Climate Dynamics (2004) 22: 605–623
DOI 10.1007/s00382-004-0405-1

T. J. Osborn

# Simulating the winter North Atlantic Oscillation: the roles of internal variability and greenhouse gas forcing

Received: 28 July 2003 / Accepted: 28 January 2004 / Published online: 4 May 2004


**Abstract** Analysis of simulations with seven coupled climate models demonstrates that the observed variations in the winter North Atlantic Oscillation (NAO), particularly the increase from the 1960s to the 1990s, are not compatible with either the internally generated variability nor the response to increasing greenhouse gas forcing simulated by these models. The observed NAO record *can* be explained by a combination of internal variability and greenhouse gas forcing, though only by the models that simulate the strongest variability and the strongest response. These models simulate inter-annual variability of the NAO index that is significantly greater than that observed, and can no longer explain the observed record if the simulated NAO indices are scaled so that they have the same high-frequency variance as that observed. It is likely, therefore, that other external forcings also contributed to the observed NAO index increase, unless the climate models are deficient in their simulation of inter-decadal NAO variability or their simulation of the response to greenhouse gas forcing. These conclusions are based on a comprehensive analysis of the control runs and transient greenhouse-gas-forced simulations of the seven climate models. The simulations of mean winter circulation and its pattern of inter-annual variability are very similar to the observations in the Atlantic half of the Northern Hemisphere. The winter atmospheric circulation response to increasing greenhouse gas forcing shows little inter-model similarity at the regional scale, and the NAO response is model-dependent and sensitive to the index used to measure it. At the largest scales, however, sea level pressure decreases over the Arctic Ocean in all models and increases over the Mediterranean Sea in six of the seven models, so that there is an increase of the NAO in all models when measured using a pattern-based index.

T. J. Osborn
Climatic Research Unit, School of Environmental Sciences,
University of East Anglia, Norwich, NR4 7TJ, UK
E-mail: t.osborn@uea.ac.uk

## 1 Introduction

The North Atlantic Oscillation (NAO, Walker 1924; van Loon and Rogers 1978) has been an important driver of circum-Atlantic climate variability during the extended boreal winter, especially over recent decades (Hurrell 1995; Wanner et al. 2001). The NAO is relevant, therefore, to seasonal predictability during the wintertime in these regions (e.g. Rodwell et al. 1999). The fact that it has also exhibited strong multi-decadal variations during the twentieth century means that the NAO has also become relevant to climate change issues (e.g. Hurrell 1996; Osborn et al. 1999; Gillett et al. 2000, 2003). Figure 1a shows a winter NAO index time series derived from Jones et al. (1997), updated through winter 2002/3 (data courtesy of Phil Jones). The multi-decadal variations that have been strong since 1900 are clear, including the strong trend from the low-index 1960s to the high-index early 1990s. Recent winters have an average NAO index that is only slightly above the 1961–1990 mean, and the trend from the 1960s to the early 1990s has not continued. Nevertheless, it is important to understand the causes of the multi-decadal variations, particularly whether they are related to anthropogenic forcing of climate or are an expression of natural climate variability.

Results from an analysis of forced and unforced coupled climate model simulations are reported here, focusing on the two issues of (1) whether the internally generated variability exhibited by the climate models is sufficiently strong to be a possible explanation of recent NAO changes; and (2) whether these climate models' responses to increasing greenhouse gas concentrations include a change in the NAO index that might have contributed to recent changes and that might continue

