Climate Dynamics (1999) 15:685–702

T. J. Osborn · K. R. Briffa · S. F. B. Tett · P. D. Jones
R. M. Trigo

Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model

Received: 20 August 1998 / Accepted: 12 May 1999

Abstract The realism of the Hadley Centre's coupled climate model (HadCM2) is evaluated in terms of its simulation of the winter North Atlantic Oscillation (NAO), a major natural mode of the Northern Hemisphere atmosphere that is currently the subject of considerable scientific interest. During 1400 y of a control integration with present-day radiative forcing levels, HadCM2 exhibits a realistic NAO associated with spatial patterns of sea level pressure, synoptic activity, temperature and precipitation anomalies that are very similar to those observed. Spatially, the main model deficiency is that the simulated NAO has a teleconnection with the North Pacific that is stronger than observed. In a temporal sense the simulation is compatible with the observations if the recent observed trend (from low values in the 1960s to high values in the early 1990s) in the winter NAO index (the pressure difference between Gibraltar and Iceland) is ignored. This recent trend is, however, outside the range of variability simulated by the control integration of HadCM2, implying that either the model is deficient or that external forcing is responsible for the variation. It is shown, by analysing two ensembles, each of four HadCM2 integrations that were forced with historic and possible future changes in greenhouse gas and sulphate aerosol concentrations, that a small part of the recent observed variation may be a result of anthropogenic forcing. If so, then the HadCM2 experiments indicate that the anthropogenic effect should reverse early next century, weakening the winter pressure gradient between Gibraltar and Iceland. Even combining this anthropogenic forcing and internal variability cannot explain all of the recent observed variations, indicating either some model deficiency or that some other external forcing is partly responsible.

T. J. Osborn (✉) · K. R. Briffa · P. D. Jones · R. M. Trigo
Climatic Research Unit, University of East Anglia, Norwich, United Kingdom
E-mail: t.osborn@uea.ac.uk

S. F. B. Tett
Hadley Centre for Climate Prediction and Research, Meteorological Office, Bracknell, United Kingdom

1 Introduction

The North Atlantic Oscillation (NAO) is one of the major modes of intermonthly to interdecadal variability in the Northern Hemisphere atmosphere. First designated by Walker (1924), it has been extensively studied since the work of van Loon and Rogers (1978) and Barnston and Livezey (1987). The NAO has a strong influence, particularly in the Atlantic/European sector, on surface climate (Hurrell 1995, 1996), surface fluxes (Cayan 1992) and maybe even on the deep ocean (Dickson et al. 1996). Interest in the NAO has been renewed recently, particularly because of two contemporaneous trends observed in the *winters* of the last three decades: a trend towards the positive phase of the oscillation, and a trend towards warmer Northern Hemisphere land temperatures (Hurrell 1996).

This raises some important questions. How much of the recent winter warming is due to a natural climate variation (see also Wallace et al. 1995)? Is a global warming signal manifesting itself by amplifying a natural mode of the climate system? How unusual is the recent dominance and strength of the positive phase of the oscillation? Some of these questions may be answered by making use of numerical climate models. To do so, however, requires a climate model that can produce a faithful simulation of the NAO and other natural climate variability. The aim of the present work is to demonstrate the strengths and weaknesses of the Hadley Centre coupled climate model (HadCM2: Johns et al. 1997) in this respect.

Using a range of observed, paleoclimate and simulated datasets (described in Sect. 2), the definition of the NAO (Sect. 3 and Appendices B and C), its

multivariate spatial structure (Sect. 4) and its temporal characteristics (Sect. 5) are evaluated in both the model and the real climate system. We consider only the 'winter' season, since this shows the strongest interdecadal variability and the strongest influence of the NAO on surface climate. Some emphasis is placed on the decadal and interdecadal time scales, since it is there that model simulations can add the most value to our observation-based understanding of the NAO.

2 Model and observed datasets

2.1 Observed datasets

Four monthly-mean gridded datasets are used. The first is sea level pressure (SLP) from 1873 to 1995 on a 5° latitude by 10° longitude grid covering much of the Northern Hemisphere (Jones 1987; Basnett and Parker 1997). This dataset is derived from the UK Meteorological Office analyses. The second set is instrumental surface temperature (IST) anomalies from 1873 to 1995 on a 5° by 5° grid box basis, also covering much of the Northern Hemisphere. This is a blended dataset (Parker et al. 1994; Jones et al. 1999) of sea surface temperature anomalies (Parker et al. 1995) and 1.5 m air temperature anomalies over land (Jones 1994). Third, a dataset (Hulme 1994) of precipitation totals (H-PREC), also on a 5° by 5° grid box basis, is used. This contains monthly data from 1900 to 1995 for land areas only, derived from rain-gauge measurements, and we use available grid boxes in the Northern Hemisphere only. This is supplemented by the fourth dataset, which is the shorter, but spatially complete (land and marine, with no missing data) precipitation dataset of Xie and Arkin (1997). This is a blend of rain-gauge data over land, satellite data over the oceans and re-analysis data where neither of the former are available/suitable. It covers the period 1979 to 1995, and we denote it by XA-PREC.

Beside the monthly-mean datasets, a *daily* version of the UK Meteorological Office SLP analyses is used to diagnose changes in storm-track-related synoptic variability. This dataset is on the same 5° by 10° grid as the monthly SLP data, and we use data from 1950 to 1995.

In addition to the gridded datasets, three time series of station pressure are used. These are Ponta Delgada in the Azores (37.7 °N, 25.7 °W) from 1865 to 1995, Gibraltar (36.0 °N, 5.5 °W) from 1823 to 1996, and a southwest Iceland time series (Jones et al. 1997), based mainly on Reykjavik (64.1 °N, 22.0 °W), also from 1823 to 1996. The locations of these stations are indicated by the circles in Fig. 1a–e.

Finally, a reconstruction of the NAO index from 1701–1980 by Cook et al. (1998), based on a network of tree-ring-width chronologies, is assessed (Appendix C) to consider its utility in defining the longer-term temporal characteristics of the NAO.

2.2. Model datasets

Monthly SLP, 1.5 m air temperature and total precipitation data for the entire Northern Hemisphere have been used from a number of integrations of the Hadley Centre's Unified Model (HadCM2: Johns et al. 1997; Tett et al. 1997; Mitchell et al. 1999). All data are on the model's 2.5° latitude by 3.75° longitude grid. HadCM2 is a fully-coupled ocean-atmosphere general circulation model, with flux adjustments used to maintain a steady and realistic climate, and it has a thermodynamic sea-ice model with ice drift driven by surface currents. We use the first 1400 y from the control simulation (CON) of this coupled climate model with fixed radiative forcing representative of present-day levels of greenhouses gases. In addition, a 50-y section of *daily* SLP data from this control integration is used for characterising synoptic variability.

Eight other simulations have also been used; these eight, each 240 y long, are all perturbed by changes in radiative forcing (Mitchell et al. 1999). Four (Ga1–4) are forced by the effects of historically-increasing greenhouse gases from 1861–1990, and thereafter by a compounded 1% y^{-1} increase in effective carbon dioxide from 1991–2099 (similar to the IS92a scenario of Leggett et al. 1992). The other four (GSa1–4) have, in addition, a representation of the direct radiative effect of sulphate aerosols. Each ensemble of four integrations has identical forcing but was begun from different initial conditions; together, these experiments facilitate the separation of an externally forced climate change signal (which should be similar in all ensemble members) from internally generated variability (which is likely to be temporally uncorrelated between ensemble members).

Note that the model does not perfectly conserve mass (Tett et al. 1997). This manifests itself as a downward trend in the global mean SLP that, in extended integrations such as those used in the present study, can be significant. An adjustment was made to the SLP data to remove this spurious trend (see Appendix A for details).

3 Definition of a winter NAO index

Various approaches have been used to define an index of the NAO, e.g. (normalised) pressure differences between various station pairs (Rogers 1984; Hurrell 1995; Jones et al. 1997), area-weighted pressure extremes (Santer 1988), or principal component time series corresponding to a pressure field principal component pattern (Rogers 1990). Similarly, various definitions of a winter season have been used, e.g. December to February (DJF, Rogers 1984; Dickson et al. submitted 1998), December to March (DJFM, Hurrell 1996), or November to March (NDJFM, Jones et al. 1997).

Our main aim is to evaluate a climate model simulation; in this respect, the precise definition of the index is of less importance provided that the comparisons are performed on a like-with-like basis (i.e. the same definition must be used for both observational and model data). Nevertheless, since the questions that we wish to answer with the help of a realistic climate model involve large spatial scales and the interdecadal time scale, a comparison between the different indices is undertaken. Interdecadal variability appears to be most coherent when the December to March season is used (see Appendix B). Using this season, various possible index definitions are investigated in Appendix C. A pattern-based index of the NAO is preferable because other modes of variability do have some power over the individual stations that are commonly used to define the NAO. However, because all indices are highly correlated at interannual and interdecadal time scales, and there is no single pattern-based index, and the pattern-based indices result in much shorter time series, we are justified in taking the length of the available record as the determining factor. We use, therefore, the absolute pressure difference between Gibraltar and SW Iceland (G-I) for the period DJFM 1823/4 to DJFM 1995/6 (Jones et al. 1997). This observed NAO index is shown in Fig. 2b.

For the December to March season, the Gibraltar and SW Iceland time series (locations marked in Fig. 1a–e) are anti-correlated ($r = -0.68$ over 173 y), indicating an adequate signal-to-noise ratio for the oscillation (see Hurrell and van Loon 1997). Time series of simulated DJFM SLP from the nearest HadCM2 grid boxes to these sites (Fig. 1f–j) are also anti-correlated ($r = -0.55$ over 1400 y).

