictand fields, or rather the PCs of these as described in Section 3.1.

Canonical correlation analysis (CCA) identifies the linear relations that give the best possible correlations between 2 sets of stochastic variables. The technique may thus be seen as an extension of multiple linear regression, allowing for 'more than one variable on both sides of the equal sign'. Hotelling (1936) was the first to introduce CCA to the scientific community. Thorough descriptions and reviews of CCA are given in Tatsuoka (1971) and Wilks (1995). Step-by-step instructions for carrying out the CCA technique can be found in Barnett & Preisendorfer (1987). More recent papers have shown different applications and developments of this technique (Barnston & Ropelewski 1992, von Storch et al. 1993, Cui et al. 1995).

Our analysis follows the method of Barnett & Preisendorfer (1987). Mathematically expressed, we are seeking linear combinations of the normalized (to unit standard deviation) predictor PCs (α_i):

Table 2. Sum of explained variance (in %) for PCs of SLP, SST, W^2 and PT

PC no.:	1	2	3	4	5	Sum
SLP	37.8	23.3	15.4	10.0	4.0	90.5
SST	28.1	13.6	12.4	10.9	5.2	70.2
W^2	38.7	27.0	12.7	8.7	6.4	93.5
PT	59.4	15.9	11.2	7.9	1.6	96.0

　　　　　　　　　　　　　　　　　　　　Clim Res 7: 97–110, 1996

$$u_j(t) = \sum_{i=1}^{10} \alpha_i(t) r_{ij} \qquad (1)$$

[the summation limit (10) reflects our choice of truncation in number of predictor PCs (5 SLP PCs, 5 SST PCs) and t is time, running through winter months, 1961–1987] and the normalised predictand PCs (β_i):

$$v_j(t) = \sum_{i=1}^{5} \beta_i(t) s_{ij} \qquad (2)$$

The u's and v's are called canonical component vectors (CCVs). The expansion coefficients r_{ij} and s_{ij} are determined such that (a) the so-called canonical correlations between the individual pairs (u_{jj}, v_{ij}) are maximised in such a way that the correlation decreases for increasing j, and (b) the canonical correlation between (u_j, v_k) is zero for $j \neq k$.

Spatial patterns associated with the CCVs are called canonical maps. They are defined for each location as temporal inner products between the (truncated) predictors or predictands and the CCV time series. As in Barnett & Preisendorfer (1987) they are called g-maps (for the predictors) and h-maps (for the predictands). The first g-map (associated with the most significant CCV) is plotted in Fig. 2 for both W^2 and PT. Having once calculated the g-maps and h-maps as well as the CCVs, the downscaling model for the W^2 predictands can be expressed as a sum of the different predictor CCVs times the corresponding canonical correlations times the corresponding W^2 h-maps:

$$\hat{W}^2{}_n(t) = \sum_{j=1}^{q''} \mu_j\, u_j(t) h_{j,n} \qquad (3)$$

where $\hat{W}^2{}_n$ indicates the hindcasted value at station n, μ_j is the canonical correlation between u_j and v_j, and $h_{j,n}$ is the jth h-map component at station n. A similar expression

$$\hat{PT}_n(t) = \sum_{j=1}^{q''} \mu_j\, u_j(t) h_{j,n} \qquad (4)$$

holds for hindcasting of PT at each station, however with different CCVs, g-maps and h-maps.

The q'' indicates the truncation (in terms of number of retained CCVs) of the downscaling model, which in our case must be
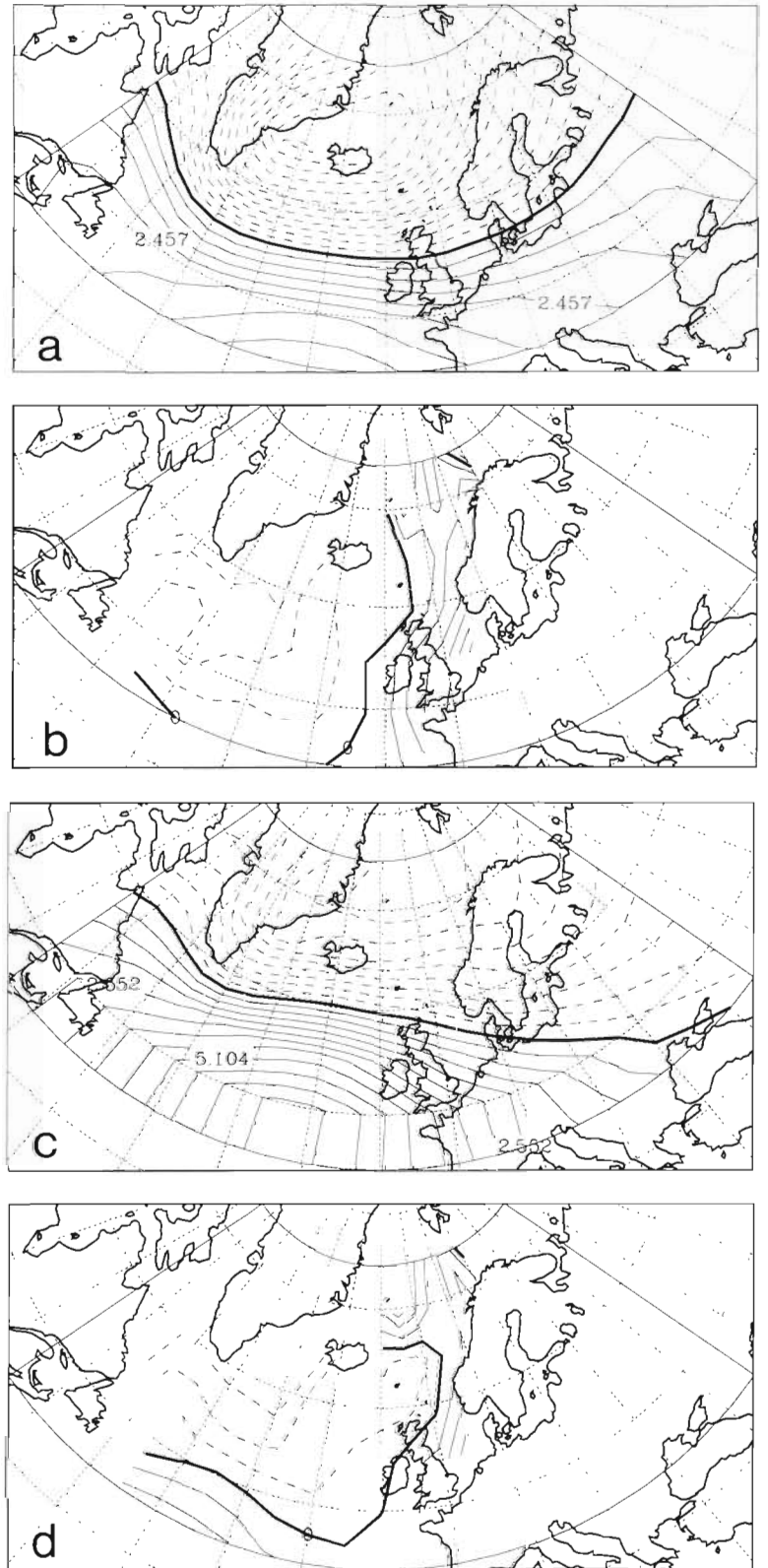