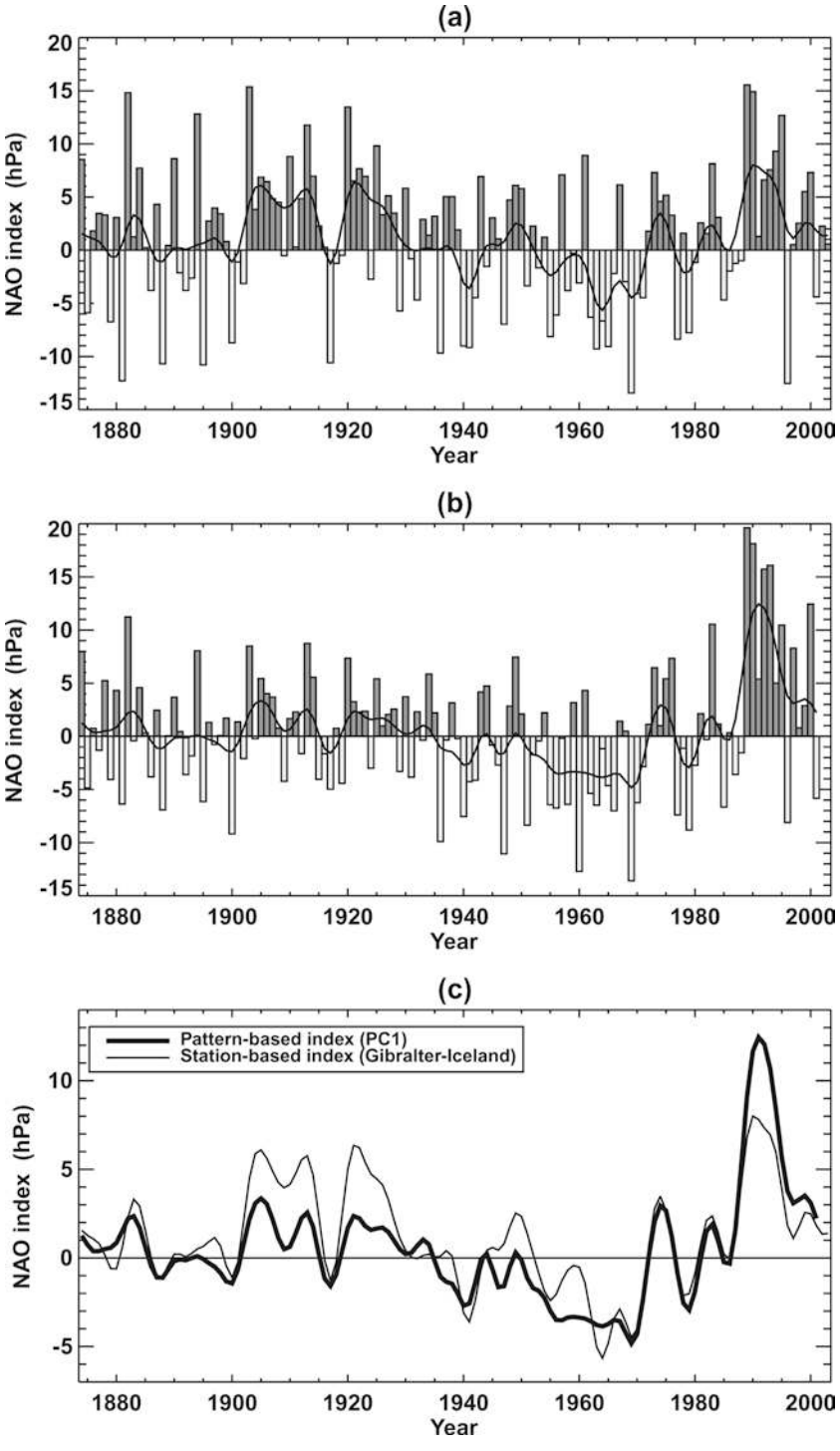


**Fig. 1** Observed winter NAO indices: **a** anomalous Gibraltar minus Iceland SLP (hPa), and **b** PC time series scaled to represent maximum pressure difference anomaly (hPa). *Bars* show individual winter values. *Curves* show decadally-smoothed values, which are repeated for comparison in panel **c**

into the future. Analyses along these lines have been undertaken before (e.g. Osborn et al. 1999; Ulbrich and Christoph 1999; Fyfe et al. 1999; among others, see review by Gillett et al. 2003), but the advance made here is that seven different climate models (Table 1) have been utilised, with very similar experiments being available from each, and all models are analysed using exactly the same methods. This enables a multi-model comparison that is quantitative rather than just qualitative in nature. Results from the average of the seven model analyses have been published in Gillett et al. (2003), but full results from the individual models are presented here for the first time, together with additional analyses of changes in the characteristics of inter-annual NAO variability under enhanced greenhouse forcing. Stephenson and Pavan (2003) have also published an analysis of NAO variability in a multiple-model comparison, but based only on control simulations and using

**Table 1** Climate model acronyms, references, resolution of atmospheric component and simulation lengths used (years)

| Acronym | Centre and reference | Resolution | | | Simulations | |
|---|---|---|---|---|---|---|
| | | Horizontal[a] | Vertical | Stratosohere[b] | Control | 1%/year |
| CCSR/NIES | CCSR and NIES, Japan Emori et al. (1999) | T21 (finer ocean) | 20 levels | medium | 210 | 1890–2099 |
| CGCM1 | CCCMA, Canada Flato et al. (2000) | T32 (finer ocean) | 10 levels | low | 200 | 1900–2099 |
| CSIRO Mk2 | CSIRO, Australia Gordon and O'Farrell (1997) | R21 | 9 levels | low | 210 | 1890–2099 |
| ECHAM4 | DKRZ, Germany Bacher et al. (1998) | T42 | 19 levels | medium | 240 | 1870–2099 |
| HadCM2 | Hadley Centre, Met Office, UK Johns et al. (1997) | 3.75°, 2.5° | 19 levels | medium | 1400 | 1860–2099 (4 runs) |
| HadCM3 | Hadley Centre, Met Office, UK Gordon et al. (2000) | 3.75°, 2.5° (finer ocean) | 19 levels | medium | 240 | 1860–2099 |
| NCAR PCM | NCAR, USA Washington et al. (2000) | T42 (finer ocean) | 18 levels | medium | 300 | 1961–2098 |