4 The spatial signature of the NAO

Both model and observations have a similar mean winter SLP pattern (Fig. 1a, f) in the Atlantic half of the

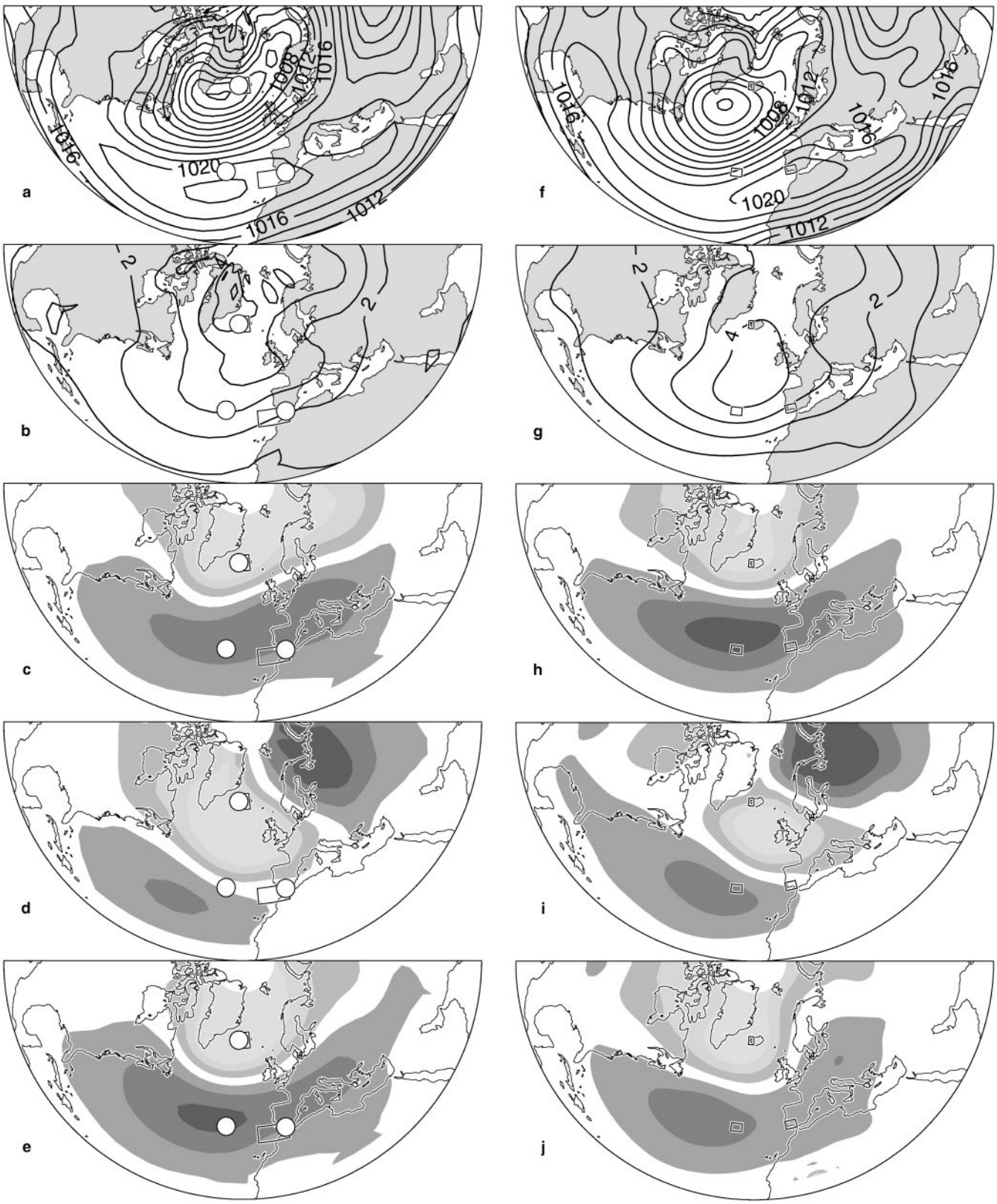

Fig. 1a–e Observed and **f–j** HadCM2 simulated DJFM sea level pressure patterns: **a, f** long-term mean (hPa); **b, g** interannual standard deviation (hPa); **c,h** leading and **d,i** second EOFs of the SLP covariance matrices; **e,j** leading rotated EOFs. All EOF patterns are dimensionless, with *light grey*, *white* and *dark grey* showing negative, near-zero and positive values, respectively. *Circles* in **a–e** show locations of station observations from Azores, Gibraltar and Iceland. *Squares* identify the grid-box data used to represent Gibraltar and Iceland observed SLP **a–e** and Azores, Gibraltar and Iceland simulated SLP **f–j**

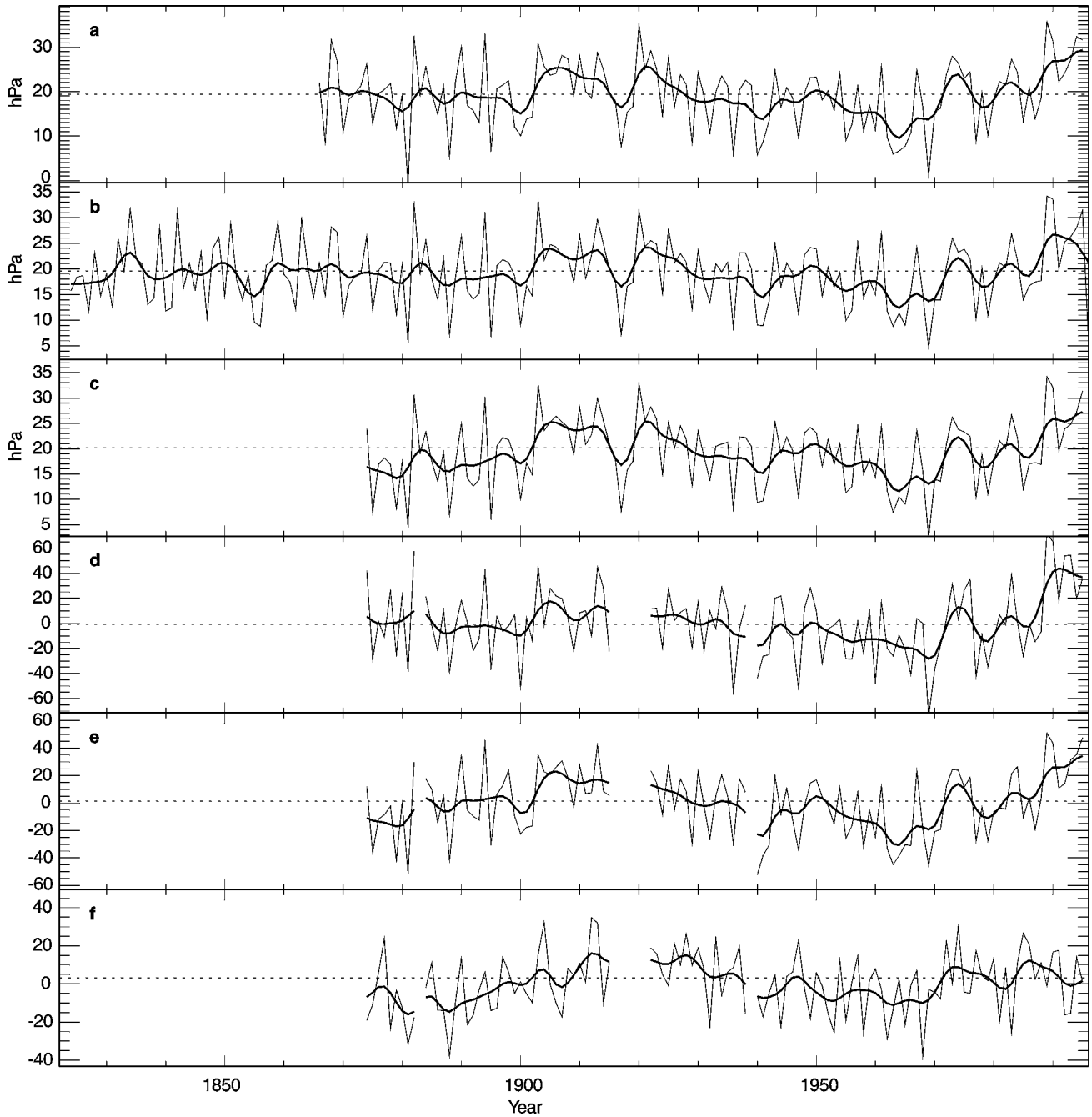

Fig. 2a–f Observed DJFM time series of **a** Azores SLP minus Iceland SLP (hPa), A-I; **b** Gibraltar SLP minus Iceland SLP (hPa), G-I; **c** difference between grid-box SLP values nearest to Gibraltar and Iceland (hPa), G-Igrid; **d** leading principal component time series, PC1; **e** as **d** but following rotation, rotPC1; **f** as **d** but second component, PC2. Each vertical scale was selected to cover a range from 3 standard deviations (SD) below to 3 SD above the mean, with mean (marked by *dotted line*) and SD computed from the 1901–1960 period of each time series. Decadally-filtered time series are also shown

Northern Hemisphere, although the simulated Azores High is a little weak and shifted towards the east. The interannual variability is similar (see, e.g. the 2 hPa standard deviation isoline in Fig. 1b, g), except in the northern North Atlantic where the observed standard deviation is >5 hPa, but the simulation has a maximum of just over 4 hPa that is also shifted to the south.

The leading principal components (PCs) of these datasets (defined from the eigenvalues of their covariance matrices) have corresponding empirical orthogonal function (EOF) patterns that both show the classic NAO signal (Fig. 1c, h; compare with Rogers 1990). The patterns are very similar, indicating that the HadCM2 CON simulation contains a realistic SLP

oscillation in this region (the leading mode accounts for 40% of the SLP variability of this region in both datasets).

The spatial characteristics of the NAO, as defined by our preferred index (Gibraltar minus Iceland DJFM pressure difference), in terms of its signature in the SLP, temperature and precipitation datasets is demonstrated in two ways: first, associated correlation patterns (ACPs) indicate the correlation coefficients between the NAO index and timeseries at all other points in the dataset; second, associated regression patterns (ARPs) indicate the linear regression coefficients in the same way. Following Hurrell (1996), the G-I NAO index (observed or simulated) was normalised before computing the regression coefficients, so that the ARP values indicate the magnitude of the response to a one standard deviation (SD) increase in the NAO.

4.1 Sea level pressure (SLP)

The NAO is an oscillation of the pressure field and we have already seen from the PCA results (Fig. 1) that the model simulation is very similar to the observations. This is now extended to the entire Northern Hemisphere and to the particular pattern associated with the G-I index of the NAO. In terms of correlation coefficients, the observed ACP (Fig. 3a) and the simulated ACP (Fig. 3c) are very similar in the Atlantic half of the hemisphere, except that the region of positive correlation with the NAO extends a little too strongly towards the western Atlantic in HadCM2. The zero-line is well located. Outside this region, the positive correlations south of the Aleutian Low in the North Pacific are simulated too strongly, the lobe of negative correlations over Siberia is too weak while that over Canada is too strong. The regression coefficients (Fig. 3b,d) support these comparisons and add the result that the southern centre shifts to the west of Iberia in both datasets, indicating that the largest absolute pressure anomalies occur there even though maximum explained variance is centred over Gibraltar (since Gibraltar is the southern station in our pressure difference index). Further, the simulated ARP is too strong in the Azores region while being too weak over Iceland (consistent with the differences in interannual SLP variability shown in Fig. 1b,g), although these minor errors compensate and the overall pressure anomaly difference is reasonably simulated.

The correlation between the NAO and the North Pacific SLP is of particular interest. Since they have opposite effects on Northern Hemisphere temperatures (Hurrell 1996), phase locking between the NAO and the North Pacific Index (Trenberth and Hurrell 1994) could reduce the combined effect on hemispheric temperature. Partial phase locking occurs in the HadCM2 simulation (note the positive regression coefficients between North Pacific SLP and the NAO index in Fig. 3d), and a similar feature is evident in the model simulations of Timmermann et al. (1998). In the real climate system, however, the full available record indicates only a weak correlation (Fig. 3a,b), although subsampling the record indicates somewhat stronger correlations during the 1900–1949 period.