Fig. 2. First g-map. Panels (a) and (b) show the SLP part and the SST part, respectively, in the W^2 case, while panels (c) and (d) show the SLP part and the SST part, respectively, in the PT case. (Units are mb in (a) and (c), while in (b) and (d) they are °C)

less than or equal to 5. The optimum value of q'' can be determined in different ways. We used a 'cross validation' technique, in which each value $q'' = 1, ..., 5$ is tested by calculating a time series of hindcasted values (for 1961 to 1987) based on the actual q'' in the following way: The hindcast of the 3 winter months (DJF) in a given winter, I, involves deletion of winter $I - 1$ (1 yr before), I and $I + 1$ (1 yr after) from the record of predictors and predictands (special treatment is needed if I is the first or last winter in the data records). This procedure is to ensure that effects caused by autocorrelations in the predictands and/or predictors are eliminated (Barnston & Ropelewski 1992). Then a downscaling model is built entirely from the ground up (including EOFs and PCs followed by CCA) based on this shorter training period. This procedure is repeated for all winters within the period 1961 to 1987. In that way we obtain 5 different (one for each value of q'') time series of hindcasted values. The optimal q'' is finally chosen as the one whose corresponding time series gives the highest correlation coefficient with the observed values. Table 3 lists this optimal q'' for each station for the W^2 as well as the PT predictands.

Once the q'' is determined, the final downscaling model is set up based on all the winters in the training period 1961 to 1987. The hindcasts can then be obtained from Eq. (3) or (4) with the predictor CCVs $[u_j(t)]$ calculated in the period 1903 to 1987 as input.

3.3. Model validation

The maximum correlations used in the previous subsection to optimise q'' cannot be considered an independent measure of the skill of the final model, since all data (1961–1987) were used to find q''. To obtain an independent hindcast we resort to a 'double cross validation' (see e.g. Michaelsen 1987), which in our application works as follows: To hindcast the 3 winter months (DJF) in a given winter, J, the 3 winters $J - 1$, J

and $J + 1$ are deleted from the record of predictors and predictands. Then the entire procedure described in Sections 3.1 and 3.2 is repeated (including the cross validation optimisation) for the reduced dataset, resulting in a specific model with specific truncation q''_J, which can be used to perform an independent hindcast for winter J. Applying the above method to all winters from 1961 to 1987, a time series of independent hindcasts is obtained. The correlation between this independently hindcasted time series and the observed time series is our measure of model skill.

The model skill for each station is listed in Table 4. The most striking feature is that the skill for PT is considerably higher than for W^2 for all stations, meaning that hindcast of PT can be considered more reliable than that of W^2. We believe the explanation for this to be the inhomogeneity of the wind measurements during the training period 1961–1987. This is not an unlikely explanation. The station history files of almost any station report the relocation of an instrument site or the replacement of instruments by more modern ones. Such changes often take places at intervals of 10 yr or less. Furthermore, the surrounding canopy changes due to growth of vegetation, construction of houses and sheds, etc. In contrast, pressure — and therefore PT — is insensitive to these factors and can be regarded as fairly homogeneous.

Of course the skill differs from station to station, particularly for W^2 where Stns 7, 8 and 10 have such low skills that the hindcasts are disregarded in the following discussion.

4. FACTORS INFLUENCING MODEL SKILL

4.1. General discussion

Overall, the most important predictor patterns for obtaining the downscaled results are the first g-maps, which are shown in Fig. 2.

Table 3. Number of retained canonical modes (q'') for the hindcasts of W^2 and PT at the 10 stations

Station	W^2	PT
1	4	4
2	1	5
3	4	4
4	2	4
5	3	4
6	2	5
7	5	4
8	4	4
9	1	5
10	4	5

Table 4. Correlation coefficients from the double cross validation for both W^2 and PT

Station	W^2	PT
1	0.53	0.72
2	0.56	0.79
3	0.65	0.76
4	0.51	0.71
5	0.50	0.64
6	0.45	0.71
7	0.39	0.80
8	0.14	0.68
9	0.53	0.80
10	0.26	0.63

Beginning with the W^2 case, we note that the SLP part of the first g-map (Fig. 2a) is very similar to the North Atlantic Oscillation (NAO), which is the dominant anomaly pattern in the SLP fields in the North Atlantic region. Therefore it is also quite similar to the first SLP EOF (not shown).

The NAO has been recognized for many years and was first discovered by Walker & Bliss (1932) and further described by van Loon & Rogers (1978) and Wallace & Gutzler (1981), among others. The NAO was originally defined as the pressure difference between the Icelandic low and the mid-Atlantic subtropical high. As such it is a measure of the strength of the westerlies over the central North Atlantic region. Many studies have documented the anomalous weather associated with the 2 extremes of the NAO. The first dates as far back as the 1770s, when the Danish priest Hans Egede Saabye (1778) wrote: 'All winters in Greenland are severe, however, with differences. The Danes have noted that when the winter has been severe, as we call it, in Denmark, the Greenland winter has — in its own way — been mild and vice versa'. The temperature seesaw described here between northern Europe and (western) Greenland is related to the NAO (van Loon & Rogers 1978). This relationship between temperature anomalies and the NAO is also indicated in the SST part of the first g-map (Fig. 2b): cold water to the south of Greenland and warm water in the North Sea and the Norwegian Sea are associated with strong average westerlies around $55°–60°$ N. Due to the similarity between NAO and the SLP part of the first g-map we can conclude that not only temperature anomalies, but also anomalies in W^2 (the corresponding h-map, not shown), are associated with the NAO.