[a]Horizontal: finite difference model resolution in degrees of longitude then latitude; spectral model resolutions are approximately equivalent to: T21 ≈ 5.6°, 5.6°; T32 ≈ 3.8°, 3.8°; T42 ≈ 2.8°, 2.8°; R21 ≈ 5.6°, 3.2°

[b]Stratosphere: a qualitative description of model resolution of the stratosphere, based on the number of model levels above 200 hPa

surface air temperature patterns to define the NAO. Thus, the study reported here is complementary to their paper and both are necessary to obtain a comprehensive view of current model simulations of the NAO.

The observed and simulated data are described in Sect. 2, together with methods for measuring the NAO in a consistent way. Recent variations in the observed NAO index are considered in Sect. 3, followed by a comparison with the internally-generated NAO variability of the climate models in Sect. 4. The influence of increasing greenhouse gas forcing is reported in Sect. 5, considering first changes in the mean NAO index and then changes in the spatial structure and temporal variability of winter circulation in the perturbed climate states. Finally, the implications of these results and the possible contribution of other forcing factors are discussed in Sect. 6.

## 2 Data and methods

### 2.1 Observed data

Monthly mean sea level pressure (SLP) from 1873 to 2001 on a 5° latitude by 10° longitude grid covering much of the Northern Hemisphere were used. This data set is derived from the UK Met Office analyses (Jones 1987; see Basnett and Parker 1997 for further discussion, though note that their GMSLP data set was not used here because it does not extend beyond 1998). A December to March (DJFM) average is formed from the monthly data to provide the mean winter SLP used here (this four month seasonal mean is used throughout this study for reasons outlined by Osborn et al. 1999). All winters are referenced in the figures and text by the year in which the January falls (i.e. 1874 is the winter from December 1873 to March 1874).

### 2.2 Simulated data

Monthly sea level pressure fields, averaged across the December to March season, were taken from the seven different coupled ocean-atmosphere climate models listed (with acronyms and references) in Table 1. Data from five of these models were obtained from the Intergovernmental Panel on Climate Change (IPCC) Data Distribution Centre (http://ipcc-ddc.cru.uea.ac.uk/), while data from the HadCM2 and HadCM3 simulations were obtained from the Climate Impacts LINK Project (http://www.cru.uea.ac.uk/link/); all the modelling centres are gratefully acknowledged for allowing their data to be distributed for research use.

For each model, data were taken from control integrations with unchanging external forcing, varying in length from 200 to 1400 years (Table 1). Winter SLP was also taken from integrations with increasing greenhouse gas (GHG) concentrations, forced with historic levels prior to 1990 and thereafter by a compounded 1% per year increase in effective carbon dioxide. For one model (HadCM2), an ensemble of four such integrations was available, each with identical forcing but begun from different initial conditions (Mitchell et al. 1999).

Table 1 also lists brief details about the resolution of each model, including some qualitative information about stratospheric resolution, because stratospheric processes may be important for simulating variations in the NAO (Shindell et al. 1999). As indicated, none of these models have sufficient levels to resolve the stratosphere in detail.

Not all models (notably HadCM2 and HadCM3) conserve mass perfectly. This can manifest itself as a trend in global-mean SLP (see Fig. 1a of Osborn et al. 1999). In all models and simulations used here, any such artificial variations in globally averaged sea level