The possible link between the atmospheric circulation patterns over the Pacific and Atlantic Oceans deserves further comment. The leading EOF of observed hemispheric sea level pressure is the Arctic Oscillation (AO) and has the Iceland and Aleutian Lows *anti-correlated* (Thompson and Wallace 1998). The Arctic Oscillation is not independent of the NAO; instead it is more zonally symmetric and represents a less regionally constrained definition of a preferred mode of the Northern Hemisphere atmosphere, differing from the NAO mainly in the Pacific. The HadCM2 SLP pattern (Fig. 3c,d) shows a closer resemblance to the observed Arctic Oscillation (Thompson and Wallace 1998), and may indicate that the Gibraltar minus Iceland index is capturing this oscillation rather than the NAO in HadCM2.

Given the strong links between the El Niño-Southern Oscillation (ENSO) and the atmosphere over the North Pacific (Trenberth and Hurrell 1994), any correspondence between Pacific and Atlantic SLP might also depend upon a link between the NAO and ENSO. If such a link exists, then it is not strong (Rogers 1984; Hurrell 1996; Huang et al. 1998). We note that the DJFM sea surface temperature (SST) time series from the Nino3.4 region of the equatorial Pacific (a useful measure of ENSO, see Trenberth 1997) does not correlate significantly with the winter NAO index over the period 1951–1996 (whether unfiltered or after the application of high-, band-, or low-pass filters). At the lower frequencies (e.g., between 30-year low-pass filtered series) the correlation reaches $+0.8$, but it is not statistically significant (due to high autocorrelation) because it is really only measuring two coincidental trends. The HadCM2 Nino3.4 index is negatively correlated with its NAO index ($r = -0.23$) and remains so when temporally smoothed; this is of the correct sign, but too weak, to explain the link between the NAO and the North Pacific in HadCM2 (Fig. 3c,d), and further study is necessary to identify the dynamics involved in the partial correlation of ENSO and NAO variability.

4.2 Synoptic SLP variability

Winter-mean SLP variability is driven, in part, by changes in shorter time scale synoptic disturbances. Serreze et al. (1997) use a sophisticated storm-tracking algorithm to diagnose the difference in storm activity between winters with high or low NAO index. High NAO index winters have more storms in the Iceland region (contributing to a deeper Iceland Low), fewer in a band stretching from Newfoundland to Iberia and

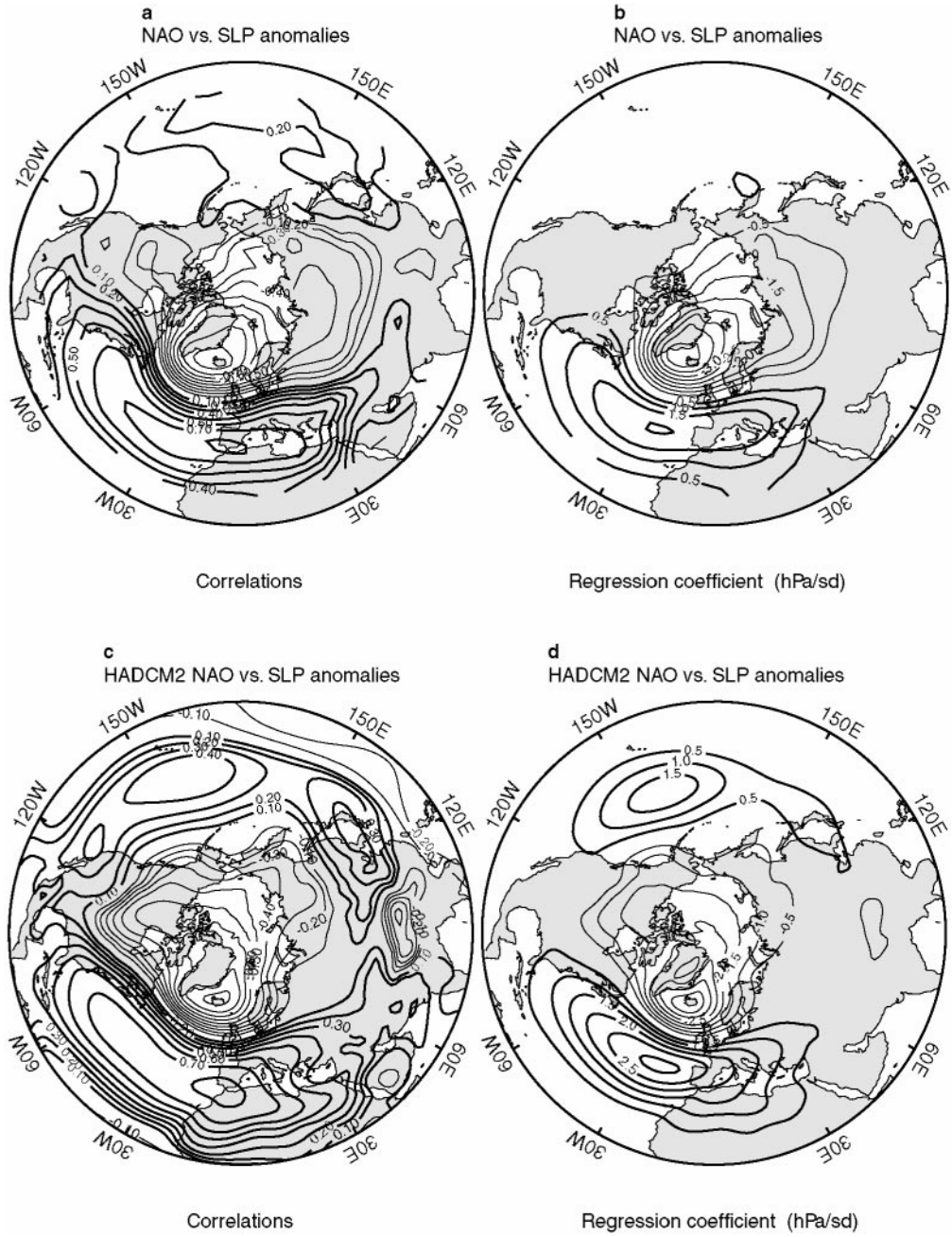

Fig. 3a,c Correlations and **b,d** regression coefficients between local SLP and the normalised NAO index for DJFM, as **a,b** observed and **c,d** simulated by HadCM2, computed over the maximum record length available. Positive isolines are *thick*; intervals are **a,c** 0.1 (dimensionless) and **b,d** 0.5 hPa per standard deviation; zero lines omitted. All correlations shown in **c** are statistically significant (at the 90% confidence level); in **a** the sample size varies slightly due to missing data but all correlations >0.2 are likely to be significant

fewer over the northern Mediterranean Sea (Serreze et al. 1997, their Fig. 6). Here we follow the simpler approach of Rogers (1997) and Hurrell and van Loon (1997) to quantify synoptic variability: the daily SLP dataset at each grid point is filtered by a digital filter that retains mainly variability on a 2–8 day time scale. The standard deviation of each December to March section of the filtered time series then provides a proxy for SLP synoptic variability at each grid point for each winter. Note that this measure of synoptic variability is driven by more than just storm centres (e.g. anticyclonic activity, and variability associated with depressions but away from their centres).

Both the observed (46-y section) and simulated (50-y section) datasets were analysed in this way. Due to differences in resolution, absolute values are not

comparable, so we concentrate on the patterns. The storm tracks simulated by HadCM2 have previously been shown to be reasonably realistic (Carnell and Senior 1998, who used a storm-tracking algorithm). As expected, therefore, the observed and simulated mean fields of synoptic variability (not shown) are quite similar and capture the storm tracks well (as defined, e.g. by Fig. 1b of Serreze et al. 1997), with maxima in the Aleutian Low region and upstream, and in the Iceland Low region and upstream (peaking near to Newfoundland). The interannual variability of synoptic activity is also strongest in the storm-track region (the observed data have an additional maxima over the Tibetan Plateau region, although data quality may be poorer there).

Observed interannual changes in winter synoptic activity are significantly positively correlated with the NAO in a region from Newfoundland via Iceland and Scandinavia to northeast Russia (Fig. 4a). There is decreased synoptic activity to the south of this band and in the Labrador Sea/Baffin Island region when the NAO is in its positive phase. This Atlantic sector

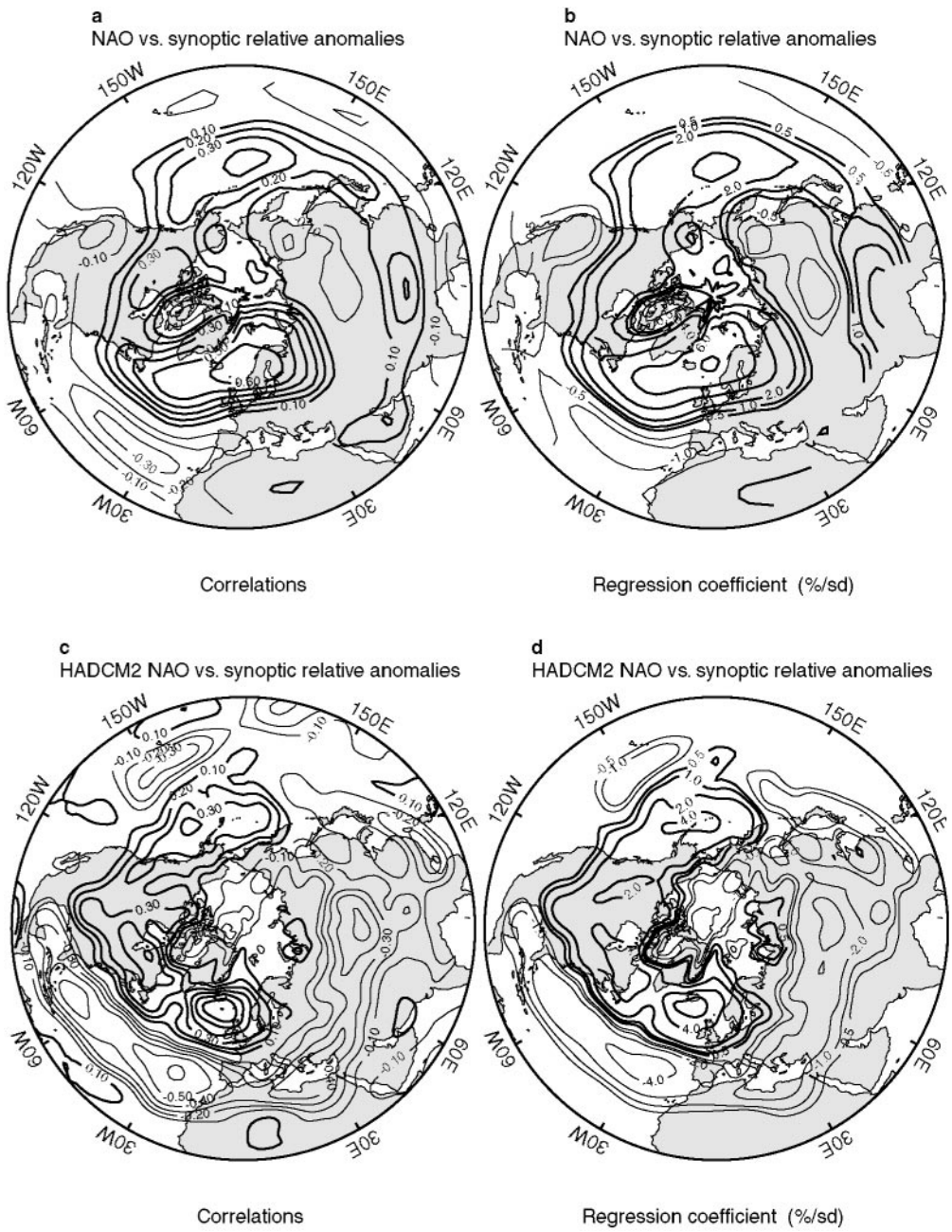

Fig. 4a–d As Fig. 3, but for local synoptic SLP variability versus the normalised NAO index. Isoline values in **b,d** are $\pm$ (0.5, 1, 2, 4, 6 and 8) % per standard deviation. Observed and simulated correlations exceeding 0.25 in **a,c** are statistically significant

pattern agrees well with the results of Hurrell and van Loon (1997) based on a 16-y 300 hPa geopotential height dataset. The pattern is also evident in the HadCM2 CON simulation (Fig. 4c), although the negative correlations are too high in the Azores High region. Both patterns are a little different to the changes in storm numbers that accompany changes in the NAO (Serreze et al. 1997, their Fig. 6), where the positive correlations are confined to a much smaller region near Iceland and the band of decreased storm counts (with high NAO) is 15° of latitude further north.