Turning to the PT case, the SLP part of the first g-map (Fig. 2c) is arguably more elongated in the east-west direction and displaced to the north, compared to the previous pattern. This is in agreement with Rogers (1995), who used monthly root mean squares of band-pass-filtered daily gridded SLP data — a quantity which is very similar to our PT— to estimate storm tracks. He found no 1:1 correspondence between variations in the dominant storm tracks and the NAO.

Rather, he found that the pattern related to strong storm activity in the Northeast Atlantic and northern Europe is more elongated in the east-west direction and displaced to the north and east, very similar to that in Fig. 2c.

The SST part of the first g-map for the PT case (Fig. 2d) shows features very similar to the W^2 case: cold water to the south of Greenland and warm water in the North Sea and the Norwegian Sea.

Finally, one should not put too much emphasis on the SST g-map patterns, since generally less than 5% of the explained variance in the hindcasts comes from SST PCs (see Table 5). Only 2 exceptions were found: namely Stns 8 and 10 in the W^2 case, where SST contributed 22.5 and 45.9%, respectively, to the total variance. But for these 2 cases the model's validation correlation values were very low (0.136, 0.262; see Table 4), and thus they should not be considered.

4.2. Optimal predictor maps

The g-maps shown in Fig. 2 are only the first and dominant g-maps, but more g-maps contribute to our downscaling model. When the downscaling model is built one can, for each individual station, construct a single predictor map which will give exactly the same predictions as those from Eqs. (3) & (4). These maps, which we call optimal predictor maps (OPMs), are quite informative when one wants to interpret the physics behind the downscaled results. A description of how OPMs are calculated is given in Appendix 1.

In Fig. 3 the W^2 OPMs are shown for Stns 2 and 3, which are the stations with highest skill. Note that only the SLP part of the OPMs is shown. For both stations high W^2 values are associated with SLP patterns, showing an anomalously strong monthly mean flow near the stations and an upstream area of anomalous confluence. It is likely that there is a sharpening of the temperature contrasts associated with this confluence, leading to an enhancement of the baroclinicity immediately upstream of the stations.

Turning to the PT case the same tendency is evident for most of the 10 OPMs (Fig. 4), i.e. a strong gradient

Table 5. Relative contribution (in %) of SLP and SST (in terms of explained variance) to the total variance of the hindcasted W^2 and PT for each of the 10 stations

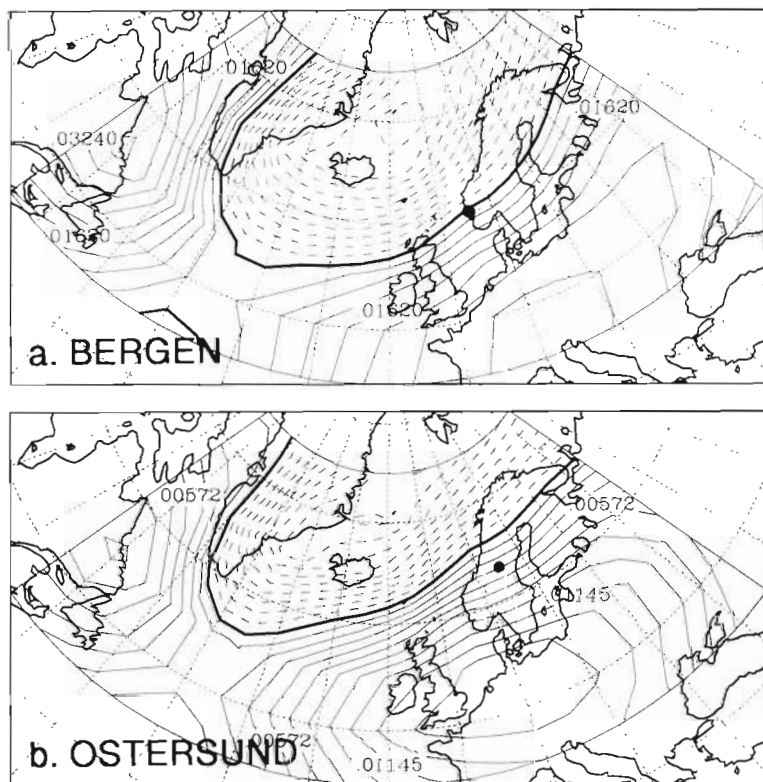|	Station:	1	2	3	4	5	6	7	8	9	10
W^2	SST	4.3	2.9	6 9	6.4	3.0	4.7	5.2	22.5	2.9	45.9
	SLP	95.7	97.1	93.1	93.6	97.0	95.3	94.8	77.5	97.1	54.1
PT	SST	3.0	1.9	5.5	6.3	4.4	6.1	0.2	4.9	0.3	4.3
	SLP	97.0	98.1	94.5	93.7	95.6	93.9	99.8	95.1	99.7	95.7

Fig. 3. Optimal predictor maps for (a) Stn 2 and (b) Stn 3 in the W^2 case. Geographical positions of the stations are marked by black dots. Units are $10^{-2}\,m^3\,kg^{-1}$

comparisons for the same station, since large pressure tendencies often occur upstream of and prior to the large winds.

5. HINDCAST RESULTS

5.1. Hindcast for W^2

The model was next used to hindcast the W^2 at the 10 synoptic stations. Fig. 5a–j show the hindcasted W^2 time series (1903 to 1987) and the 27 yr of observed W^2. For clarity both time series have been filtered by a 3 yr running mean before plotting. The long-term average values are also plotted for comparison. A visual inspection of these plots reveals no overall trends, but considerable variations on decadal time scale are seen.

We cannot rely on all the hindcasted time series in Fig. 5a–j with equal confidence, since the model's skill is different from station to station (Table 4). As already mentioned in Section 4.1, Stns 7, 8 and 10 should be disregarded for the W^2 case.

near the relevant station and an upstream anomalous confluence area. There is a high level of consistency between the maps in the sense that stations close to one another have very similar OPMs, e.g. Stns 1 and 5. From the maps for Stns 7 and 9 it is seen that high Central Atlantic cyclonic activity (i.e. high PT value) is related to NAO-like patterns — a fact which is also well known from synoptic experience and not inconsistent with Rogers' (1995) findings. For the western Greenland station strong cyclonic activity is associated with an anomalous ridge over the Central Atlantic. This is also consistent with synoptic experience indicating that 'many lows migrate up along the west coast of Greenland when the pressure over the Central Atlantic is high' (pers. comm. from forecasters at the Danish Meteorological Institute).

It is tempting to try to inter-compare OPMs for the W^2 and PT case at individual stations. For Stns 2 and 3, for example, the W^2 OPM shows a southwesterly anomaly flow at the station, while the PT OPM shows a westerly anomaly flow. This veering seems to be a general feature for the other stations as well, which is also reflected in the g-map patterns in Fig. 2a and 2c. The reason for this is not obvious, but one must bear in mind that it is not necessarily relevant to make inter-

5.2. Hindcast for PT

Fig. 6a–j is similar to Fig. 5a–j, but for the PT case. Again visual inspection reveals no overall trends but considerable variations on the decadal timescale. No large change is apparent in terms of the amplitudes of the variations, though a more regular variation is seen at all the 10 stations during the latest part of the series (1970s to 1987).

Since PT OPMs (Fig. 4) are very similar for stations close to one another (e.g. Stns 1 and 5), the PT hindcasts are also similar to each other for these stations. This can be clearly seen on the hindcast plots.

Just before finishing the manuscript, digitised long-term records of SLP at Stns 2 and 3, containing data collected 2 to 4 times daily and ranging back to approximately 1875, became available to us. From these observed pressure values we calculated the PT values for the entire period (1903 to 1987) and plotted them (Fig. 6b, i). In these plots it is apparent that the downscaled values are very similar to the observed data throughout the entire period, even though the downscaling model was only trained on data from the last 27 yr. The overall conclusion of no trends is thus confirmed by considering the directly observed data for these 2 stations.