The magnitude of the variations is relatively small (Fig. 4b, d), peaking at an observed increase of 7% of the mean synoptic variability northeast of Iceland for a unit standard deviation increase in the NAO index. The decreased variability to the south of this region, over both the Atlantic and Asia, is simulated too strongly (Fig. 4d).

4.3 Surface temperature

Similar analyses of the observed (IST) and simulated near-surface temperatures reveal a well simulated set of four major temperature anomalies (Fig. 5; also see Fig. 3a of Hurrell 1996). An increase in the NAO index corresponds to two regions of warming (centred over the southeast USA and over Scandinavia) and two of cooling (centred over the Labrador Sea and northern Africa). The latter cooling is elongated towards the tropical North Atlantic and towards southern Asia. The Scandinavian warming extends into Siberia, peaking in northern Russia at 1.2 °C/SD in the observations (Fig. 5b) but only at 0.8 °C/SD in HadCM2 (Fig. 5d). The correlations (and hence the fraction of variance explained) peak further west where the interannual variability of temperature is lower: over the Baltic in the observed record (Fig. 5a) and (even further west) over the British Isles in HadCM2 (Fig. 5c).

These temperature anomalies can be explained to a large extent by the large-scale atmospheric flow associated with the NAO (indeed, Hurrell and van Loon 1997, their Fig. 12, show that, for the Atlantic sector, heat flux convergence/divergence by transient eddies tends to destroy the temperature anomaly pattern). The SLP patterns (Fig. 3) indicate warm anomalous southerly flow over southeast USA, cold anomalous northerly flow over the Labrador Sea, warm anomalous south-westerly flow over northwest Europe, and cold anomalous northerly flow over the Mediterranean region.

There are two aspects where the simulation differs more noticeably from the observations and that are not as easily explained by the mean atmospheric flow. First, HadCM2 simulates a significant cooling over the Bering Strait when the NAO index is high (Fig. 5c, d), and second, significant warming occurs over the middle of the North Pacific Ocean in HadCM2. Both of these features are probably related to the overly strong response of Pacific SLP (Fig. 3c, d), although not simply a response to the mean circulation anomaly.

As noted in Appendix C, the observed NAO index explains 10.8% of the *point-wise* near-surface temperature variability in winter, when totalled over the Northern Hemisphere. A similar calculation yields a value of 7.7% for HadCM2, but direct comparison is not valid due to the differing grid resolutions. A more useful comparison is with the mean temperature for the entire extratropical (north of 20 °N) Northern Hemisphere; for winter, Hurrell (1996) found that 31% ($r = 0.56$) of its variance was related to the winter NAO over the period 1935–1994. This relationship is, in fact, rather time-dependent in both the observations and in the model simulation (Fig. 6, computed using sliding 50-y windows). The changes are generally larger and more rapid in the observations, however, with potentially important implications for the analysis of Hurrell (1996). Further investigation should be made to identify why the correlation was 0.57 for the 1946–1995 period, yet the NAO explains none of the variability for 50-y periods beginning between 1895 and 1920. Even earlier, the correlation was apparently above 0.5 again (although we note that observational coverage was less, see Jones 1994, and may have been biased towards regions where the NAO has a stronger control on temperature). In terms of model evaluation, over the 1874–1995 period the observed correlation is 0.34, compared to 0.14 over the simulation period shown in Fig. 6b, although it is clearly higher and closer to the observed value over the earlier part of the simulation: $r = 0.31$ over 1900–2099 (model years are arbitrary in this control integration with constant radiative forcing).

4.4 Precipitation

The gauge-based land-only observed precipitation (H-PREC) allows similar ACP (Fig. 7a) and ARP (not shown) fields to be generated. The spatially-complete XA-PREC dataset is too short for such calculations, but composites (see Hurrell and van Loon 1997) can be formed from winters when the NAO has been very high (1983, 1989–90, 1992 and 1995) or slightly lower than normal (1979, 1982, 1985–86, 1988). Given the dependence of the magnitude of precipitation variations on mean precipitation, extra detail is evident if precipitation is expressed as a per cent anomaly from its mean at each point (as was the synoptic SLP variability shown in Fig. 4b,d). Figure 7b shows the difference between the high and low NAO composites of relative precipitation anomalies. Similar calculations with the H-PREC dataset (not shown) produced very similar patterns and magnitudes over the land areas. Figure 7c, d shows HadCM2 control integration results for comparison with Fig. 7a,b.

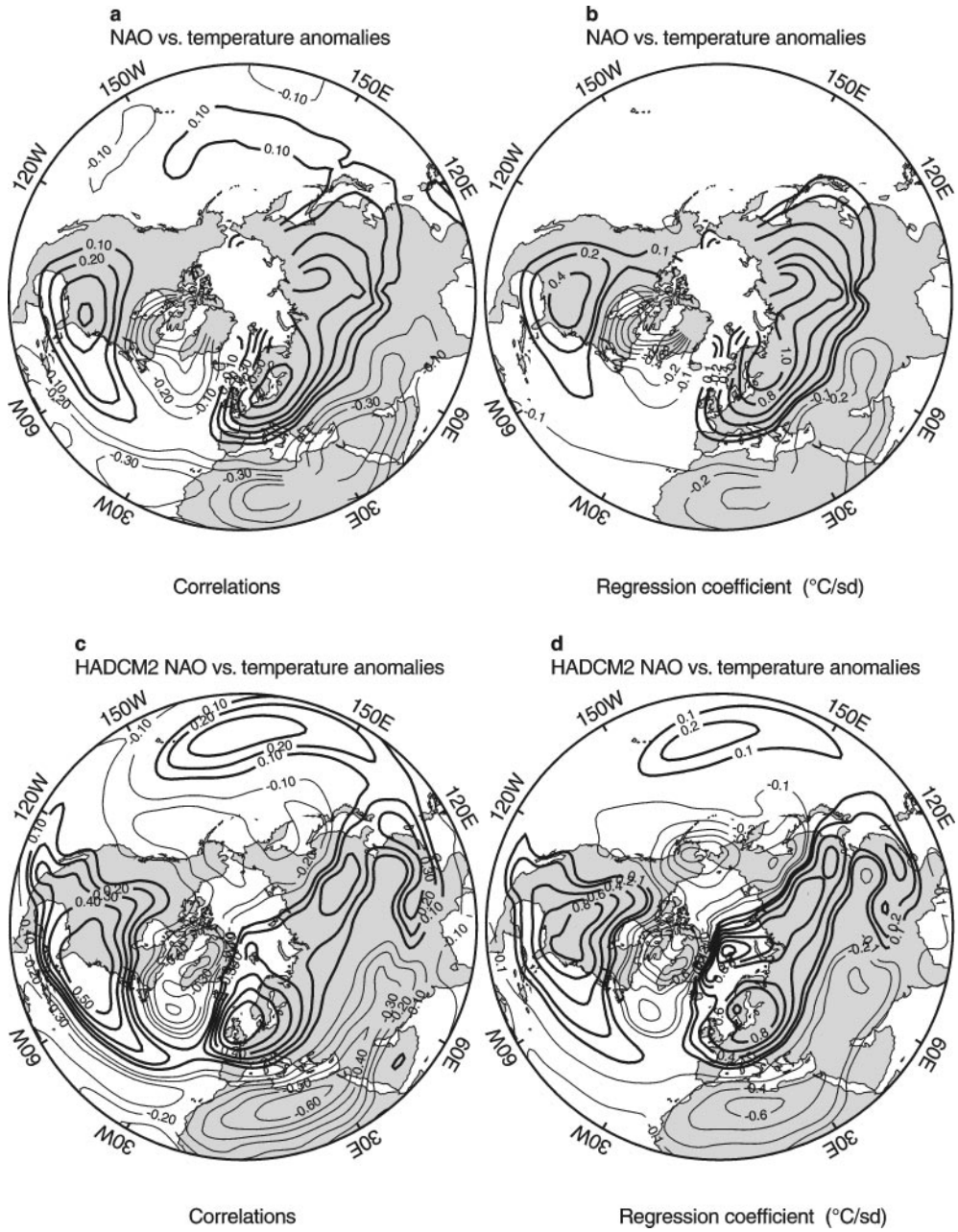

Fig. 5a–d As Fig. 3, but for local near-surface temperature versus the normalised NAO index. Isoline values in **b,d** are $\pm$(0.1, 0.2, 0.4, 0.6, 0.8, 1 and 1.2) K per standard deviation

The observed results agree well with earlier studies (compare Fig. 7a with the European-only Fig. 13 of Hurrell and van Loon 1997; compare Fig. 7b with the Atlantic-region Fig. 4 of Hurrell 1995). The wet northern and dry southern Europe pattern associated with high NAO index conditions is now well known, and HadCM2 reproduces it well (Fig. 7c, d). The strongest drying is centred over Iberia, although the relative anomalies used for Fig. 7b shift southwards into Algeria where the mean precipitation is lower. Other features include the (observed and simulated) positive coefficients over Libya and Egypt, possibly extending eastwards in a band of similar sign (these bands are discussed further in the next subsection). Looking west, the high precipitation in northwest Europe is connected to higher precipitation in North America by a storm-track-associated (Fig. 4) band of enhanced precipitation across the Atlantic Ocean. This band is oriented too zonally in the model simulation (Fig. 7c,d) connecting it to a wetter eastern Canada, compared to the results from the Xie and Arkin (1997) dataset which shows the Atlantic band of positive anomalies having a more SW-NE orientation (Fig. 7b). Perhaps the most notable difference between the HadCM2 simulation and the observations is that both XA-PREC and H-PREC suggest that eastern Canada is dry during

Fig. 6 a Observed and **b** simulated (by HadCM2 CON) correlations between DJFM NAO index and DJFM extra-tropical Northern Hemisphere temperatures, from sliding 50-y windows plotted against start year of window. Horizontal scaling is the same to facilitate comparison; simulated results are from the first 650 y of the simulation. Correlations of 0.25 are significant in the absence of autocorrelation

high NAO winters and southern USA is wet, while HadCM2 simulates the reverse. Observations and model simulations agree, however, that the Labrador Sea precipitation is negatively correlated with the NAO.

4.5 Multivariate considerations

Evaluating the model in terms of individual variables has shown that HadCM2 simulates a realistic NAO in spatial terms, even for variables that are often considered to be less reliable, such as synoptic variability and precipitation. Considering the results in combination, it is clear that much of the temperature response (Fig. 5) can be explained by anomalous mean flow (Fig. 3), while synoptic variability (Fig. 4) is the determining factor for precipitation variations (Fig. 7).

Hurrell and van Loon (1997) demonstrate that synoptic variability contributes to the maintenance of the Atlantic-sector large-scale flow (and hence SLP) anomalies associated with the NAO. Perhaps, then, differences in the observed and simulated strength of the teleconnection with North Pacific SLP (Fig. 3) might be explained by the synoptic variability changes (Fig. 4). A noticeable difference is that the southern disturbance track responds more strongly to the NAO in HadCM2 than in the observations, and this stronger response continues across Asia to the Pacific coast. The impact on the Pacific region of this difference in synoptic variability might explain the bias in HadCM2 that occurs there, but we have not considered the dynamics of such an interaction.

Comparing both the synoptic variability and the precipitation (Figs. 4 and 7), there is an apparent spiral of bands of alternating sign anomalies over much of the hemisphere. Positive values over North Africa extend eastwards over the Tibetan Plateau towards southern Japan, through the Aleutian Low, across all but the far north of Canada, across the northern North Atlantic towards Scandinavia and into the Eurasian Arctic. Negative values span from central America, through the Mediterranean and eventually into northeast Siberia. This band of negative values then becomes rather broken, perhaps following the Canadian Arctic coast and reaching Baffin Bay and the Labrador Sea. This band structure is not clearly defined, and is not everywhere statistically significant in the observed data, but the similarity between the model and the observations and between the synoptic variability and the precipitation lends some support to its existence. If such banding is real, it implies that part of the North Atlantic Oscillation is driven by events upstream from the Atlantic (although if the NAO is part of a hemispheric

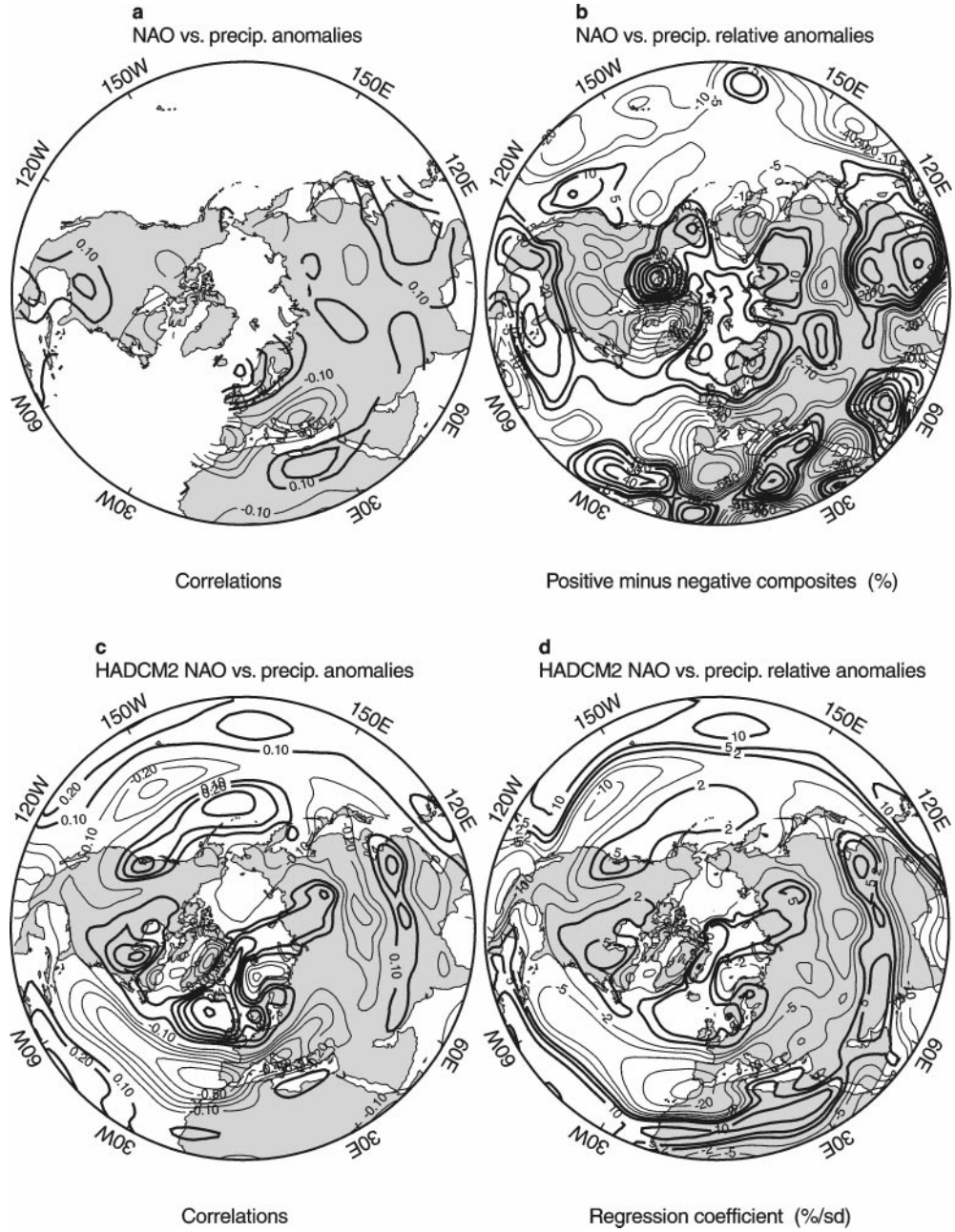

Fig. 7a–d As Fig. 3, but for local precipitation versus the normalised NAO index, and **b** is the difference in precipitation per cent anomalies from a composite of winters with highly positive NAO index compared to a composite with negative NAO index. Isoline values in **b** are $\pm$(5, 10, 20, 30, 40 etc.) % and in **d** are $\pm$(2, 5, 10, 20 and 30) % per standard deviation. Most observed correlations > 0.25 in **a** are significant; all simulated correlations shown in **c** are significant

pattern, see Thompson and Wallace 1998, then 'upstream' and 'downstream' become meaningless).

5 Temporal characteristics

5.1 Variability and trends

Power spectra (Fig. 8a, b) of the observed and simulated Gibraltar minus Iceland NAO indices indicate near-white behaviour. (Using the Azores minus Iceland version also results in a white spectrum for HadCM2, but a redder spectrum for the observations. This is in accord with the discussion at the end of Appendix C of differences in the "redness" of the various observed NAO indices, and is in accord with published power spectra of the NAO, see Hurrell and van Loon 1997.) The HadCM2 spectrum exhibits a number of peaks that differ significantly from an autoregressive spectrum, notably at periods of around 4–5, 8 and 15 y. These have not been investigated further, since they do not appear in the observed spectrum (although *near*-significant peaks occur at around 2.5 and 8 y).

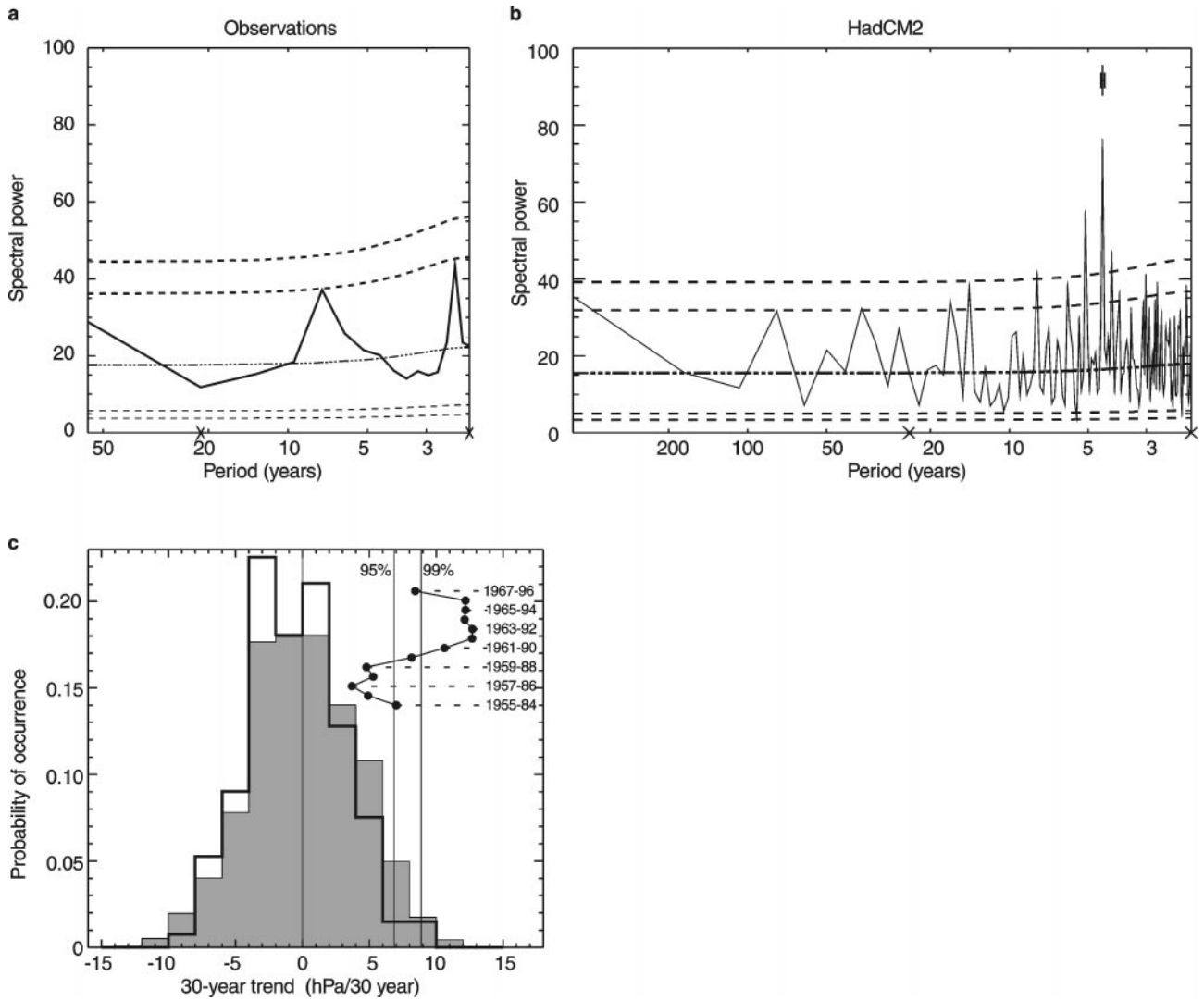

Fig. 8 Power spectra (*solid lines*) of **a** observed and **b** HadCM2 NAO indices, together with best-fit AR1 process (*dash-dotted*) and 95% and 99% confidence intervals for AR1 process (*dashed lines*). *Crosses* on the *x axis* indicate the range of frequencies where the AR1 fit was done (periods between 2 and 22 y). **c** Distribution of 30-y trends (hPa per 30 y) in the DJFM NAO index from HadCM2 CON (*shaded*) and observations (*thick line*, using pre-1984 trends only). *Vertical lines* show the zero-trend line, and the 95 and 99 percentiles from the HadCM2 distribution. Overlaid (*dots connected by lines*) are the 30-y trends observed between 1955–84 and 1967–96, some of which fall outside the 99 percentile of the HadCM2 distribution