b. BERGEN

d. STOCKHOLM

f. KUOPIO

a. TROMSO
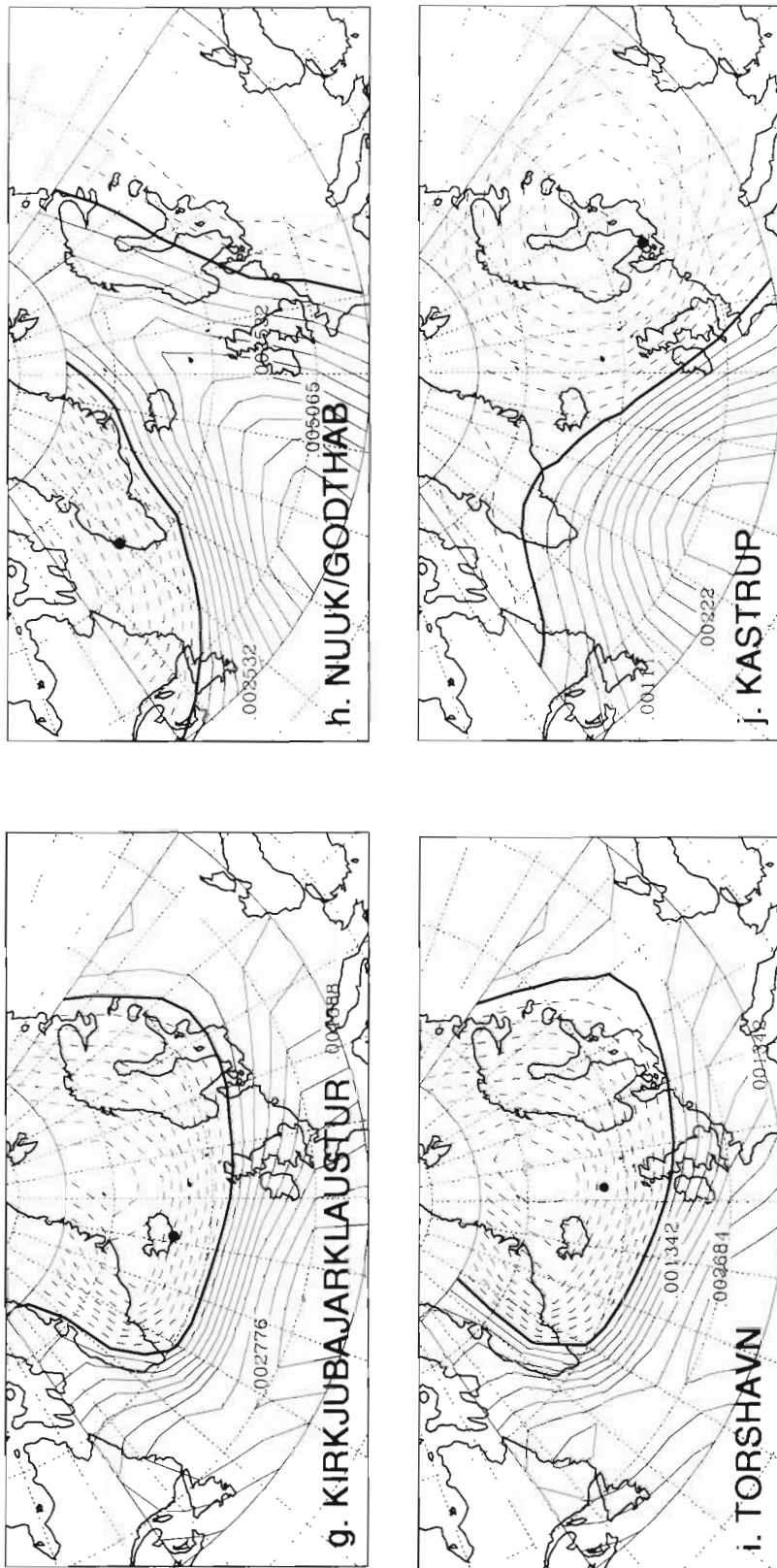
c. OSTERSUND

e. SODANKYLA
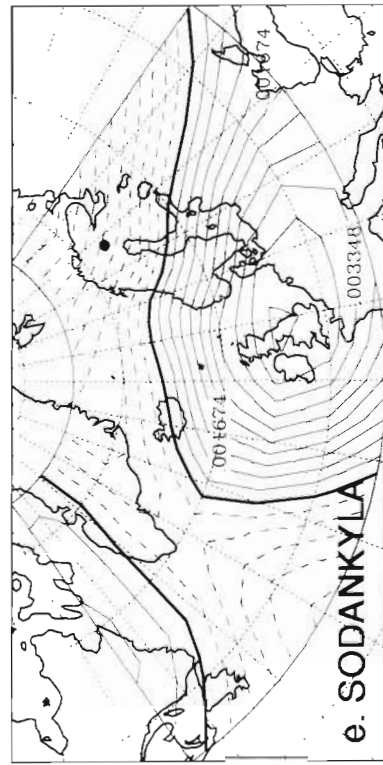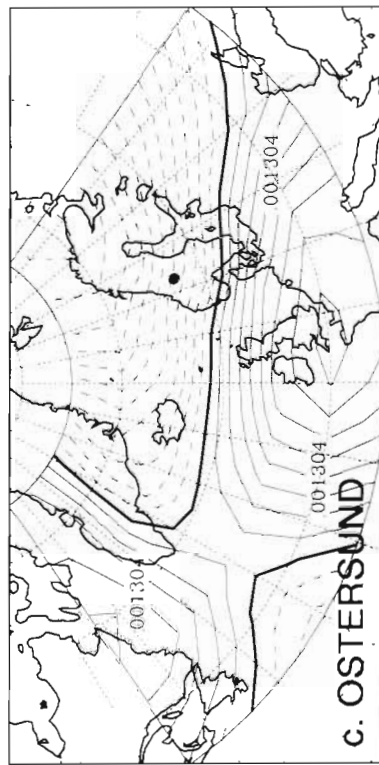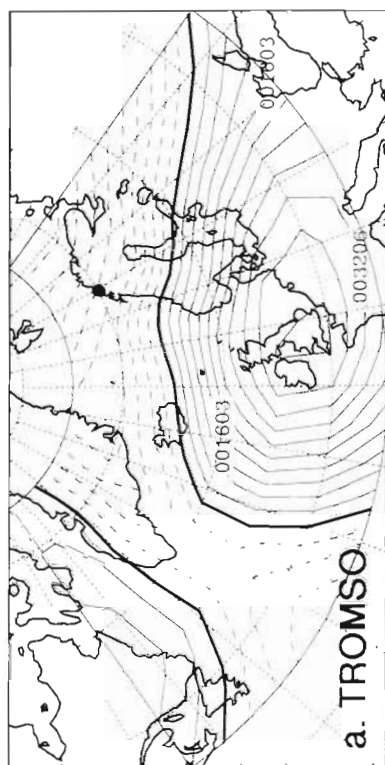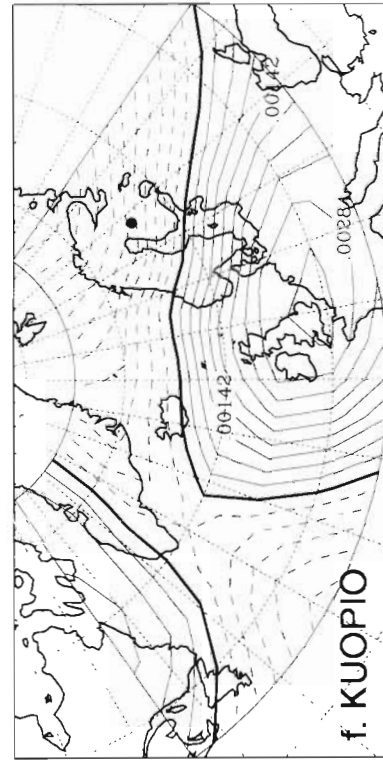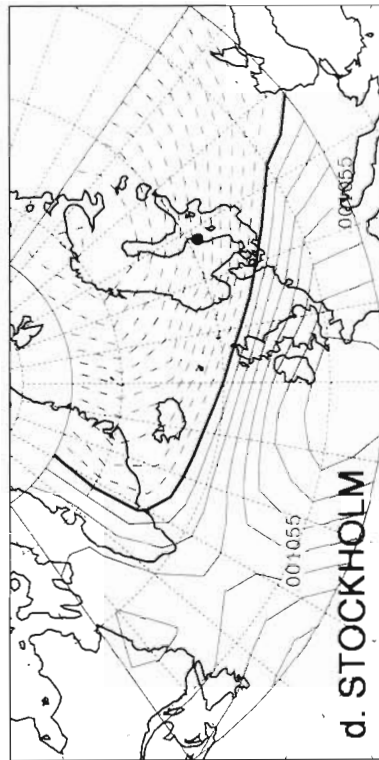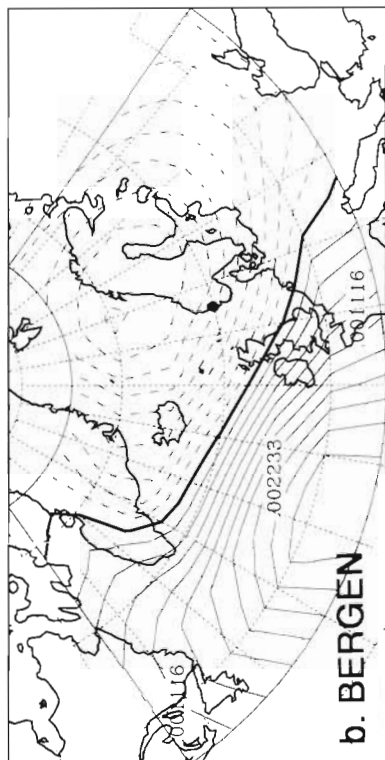
Fig. 4. Optimal predictor maps as Fig. 3 but for the *PT* case and for all 10 stations (a) through (j). Geographical positions of the stations are marked by black dots. Units are (12 h)$^{-1}$
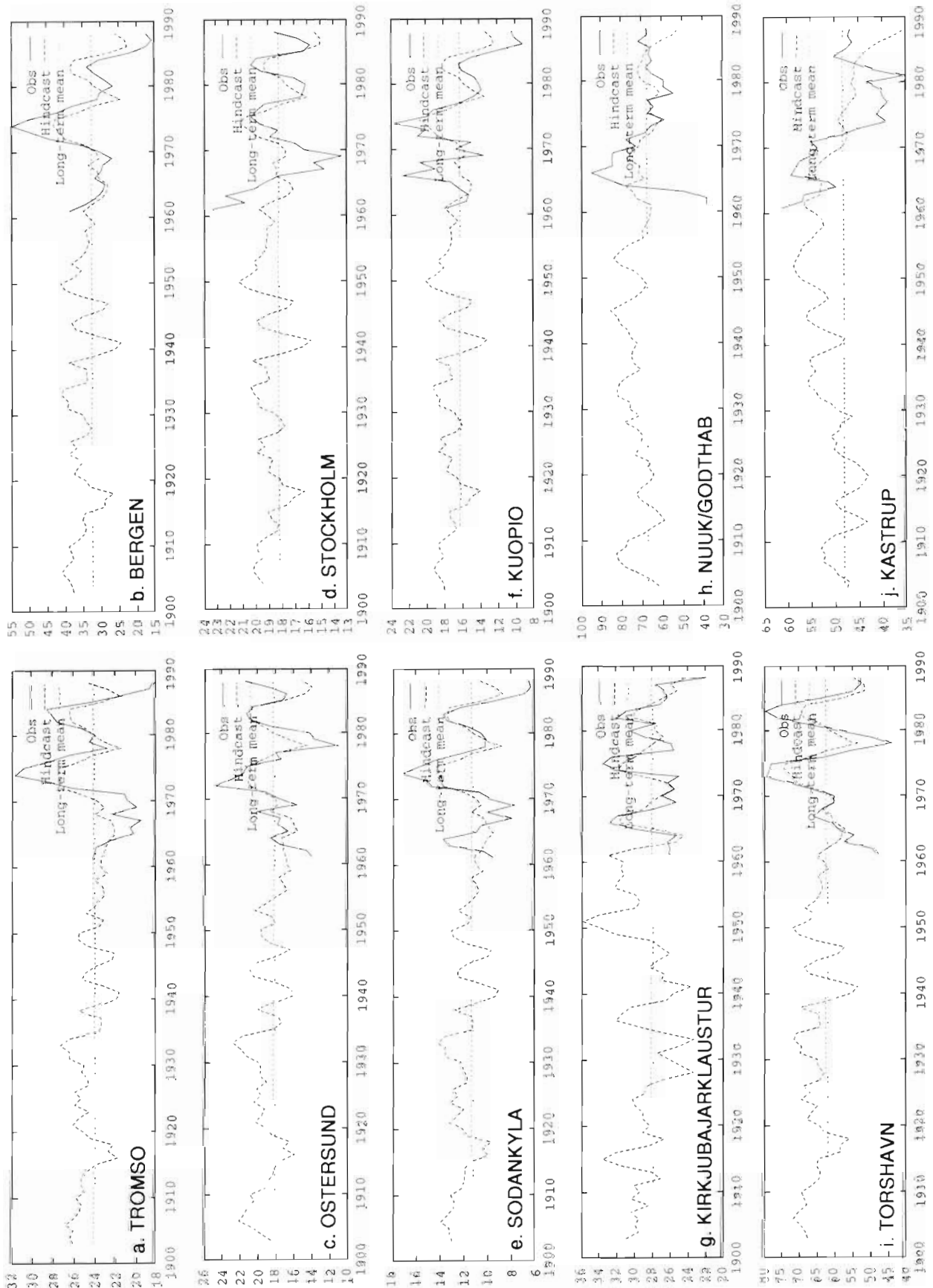
Clim Res 7: 97–110, 1996