Time series of the observed and simulated NAO indices reveal a difference at lower frequencies (periodicities of 15 y and longer). While the interannual (high-pass filtered) standard deviation of the winter NAO index is well simulated (6.1 hPa in HadCM2, compared to an observed value of around 6.3 hPa), if the indices are first smoothed with a 10-y low-pass filter, then the HadCM2 standard deviation of 2.4 hPa is noticeably lower than the 3.2 hPa observed over 1895–1995. This difference, however, is due entirely to the most recent large positive trend in the observed data between the 1960s and the 1990s (Fig. 2b); if the comparison is made using the observed period 1850–1950, then the interdecadal standard deviation is close (2.2 hPa) to that simulated by HadCM2 (2.4 hPa).

To consider this difference between model and observations further, we generated the frequency distributions of 30-y trends in each NAO index (Fig. 8c). The observed winter NAO index has its largest positive trend from 1963–1992 ($+12.7$ hPa/30 y), and this is highly unusual compared to the variability of the HadCM2 control integration (95% of 30-y trends in this simulation were between -8 and $+7.8$ hPa/30 y, with extremes of ± 12 hPa/30 y). The post-1950 period contributes greatly to the positive tail of the observed distribution of trends (Fig. 8c), with five trends (those starting between 1962 and 1966) exceeding the highest HadCM2 trend.

5.2 Anthropogenic forcing

The recent observed trend in the observed NAO index exceeds any simulated during the model control run, indicating either that HadCM2 is deficient in its simulation of interdecadal variability or that the observed changes are externally forced. While other forcings cannot be ruled out, the most likely candidate is anthropogenic forcing of the climate system. Figure 9 shows the ensemble-mean NAO index time series (filtered to highlight the interdecadal and longer changes) from the perturbed integrations, together with the mean and appropriate significance levels computed from the control integration. The ensemble means highlight mainly the underlying signal, to which variability is added to produce any one realisation. The filtered observed index (Fig. 9) is, of course, a single realisation, so it combines signal and full variability and the same significance levels do not apply.

Immediately obvious is a statistically significant decline in the mean NAO index next century, particularly in the ensemble that includes sulphate aerosol forcing as well as greenhouse gases (GSa). Global warming, as simulated by HadCM2, does *not* manifest itself in the long term by amplifying the winter pressure gradient between Gibraltar and Iceland (which is our NAO index and which weakens during the next century). The reduction in the mean NAO index in these ensembles of perturbed integrations is due to a long-term change in the distribution of atmospheric mass (not shown); this shift does not resemble the NAO pattern itself, but does act to weaken the SLP gradient between the eastern Azores High (where Gibraltar lies) and the Iceland Low. Other definitions of the NAO index might yield different signals.

Earlier in the perturbed simulations (Fig. 9), however, there are weakly significant maxima in the NAO index in the Ga ensemble in year 1915 and in the GSa ensemble in year 1990, and a minima in the Ga ensemble in year 1953. The maximum in the GSa ensemble has a similar timing to the recent observed maximum, and is the highest 30-y filtered value in the entire ensemble mean time series. This suggests some anthropogenic contribution to the recent high extreme of the NAO index, a contribution that HadCM2 suggests will reverse during the next century. The statistical significance of the result is, however, somewhat equivocal. First, we would expect some values to exceed the lowest or highest 5% of control-integration-derived values by mere chance, so the 1990 NAO index maximum in the GSa ensemble might be internal variability rather than an externally forced signal. Second, even if the trend from the 1950s to 1990s in the GSa ensemble (Fig. 9) is a real signal, it makes only a small contribution to the observed trend (no more than +2 hPa/30 y). Internal variability will occur on top of this signal, but 95% of 30-y trends from the control integration are less than +7.8 hPa/30 y indicating a combined 95% limit (signal plus variability) of no more than + 9.8 hPa/30 y. This is weaker than the observed maximum trend of + 12.7 hPa/30 y from 1963–1992, and so we re-iterate that this observed trend is outside the range of variability simulated by HadCM2 with or without anthropogenic greenhouse gas and sulphate aerosol forcing.

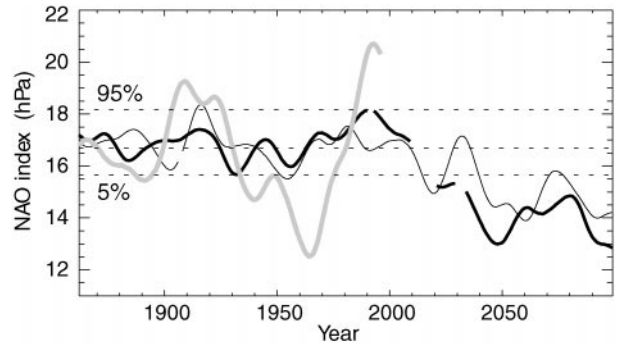

Fig. 9 Time series of filtered DJFM NAO indices (hPa). NAO indices are from four Ga experiments (*thin line*) or four GSa experiments (*thick line*), filtered with a 30-y low-pass filter, then averaged together to produce just the ensemble means (*gaps* indicate unavailable data). *Dotted lines* indicate the mean NAO index and the 5% and 95% levels (from the mean of four 30-y filtered values) computed from the control integration. Also shown (*grey line*) is the filtered observed NAO index, adjusted so that its mean over the first 50 y matches the control integration. The 5% and 95% levels do not apply to the observed series, being a single filtered time series rather than the average of four

6 Summary and discussion

The Hadley Centre's coupled climate model (HadCM2) generates a realistic pattern of winter mid-latitude synoptic variability in the Atlantic sector, maintaining a seasonal-mean sea level pressure field that is close to that observed. The leading mode of interannual variability of the winter-mean SLP field in both the model output and the observations is the well-known North Atlantic Oscillation, a shift of atmospheric mass between the Azores High and the Iceland Low. The mean flow anomalies associated with this oscillation generate surface temperatures with the same four anomaly regions as observed.

The interannual variations in synoptic variability involved in the generation of the oscillation are similar in the observed and simulated data, and drive precipitation variations (particularly over Europe, where the HadCM2 response is realistic). The changes in synoptic variability are strongest in the Atlantic/European region, but both upstream and downstream changes are also apparent. In spatial terms, the model's main deficiency appears to be that it overestimates the correspondence between the NAO and SLP over the North Pacific Ocean, perhaps related to differences in the

synoptic variability fields: high NAO index winters are associated with an overly weakened southern synoptic disturbance track across Asia in HadCM2.

Although the occurrence of the NAO and its spatial structure are well simulated, it should be noted that this is mainly a test of the realism of the atmospheric component of the model, since that is all that is needed to produce a North Atlantic Oscillation (Glowienka-Hense 1990; although Saravanan 1998 found that a coupled model was required to simulate the finer details of the spatial structure). Nevertheless, a more stringent test of the *coupled* model is its ability to simulate the temporal characteristics of the NAO, particularly any decadal variability for which the oceans may provide the inertia of the coupled system. Omitting the most recent upward trend in the NAO index from the 1960s to the early 1990s, the observed temporal variability has been shown to be well reproduced by HadCM2.

The recent NAO trend, however, lies outside the range of variability exhibited by the 1400-y control integration. This implies that either the model is deficient or that the cause of the trend is not internally-generated natural variability of the climate system. External anthropogenic forcing in the form of increased greenhouse gas and sulphate aerosol concentrations may be able to explain a small part (perhaps 20%) of the recent trend, although increased future forcing would yield a decline in the NAO index next century (at least for the pressure index used here, due to a shift in the mean sea level pressure pattern; different definitions of the NAO might yield a different anthropogenic signal). Different climate models also produce different anthropogenic signals (e.g. Hupfer et al. submitted 1999 find a slightly enhanced winter NAO in the future).

Other forcings that we have not considered are natural variations in solar or volcanic activity. Additionally, the coupling between the winter troposphere and stratosphere (Perlwitz and Graf 1995) suggests that the anthropogenically-induced decrease in stratospheric ozone concentrations may result in a strengthening of the North Atlantic Oscillation (Thompson and Wallace 1998), although cause-and-effect could be reversed.

The alternative possibility is that the HadCM2 model is deficient in either generating or responding to sea surface temperature variations. An earlier version of the AGCM used in HadCM2 was unable to reproduce the observed evolution of the winter NAO index when forced by the observed sequence of SST and sea-ice distribution (Davies et al. 1997); however, similar experiments with a more recent version of the Hadley Centre AGCM forced by an improved dataset of observed SST and sea-ice indicates that the low-frequency aspects of the NAO variability *are* driven in part by variations in the lower boundary conditions, at least for the post-1940 period (Rodwell et al. 1999; we note that the inappropriate scaling applied to their Fig. 1a overstates the link between SST and winter NAO). Such experiments identify the atmospheric response to external forcings and natural internal variations to the extent that these are recorded in the SST fields. These cannot be easily separated, but we note that (1) Rodwell et al. 1999, find only a minor impact of equatorial Pacific SSTs, implying little role for multi-decadal ENSO variations in driving the recent NAO trend; (2) the pattern of Atlantic SST anomalies that drives an NAO response in their model resembles that linked to variability in the wind-driven subtropical gyre (e.g. Grotzner et al. 1998); and (3) this pattern does not resemble that linked to variability in the HadCM2 thermohaline circulation (Tett et al. 1997), although we have found that an enhanced thermohaline circulation does drive a slightly enhanced winter NAO in HadCM2. Further model-based experiments are clearly required to identify the driving mechanisms of low-frequency variability of the North Atlantic Oscillation.

Acknowledgements Support was provided for TJO by UK Natural Environment Research Council (GR9/02522) and by UK Department of Environment, Transport and Regions (DETR EPG 1/1/48), for KRB by Commission of the European Communities (ADVANCE-10K, ENV4-CT95–0127), for SFBT (and HadCM2 integrations) by UK DETR (PECD 7/12/37), for PDJ by US Department of Energy (DE-FG02–98ER62601) and for RMT by Fundação para a Ciência e a Tecnologia Lisbon (PRAXIS XXI, BD/5734/95-RN). Hadley Centre model output was obtained from the Climate Impacts LINK Project (UK DETR EPG 1/1/16). The XA-PREC dataset was kindly provided by Jim Hurrell. Input from a number of sources is acknowledged, particularly Dave Stephenson, John Mitchell and two anonymous reviewers.

Appendix A: numerical drift in sea level pressure

The HadCM2 model does not perfectly conserve atmospheric mass, resulting in a long-term drift in globally-averaged sea level pressure (Fig. A1). The decrease appears to be dependent on model state, being more rapid for warmer states (compare the Ga, GSa and CON results in Fig. A1). In year 2090 of the CON integration, the rate of this numerical drift was reduced to about one quarter by increasing the precision used for storing the model state (Tett et al. 1997).

It can be assumed that this numerical drift affects all points equally. It does not, therefore, affect our NAO index results that are based on the pressure difference between two points. But other analyses that make use of the pressure at individual points, or principal component analysis of the pressure field, could be biased by this drift. The longer the time period considered, the more important the drift becomes relative to other variations.

In the present study, therefore, we have computed the area-weighted global-mean SLP for each winter (Fig. A1), subtracted it from the SLP at all individual points for that winter, and added back a constant (1009.07 hPa, the global-mean SLP from the mean of the first 30 winters of the control integration). This 'corrected' dataset was then used in all analyses presented here.

Appendix B: definition of a winter season

The absolute pressure difference between Gibraltar and SW Iceland on a monthly time scale is used as an NAO index to select a winter season, using the period 1865 to 1995. Table B1 shows the

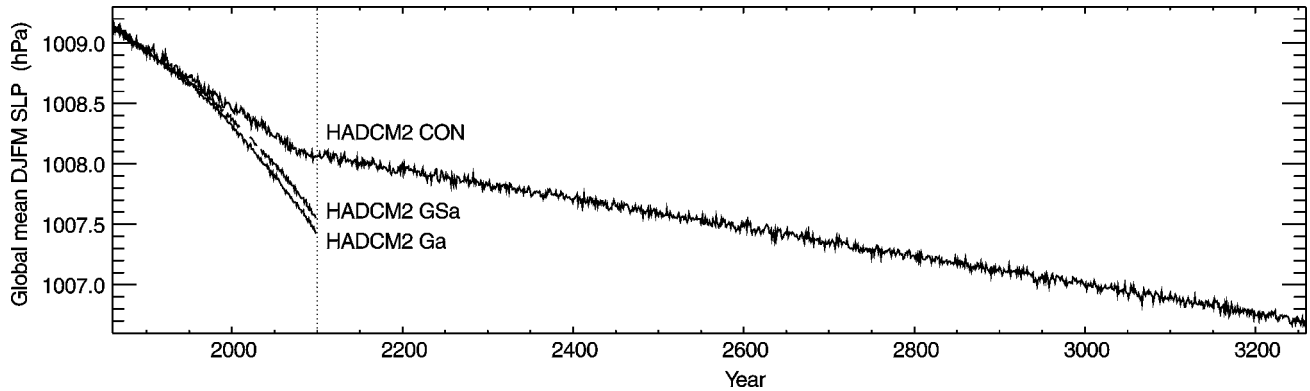

Fig. A1 Global mean DJFM SLP (hPa) from the 1400-y HadCM2 control integration and the 240-y ensemble-mean perturbed integrations.

Table B1 Correlations (×100) between separate monthly time series of the observed NAO index (Gibraltar minus-Iceland pressure) for November to April. Subdiagonal values are between the raw time series; above-diagonal values are between decadally filtered monthly time series. Positive correlations significant (accounting for autocorrelation) at the 90% level are bold, and at the 95% level are also underlined. Equivalent correlations for the HadCM2 NAO index are given in parentheses (the 1400 year record means that lower correlations can still be statistically significant)

	N	D	J	F	M	A
N	—	**39** (7)	− 6 (8)	− 6 (4)	21 (6)	− 21 (− 8)
D	**21** (**11**)	—	**29** (**16**)	24 (− 8)	15 (**9**)	24 (**15**)
J	− 12 (2)	**15** (0)	—	**44** (**15**)	22 (− 8)	− 2 (6)
F	− 7 (0)	3 (− 2)	**16** (**15**)	—	**28** (**18**)	21 (7)
M	4 (0)	2 (1)	5 (− 1)	**21** (**20**)	—	− 7 (**23**)
A	− 2 (− 3)	**21** (2)	6 (2)	12 (3)	**− 15** (**20**)	—

correlations between each monthly time series and all other months, for the months from November to April (note that November and December monthly values are correlated against January to April monthly values from the following year). The subdiagonal correlations were calculated using raw data, while those above the diagonal were computed by correlating monthly time series that had first been smoothed to remove variations on timescales of less than 10 y (all the Januarys are filtered separately from all the Februarys, etc.).

The interannual correlations (Table B1) are of interest in their own right because they demonstrate that the potential for predicting the future state of the NAO simply from its current state is very low. Correlations are significant, but low, for one month ahead from November, December, January, February or March, and four months ahead from December (but not for intervening months). Once each time series is filtered, correlations become stronger between adjacent months (though the levels of significance are no better, due to the increased autocorrelation after filtering). All adjacent months from November to March are significantly correlated.

Starting from the premise that our winter season must contain at least January, we see from Table B1 that the interdecadal variations in January are significantly correlated with the interdecadal variations of the NAO index that occur in December and February. Forming a DJF mean and re-filtering, interdecadal time scale correlations are significant with March ($r = 0.30$) but not with November ($r = 0.07$) or April ($r = 0.17$). Next, forming a DJFM mean, decadal correlations remain non-significant with November and April (both have $r = 0.12$).

On the basis of the decadal coherence, therefore, we define DJFM as our winter season, while acknowledging that other objective approaches (e.g., comparing monthly patterns rather than time series) might yield different definitions. Throughout the remainder of the study, only seasonal mean values are used, and any decadal filtering is applied to seasonal rather than monthly values.

A similar analysis of a "Gibraltar" minus "Iceland" NAO index from the model simulation CON (formed by differencing the SLP from the nearest grid boxes to the stations involved, indicated in Fig. 1f–j), yields the correlations shown in parentheses in Table B1. The interannual correlations between adjacent months are of similar magnitude to those observed, although correlations between non-adjacent months are smaller. Decadal smoothing does not raise the inter-monthly correlations in HadCM2 as it did for the observed record; this implies that there is some controlling factor in the real world that is acting to co-ordinate the decadal behaviour of the NAO across the winter months, but is missing from HadCM2. This helps to explain why the HadCM2 NAO index values for individual months show interdecadal variability that is as strong as is observed, but the lower correlations between the simulated months yield weaker than observed low frequency variability of the DJFM-mean NAO (see also Sect. 5.1).

Appendix C: definition of an NAO index

The North Atlantic Oscillation is effectively a shift in atmospheric mass from the region of the Iceland Low to that of the Azores High (positive phase of the index) or vice versa (negative phase). The DJFM sea level pressure timeseries at the Azores and at SW Iceland are, therefore, negatively correlated ($r = -0.63$), and their difference (A-I) could be used as an index of the NAO (Fig. 2a). Taking a station further east for the southern centre-of-action (e.g., Hurrell and van Loon 1997 used Lisbon; Jones et al. 1997 used Gibraltar) can yield a stronger winter signal (i.e. inverse correlation: $r = -0.68$

Table C1 Correlations (×100) between the five observed NAO indices considered in Appendix C, between raw time series below the diagonal and decadally-filtered above the diagonal. Also included are correlations between these indices and the time series of the *second* principal component of the gridded SLP (PC2)

	A-I	G-I	G-Igrid	PC1	rotPC1	PC2
A-I	—	95	93	88	95	52
G-I	95	—	95	91	92	58
G-Igrid	95	98	—	83	94	63
PC1	84	88	85	—	85	41
rotPC1	96	93	94	82	—	57
PC2	29	29	30	0	36	—

for DJFM), although not for the other seasons (Jones et al. 1997). The Gibraltar minus SW Iceland SLP series (G-I) is, therefore, slightly preferable to A-I (*and* the G-I series is also 43 y longer, Fig. 2b). Nevertheless, both G-I and A-I time series are well correlated at interannual and interdecadal time scales (Table C1).

In the present study, we use the absolute pressure difference between the two selected stations; the alternative (followed in most earlier studies) is to scale (normalise) each series so that it has zero mean and unit variance and then to difference them. The reason for our use of the absolute difference is that normalisation could hide errors in the model simulation. The difference between the normalised and absolute difference indices is not critical: they are highly correlated, and both explain similar amounts of variability of Northern Hemisphere gridded fields of SLP (both 17.2%), surface temperature (10.8% without normalisation, 11.3% with) and precipitation (3.0% without, 3.2% with).

To compare the G-I index based on station time series with an equivalent index based on gridded SLP data (which is how the index has to be defined from the model data), a G-Igrid index is derived by differencing the observed time series from the grid boxes nearest to Gibraltar and Iceland (positions marked on Fig. 1a-e) in the observed SLP dataset. This also provides a brief assessment of the gridded SLP dataset. While G-I and G-Igrid are very highly correlated (Table C1), there is a slight deviation pre-1895, when G-Igrid (Fig. 2c) was consistently lower than G-I (Fig. 2b).

Principal component analysis (PCA) provides an alternative method of defining the NAO (Rogers 1990) that includes more information than the pressure at only two points. It is limited, however, by the availability of the gridded SLP dataset (1873 to 1995, although for the PCA we also omit years 1883, 1916–21 and 1939 due to the extensive missing data in these years). To focus on the variability of the North Atlantic (but without constraining model and observed patterns to be similar by the use of too small a region), all principal component analyses use only DJFM SLP data from the region 20 °N to 80 °N and 110 °W to 70 °E, whether observed or simulated (we use only a 200-y period of the HadCM2 control integration).

As discussed in Sect. 4, PCA of the winter SLP in this region yields a leading EOF that captures the NAO signal (Fig. 1c,h for the observed and HadCM2 EOFs, respectively). The time series (PC1: Fig. 2d) corresponding to the leading observed EOF shows a similar, but not identical, evolution to the station-based indices (Fig. 2a–c). The correlations vary from 0.83 to 0.91 between PC1 and the other indices, for raw or decadally-filtered time series (Table C1). The differences are clearer on the decadal time scale (thick line in Fig. 2d), with the 1873–1895 low NAO period and the 1902–1922 high NAO period less prominent in PC1, making the trend from the 1960s to the 1990s particularly prominent relative to the other variability.

The differences between the station-based NAO indices and PC1 come about because other modes of variability have some power over the station locations. In particular, the second EOF modes of both the observed and simulated datasets (Fig. 1d,i) capture some of the SLP variability over Iceland (and from a validation point of view, note the similarity between the second modes of SLP variability and that they each explain 14% of the variability in their respective datasets). The influence of other modes of variability over Iceland and the southern stations demonstrates a disadvantage of using a station-based NAO index rather than a pattern-based index such as PC1.