Fig. 5. Three-year running mean of the hindcasted W^2 time series (1903 to 1987), three-year running mean of the observed W^2 time series (1961 to 1987) and long-term mean (1961 to 1987). Stns 1 through 10 are shown in (a) through (j) respectively. Units are m^2 s^{-2}
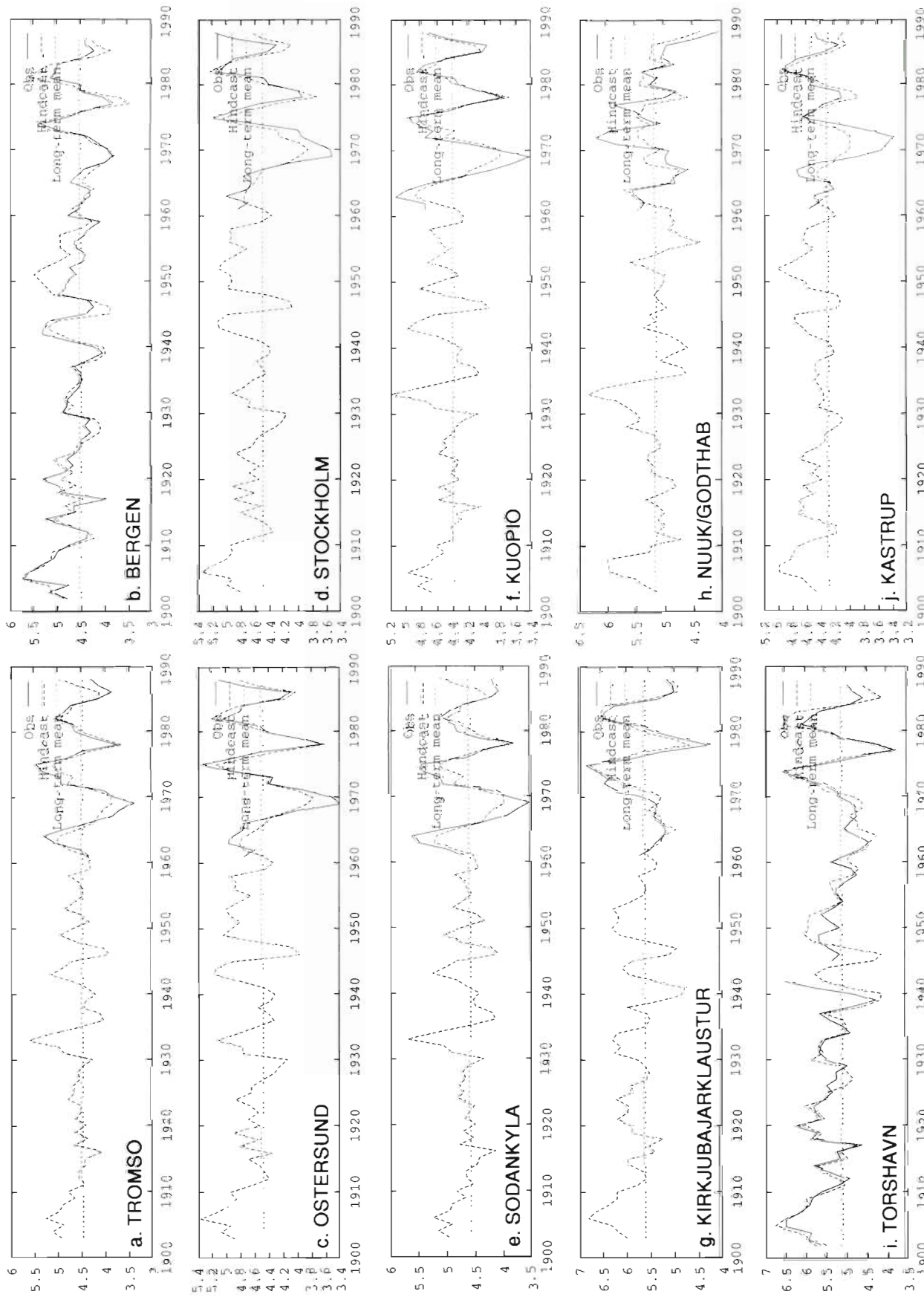
Fig. 6. Three-year running mean of the hindcasted *PT* time series (1903 to 1987), three-year running mean of the observed *PT* time series (1961 to 1987) and long-term mean (1961 to 1987). Stns 1 through 10 are shown in (a) through (j) respectively. For Stns 2 and 9 the observed time series covers the full period 1903 to 1987. Units are hPa per 12 h

6. RELATIONSHIP BETWEEN PRESSURE TENDENCY AND WIND SPEED

High local pressure tendencies indicate developing and/or moving strong synoptic disturbances. However, a high pressure tendency generally will not occur at the same location and time as high wind speeds. For this reason one cannot use station data to show a firm relationship between high winds and large pressure tendencies. Instead we chose to use gridded data from a 5 yr experiment with the ECHAM3 general circulation model run at T106 resolution (see Bengtsson et al. 1995). This dataset is, of course, homogeneous and contains a realistic number of storms in the North Atlantic, with core pressures down to approximately 930 hPa and pressure tendencies up to more than 40 hPa per 12 h. Maximum wind speeds over the ocean are generally between 25 and 35 m s^{-1}.

Twice a day the model near-surface wind speed and the absolute values of the 12 h pressure tendency were extracted at all sea points in the Northeast Atlantic from approximately 50° N to 70° N and 20° W to 15° E. To account for the varying Coriolis effect, the tendencies were scaled with by factor of $\sin(45°)/\sin(\varphi)$ where φ is the local latitude. Only extended winter seasons (NDJF) were considered, implying a total of 20 winter months. Next, 4 time series were calculated twice daily from these data:

Series 1: the number of grid points where the wind speed exceeded the local 95% percentile level;

Series 2: the average absolute 12 h pressure tendency in the area;

Series 3: the maximum wind speed occurring anywhere in the area;

Series 4: the maximum absolute 12 h pressure tendency occurring anywhere in the area.

To see whether strong winds were closely related to strong pressure tendencies, time series 3 and 4 were compared. A 5 d low pass filter was used to account partly for the fact that strong winds may be related to older, occluded systems (with small pressure tendencies) which originally were generated during a baroclinic instability process dominated by high pressure tendencies. The correlation between the filtered time series 3 and 4 was 0.78, which shows that short periods of a few days dominated by high extreme (in space) pressure tendencies are also dominated by high extreme winds, and vice versa.

In the present paper monthly mean pressure tendencies are used as a proxy for cyclonic activity. The correlation between the 20 monthly averages of time series 1 and 2 will give an estimate of the extent to which a monthly average pressure tendency is a good indicator of strong winds occurring during the

month. This correlation was found to be 0.80, which means that high/low monthly mean pressure tendencies imply a high likelihood of many/few strong winds.

Our station *PT* values only represent local average windiness, but the relationships identified above show that extreme winds are most likely to occur during months with high *PT* values. The strong winds may, however, occur remotely from the actual station.

7. DISCUSSION

There are several issues concerning the methodology we have used that should be discussed. Those which we find most important are listed below.

(1) *Physical processes captured by the downscaling model:* The linear downscaling model is by no means perfect, and monthly mean SLP and SST maps can never describe all aspects of W^2 and *PT*. One important aspect not covered by the model is the atmospheric static stability, which plays an important role in baroclinic instability (e.g. Eady 1949). This might have changed due to increased radiative forcing of the atmosphere, and thus may have influenced the e-folding time and size of the lows. The 2 observed *PT* time series at Stns 2 and 9 covering the years 1903 to 1987, however, bear no indication that this is the case.

(2) *Model skill:* A key question concerning the downscaling model is its skill. To obtain an estimate of the validity of the correlations obtained during the double cross validation, we correlated the independently observed *PT* time series at Stns 2 and 9 (see Section 5) with the *PT* hindcasts for the period 1903 to 1960 and obtained the correlations 0.74 for Stn 2 and 0.79 for Stn 9. (Note that the curves in Fig. 6b and i show 3 yr running mean data, and not the raw data on which the correlations are based.) These raw correlations have the same order of magnitude as those in Table 4. We can conclude that, at least for these 2 stations, double cross validation works as a fair method of calculating skill.

(3) *Choice of predictors:* One of our basic results is that the SSTs contribute very little to the hindcasted time series (see Table 5). Atmospheric near-surface temperatures would definitely have been a better choice for capturing elements of monthly mean baroclinicity. Recently a new dataset, the North Atlantic Climatologic Dataset (Frich et al. 1996), which covers a period of 100 yr, was compiled from a large number of stations in the North Atlantic region. These monthly mean data are quality-controlled and homogenised and cover a range of variables — including near-surface temperature — which may be used for down-

scaling purposes. The NACD dataset will be used in future applications of the downscaling model.

(4) *WASA pressure data set:* Pressure data collected 2 to 4 times daily at 20 North Atlantic stations, covering the years 1875 to 1995, are being digitised as part of the WASA project (Schmith et al. in press). An analysis of these data is now being carried out and both *PT* and geostrophic wind (based on nearby pressure data) time series will be studied in the WASA project.

Appendix 1. Calculation of optimal predictor maps

Considering the predictand W^2 (the equations are similar for the *PT* case) at a given station n, this appendix derives a single predictor map $e_{m,n}$, which includes 1 component from SLP as well as 1 component from SST, with m representing the number of gridpoints. The map $e_{m,n}$ is defined so that the same prediction $\hat{W}^2_n(t)$ as in Eq. (3) (see Section 3.2) can be obtained as a projection of SLP and SST fields onto $e_{m,n}$:

$$\hat{W}^2_n(t) = \sum_{m=1}^{161} SLP_m(t)e_{m,n} + \sum_{m=1}^{77} SST_m(t)e_{m+161,n} \quad (A1)$$