Further comparison, however, reveals that the time series (PC2: Fig. 2f) corresponding to the second observed EOF explains some of the low frequency variability exhibited by the station-based NAO indices (see also Table C1) and, while PC1 and PC2 are uncorrelated (due to their orthogonality), if they are each decadally filtered then they become significantly correlated ($r = 0.41$; significance is only marginal at the 90% level due to the autocorrelation in the filtered time series). These results highlight the need to perform PCA on the time scale of interest. For example, if the observed gridded SLP data are temporally filtered with low-pass decadal filters on a point-wise basis, prior to PCA, then the leading EOF (explaining 43% of the filtered SLP variability in the PCA region) combines some of the temporal and spatial aspects of PC1 and PC2 of the unfiltered SLP. A similar result is also obtained by applying a Varimax (Kaiser 1958) rotation to the leading four EOFs of the unfiltered SLP (similar in that the leading mode after rotation also appears to combine some of the temporal and spatial aspects of the unrotated PC1 and PC2 of the unfiltered SLP data). The station-based NAO indices correlate more strongly (Table C1) with PC1 or with the time series corresponding to the leading mode after rotation (rotPC1: Fig. 2e) than does PC1 with rotPC1. Spatially, the rotPC1 pattern (Fig. 1e) is similar to that of PC1 (Fig. 1c) but with a westward shift of the Azores region and a southward shift and retreat from northern Eurasia of the Iceland region, changes that are evidently derived from the PC2 pattern (Fig. 1d). The HadCM2 simulation shows a similar change when rotation is applied (Fig. 1j).

The difference between the two pattern-based indices (PC1 and rotPC1) is, therefore, as large over the instrumental period as the uncertainty introduced by using a station-based index. The choice of NAO index does not appear to be critical. In that case, the choice is best determined by the length of available record; the absolute pressure difference between Gibraltar and SW Iceland (G-I) as derived from Jones et al. (1997), is superior in that respect. We also applied the approach of Santer (1988) to measuring the strength of the Iceland Low and the Azores High, which allows the construction of an NAO index with moving centres-of-action. It was successful for the observations but not for HadCM2 (because in the latter, the centres of the oscillation were slightly offset from the absolute pressure centres).

Note that further intercomparisons of the various index time series shown in Fig. 2 might identify larger differences between them. For example, the 'redness' of the various time series (defined as the ratio of the standard deviation of low-pass-filtered time series to the standard deviation of high-pass-filtered time series, with filter cutoffs of 5, 10, 15 or 20 y) varies by up to 35% between definitions. The selected index (G-I) is one of the least 'red' (see also Sect. 5.1). The reddest is the time series associated with the leading EOF after rotation (rotPC1); this is, perhaps, not surprising given our earlier comment regarding the similarities between this rotated mode and the leading EOF of the temporally-smoothed SLP dataset (which we would expect, by definition, to have a particularly red time series). Nevertheless, this result does not alter our choice of NAO index since the redder indices may be measuring a somewhat different phenomenon to the NAO. Kushnir (1994), for example, identifies differences between the atmospheric circulation patterns involved in interannual and interdecadal variability; the latter has a southern centre-of-action further to the west than the former. This agrees with the intercomparison of NAO indices considered here, since rotPC1 has its southern centre-of-action to the west of that of PC1 (Fig. 1e compared with 1c) and rotPC1 has the redder time series of the two, while A-I uses a more western southern station than does G-I and A-I is the redder of these two time series.

There is the potential to augment the instrumental NAO index with paleo-based reconstructions (Cook et al. 1998; Appenzeller et al. 1998; Stockton and Glueck 1999), to provide additional information at the longer time scales and for further consideration of how unusual the recent observed changes are. Here we consider only the reconstruction by Cook et al. (1998) back to 1701 which was calibrated and validated using the 1874–1980 NAO index of Rogers (1984), taking the difference between the normalised DJF SLP from the Azores and Iceland. The instrumental extension by Jones et al. (1997), back to 1823/4, provides us with the opportunity to further test the Cook et al. (1998) reconstruction. Over the period 1874–1980, the Cook et al. (1998) and Jones et al. (1997) indices correlate at 0.56, whereas over the completely independent 1824–1873 period this falls to 0.20 (with this residual correlation entirely due to skill at time scales shorter than a decade). In terms of explained variance this is inadequate for comparison with the model simulation, although we do not know whether it is the reconstruction or the early instrumental data that are in error. The reconstructed NAO index is not, therefore, used further.

References

Appenzeller C, Stocker TF, Anklin M (1998) North Atlantic Oscillation dynamics recorded in Greenland ice cores. Science 282: 446–449

Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. Mon Weather Rev 115: 1083–1126

Basnett TA, Parker DE (1997) Development of the global mean sea level pressure data set GMSLP2. Climate Research Technical Note 79, Hadley Centre, UK, 16pp

Carnell RE, Senior CA (1998) Changes in mid-latitude variability due to increasing greenhouse gases and sulphate aerosols. Clim Dyn 14: 369–383

Cayan D (1992) Latent and sensible heat flux anomalies over the northern oceans: the connection to monthly atmospheric circulation. J Clim 5: 354–369

Cook ER, D'Arrigo RD, Briffa KR (1998) A reconstruction of the North Atlantic Oscillation using tree-ring chronologies from North America and Europe. Holocene 8: 9–17

Davies JR, Rowell DP, Folland CK (1997) North Atlantic and European seasonal predictability using an ensemble of multi-decadal atmospheric GCM simulations. Int J Climatol 17: 1263–1284

Dickson RR, Lazier J, Meincke J, Rhines P, Swift J (1996) Long-term coordinated changes in the convective activity of the North Atlantic. Prog Oceanogr 38: 241–295

Glowienka-Hense R (1990) The North Atlantic Oscillation in the Atlantic-European SLP. Tellus 42A: 497–507

Grotzner AM, Latif M, Barnett TP (1998) A decadal climate cycle in the North Atlantic ocean as simulated by the ECHO coupled GCM. J Clim 11: 831–847

Huang J, Higuchi K, Shabbar A (1998) The relationship between the North Atlantic Oscillation and El Nino-Southern Oscillation. Geophys Res Lett 25: 2707–2710

Hulme M (1994) Validation of large-scale precipitation fields in general circulation models. In: Desbois M, Desalmand F (eds) Global precipitations and climate change, Springer-Verlag, Berlin, Heidelberg, New York, pp 387–405

Hurrell JW (1995) Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. Science 269: 676–679

Hurrell JW (1996) Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. Geophys Res Lett 23: 665–668

Hurrell JW, van Loon H (1997) Decadal variations in climate associated with the North Atlantic Oscillation. Clim Change 36: 301–326

Johns TC, Carnell RE, Crossley JF, Gregory JM, Mitchell JFB, Senior CA, Tett SFB, Wood RA (1997) The second Hadley Centre coupled ocean-atmosphere GCM: model description, spinup and validation. Clim Dyn 13: 103–134

Jones PD (1987) The early twentieth century Arctic high – fact or fiction? Clim Dyn 1: 63–75

Jones PD (1994) Hemispheric surface air temperature variations: a reanalysis and an update to 1993. J Clim 7: 1794–1802

Jones PD, Jónsson T, Wheeler D (1997) Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and southwest Iceland. Int J Climatol 17: 1433–1450

Jones PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperature and its variations over the last 150 years. Rev Geophys 37: 173–199

Kaiser HF (1958) The Varimax criterion for analytic rotation in factor analysis. Psychometrika 23: 187–200

Kushnir Y (1994) Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. J Clim 7: 141–157

Leggett J, Pepper WJ, Swart RJ (1992) Emissions scenarios for IPCC: an update. In: Houghton JT, Callander BA, Varney SK (eds) Climate change 1992: the supplementary report to the IPCC scientific assessment, Cambridge University Press, Cambridge, pp 69–95

Mitchell JFB, Johns TC, Eagles M, Ingram WJ, Davis RA (1999) Towards the construction of climate change scenarios. Clim Change 41: 547–581

Parker DE, Jones PD, Folland CK, Bevan A (1994) Interdecadal changes of surface temperature since the late nineteenth century. J Geophys Res 99: 14373–14399

Parker DE, Folland CK, Jackson M (1995) Marine surface temperature: observed variations and data requirements. Clim Change 31: 559–600

Perlwitz J, Graf H-F (1995) The statistical connection between tropospheric and stratospheric circulation of the Northern Hemisphere winter. J Clim 8: 2281–2295

Rodwell MJ, Rowell DP, Folland CK (1999) Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. Nature 398: 320–323

Rogers JC (1984) The association between the North Atlantic Oscillation and the Southern Oscillation in the Northern Hemisphere. Mon Weather Rev 112: 1999–2015

Rogers JC (1990) Patterns of low-frequency monthly sea level pressure variability (1899–1986) and associated wave cyclone frequencies. J Clim 3: 1364–1379

Rogers JC (1997) North Atlantic storm track variability and its association to the North Atlantic Oscillation and climate variability of northern Europe. J Clim 10: 1635–1647

Santer BD (1988) Regional validation of general circulation models. PhD Thesis, University of East Anglia, Norwich, UK, 375pp

Saravanan R (1998) Atmospheric low-frequency variability and its relationship to midlatitude SST variability: studies using the NCAR climate system model. J Clim 11: 1386–1404

Serreze MC, Carse F, Barry RG, Rogers JC (1997) Icelandic Low cyclone activity: climatological features, linkages with the NAO, and relationships with recent changes in the Northern Hemisphere circulation. J Clim 10: 453–464

Stockton CW, Glueck MF (1999) Long-term variability of the North Atlantic Oscillation. In: Proc 10th Symp on global change studies, January 1999, Dallas, Texas, American Meteorological Society, Boston, MA, pp 290–293

Tett SFB, Johns TC, Mitchell JFB (1997) Global and regional variability in a coupled AOGCM. Clim Dyn 13: 303–323

Thompson DWJ, Wallace JM (1998) The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. Geophys Res Lett 25: 1297–1300

Timmermann A, Latif M, Voss R, Grotzner A (1998) Northern Hemisphere interdecadal variability: a coupled air-sea mode. J Clim 11: 1906–1931

Trenberth KE (1997) The definition of El Nino. Bull Am Meteorol Soc 78: 2771–2777
Trenberth KE, Hurrell JW (1994) Decadal atmosphere-ocean variations in the Pacific. Clim Dyn 9: 303–319
van Loon H, Rogers JC (1978) The seesaw in winter temperatures between Greenland and Northern Europe, part I: general description. Mon Weather Rev 106: 296–310
Walker GT (1924) Correlations in seasonal variations of weather IX. Mem Ind Meteorol Dept 24: 275–332
Wallace JM, Zhang Y, Renwick JA (1995) Dynamic contribution to hemispheric mean temperature trends. Science 270: 780–783
Xie P, Arkin PA (1997) Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. Bull Am Meteorol Soc 78: 2539–2558