At the RHS of Eq. (A1) the first sum is the projection onto the SLP component of $e_{m,n}$ with 161 grid points, while the last sum is the projection onto the SST component of $e_{m,n}$, including only 77 grid points. We can rewrite Eq. (3) as follows:

$$\hat{W}^2_n(t) = \sum_{j=1}^{q'} \mu_j u_j(t)h_{j,n}$$

$$= \sum_{j=1}^{q'} \mu_j h_{j,n} \sum_{i=1}^{10} \alpha_i(t)r_{ij} \quad (A2)$$

$$= \sum_{i=1}^{10}\sum_{j=1}^{q'} \mu_j h_{j,n} r_{ij}\sigma_i^{-1/2} a_i(t)$$

where σ_i is the variance of the ith raw principal component, $a_i(t)$, and the remaining terms are the same as in Eq. (1) (see Section 3.2). Note that a_i with $i = 1, ..., 5$ is the ith non-

normalized SLP PC and a_i with $i = 6, ..., 10$ is the $(i-5)$th non-normalized SST PC. Using the definition

$$U_{i,n} \equiv \sum_{j=1}^{q'} \mu_j h_{j,n} r_{i,j}\sigma_i^{-1/2} \quad (A3)$$

Eq. (A2) can be written as

$$\hat{W}^2_n(t) = \sum_{i=1}^{5} U_{i,n} a_i(t) + \sum_{i=6}^{10} U_{i,n} a_i$$

$$= \sum_{m=1}^{161} SLP_m(t)\sum_{i=1}^{5} U_{i,n} EOF^{SLP}_{i,m}(t) + \quad (A4)$$

$$\sum_{m=1}^{77} SST_m(t)\sum_{i=6}^{10} U_{i,n} EOF^{SST}_{i-5,m}$$

where EOF^{SLP}_i is the ith SLP EOF and EOF^{SST}_i the ith SST EOF. From Eq. (A4) the predictor map $e_{m,n}$ can be written as:

$$e_{m,n} = \sum_{i=1}^{5} U_{i,n} EOF^{SLP}_{i,m} \quad (A5)$$

for $m \le 161$ and as

$$e_{m,n} = \sum_{i=6}^{10} U_{i,n} EOF^{SST}_{i-5,m-161} \quad (A6)$$

for $m \ge 162$.

LITERATURE CITED

Bacon S, Carter DJT (1991) Wave climate changes in the North Atlantic and North Sea. Int J Climatol 11:545–558

Barnett TP, Preisendorfer R (1987) Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. Mon Weather Rev 115:1825–1850

Barnston AG, Ropelewski CF (1992) Prediction of ENSO episodes using canonical correlation analysis. J Climate 5:1316–1345

Bengtsson L, Botzet M, Esch M (1995) Hurricane-type vortices in a general circulation model. Tellus 47A:175–196

Berz G, Conrad K (1994) Stormy weather: the mounting windstorm risk and consequences for the insurance industry. ECODECISION 12:65–69

Cui M, von Storch H, Zorita E (1995) Coastal sea level and the large-scale climate state: a downscaling exercise for the Japanese Islands. Tellus 47A:132–144

Eady ET (1949) Long waves and cyclone waves. Tellus 1:33–52

Eegede Saabye H (1778) Brudstykker af en dagbog holden i Groenland i aarene 1770–1778. In: Meddelelser om Groenland 129, no. 2, 1949. C. A. Reitzels Forlag, Copenhagen (in Danish)

Frich P, Alexandersson H, Ashcroft J, Dahlstrom B, Demaree GR, Drebs A, van Engelen AFV, Foerland EJ, Hanssen-Bauer I, Heino R, Jonsson T, Jonasson K, Keegan L, Nordli PO, Schmith T, Steffensen P, Tuomenvirta H, Tveito OE (1996) North Atlantic Climatological dataset (NACD Version 1) — final report. DMI Scientific Report 96-1. Danish Meteorological Institute, Copenhagen

Gyakum JR, Anderson JR, Grumm RH, Gruner EL (1989) North Pacific cold-season surface cyclone activity: 1975–1983. Mon Weather Rev 117:1141–1195

Hotelling H (1936) Relations between two sets of variates. Biometrika 28:321–377

Li TS (1990) Temporal variability of North Pacific Ocean surface cyclone activity. MSc thesis, Department of Meteorology, McGill University, Montreal

Michaelsen J (1987) Cross-validation in statistical climate
    forecast models. J Climate Appl Meteorol 26:1589–1600
Nakamura H, Wallace JM (1990) Observed changes in
    baroclinic wave activity during the life cycles of low-
    frequency circulation anomalies. J Atmos Sci 47:
    1100–1116
Rayner NA, Folland CK, Parker DE, Horton EB (1995) A new
    global sea-ice and sea surface temperature (GISST) data
    set for 1903–1994 for forcing climate models. Hadley Cen-
    tre Internal Note no. 69. Hadley Centre for Climate Pre-
    diction and Research, Bracknell, UK
Rogers E, Bosart LF (1985) An investigation of explosively
    deepening oceanic cyclones. Mon Weather Rev 114:
    702–718
Rogers JC (1995) The North Atlantic storm track and regional
    climate variability. Book of extended abstracts, 6th Inter-
    national Meeting on Statistical Climatology. The Steering
    Committee for International Meetings on Statistical Cli-
    matology, p 79–82
Sanders F, Gyakum JR (1980) Synoptic-dynamic climatology
    of the 'bomb'. Mon Weather Rev 108:1589–1606
Schinke H (1993) On the occurrence of deep cyclones over
    Europe and the North Atlantic in the period 1930-1991
    Beitr Phys Atmosph 41:223–237
Schmidt H, von Storch H (1993) German bight storms

analysed. Nature 365:791
Schmith T, Alexandersson H, Iden K, Tuomenvirta H (in
    press) Historical records of daily msl-pressure at 20 sta-
    tions in the Northeast Atlantic area (WASA pressure
    dataset version 1.0). DMI Scientific Report, Danish Meteo-
    rological Institute, Copenhagen
Stein O, Hense A (1994) A reconstructed time series of the
    number of extreme low pressure events since 1880. Mete-
    orol Z NF 3:43–45
Tatsuoka MM (1971) Multivariate analysis; techniques for
    educational and psychological research. Wiley, New York
van Loon H, Rogers JC (1978) The seesaw in winter tempera-
    tures between Greenland and Northern Europe. Part I:
    General description. Mon Weather Rev 106:296–310
von Storch H, Zorita E, Cubasch U (1993) Downscaling of
    global climate change estimates to regional scales: an
    application to Iberian rainfall in wintertime. J Climate 6:
    1161–1171
Walker GT, Bliss EW (1932) World weather V. Mem R Meteo-
    rol Soc 4:53–84
Wallace JM, Gutzler DS (1981) Teleconnections in the geopo-
    tential height field during the northern hemisphere win-
    ter. Mon Weather Rev 109:784–812
Wilks DS (1995) Statistical methods in the atmospheric sci-
    ences. Academic Press, San Diego