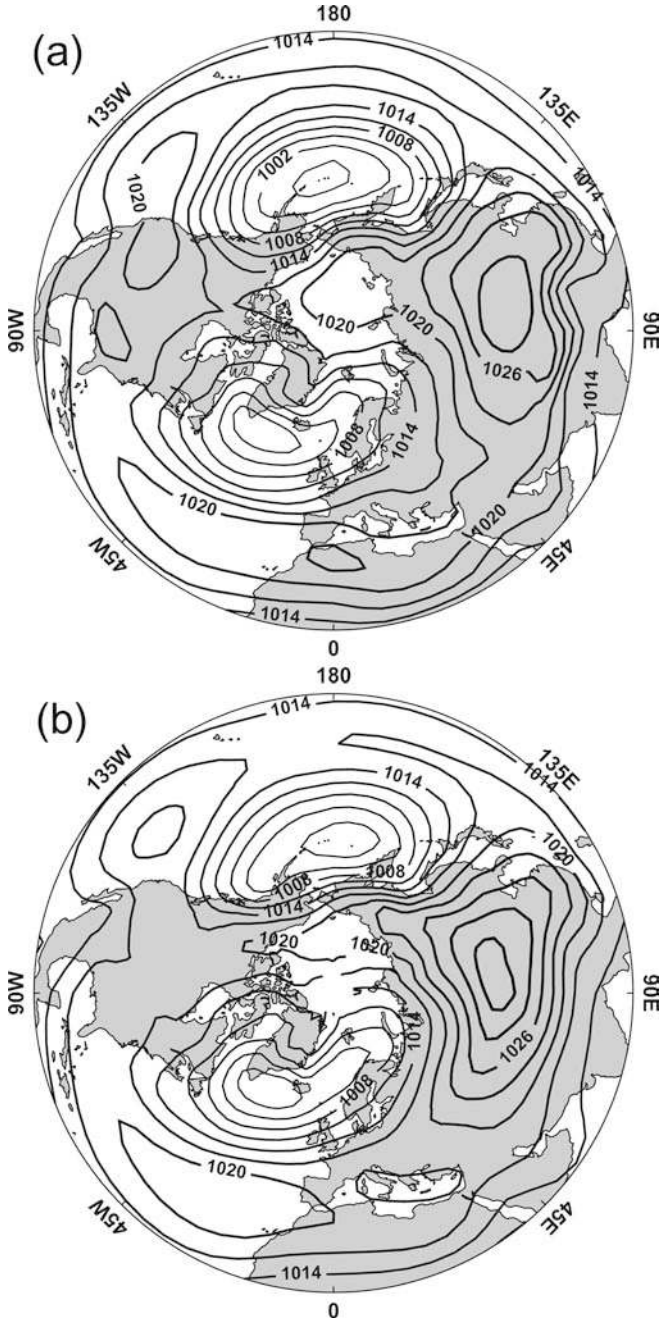


**Fig. 2** Long-term mean December-March SLP from **a** average of seven control simulations, and **b** observations. Contour interval is 3 hPa

pressure are removed by making a geographically uniform adjustment to the pressure at each grid box.

## 2.3 Measuring the North Atlantic Oscillation

Osborn et al. (1999) discuss and illustrate (see their Fig. 2) a range of indices of the observed NAO, and find that they are all quite highly correlated (in the range 0.83 to 0.98). Two distinct approaches are to use (1) pressure differences between station pairs; or (2) a pattern-based measure of the NAO. The station-based approach has the advantage of producing a longer observed record; for the Gibraltar and Iceland stations, Jones et al. (1997) produce a winter series from 1823 to the present (Fig. 1a shows a post-1874 series derived from this record). A disadvantage is that circulation variability that is unrelated to the NAO (i.e. with a different spatial structure) can have an influence on either or both stations used, and hence contaminate the NAO index with other signals. Osborn et al. (1999) point this out when analysing the HadCM2 simulations under enhanced greenhouse forcing: the simulated NAO index decreases during the simulation, but it is noted that this is due to a long-term trend pattern in atmospheric SLP that does *not* resemble the NAO pattern.

Because of these disadvantages with the station-based index, a pattern-based approach is mainly used here, identifying the pattern with principal component analysis (PCA) and then projecting time series of SLP anomaly fields on to this pattern (i.e. computing the scalar or dot product between field and pattern) to obtain an index time series. PCA (using the covariance matrix and with no rotation) is always performed on the Atlantic half (110°W to 70°E) of the Northern Hemisphere SLP field and the leading empirical orthogonal function (EOF) is retained as the NAO pattern. If SLP from the entire hemisphere had been used, then the resulting pattern might resemble the Arctic Oscillation (Thompson and Wallace 1998) rather than the NAO. Nevertheless, it is of interest to know whether the NAO is related to anomalous atmospheric circulation in the Pacific half of the hemisphere. Having obtained an Atlantic EOF pattern and corresponding principal component (PC) time series, a pattern covering the entire hemisphere is generated by computing the regression slope coefficients between the PC time series and the time series of winter SLP from each point in the data set, and then scaling this pattern of regression slopes to resemble an EOF (see discussion of scaling later).

One possible approach would be to use the same pattern to define the NAO in all data sets, perhaps the EOF of the observed data or from one particular model simulation (e.g. Gillett et al. 2001). A disadvantage is that the EOF used will be optimised to one particular data set (explaining, by definition, the maximum amount of variance for that data set), but will not be the optimal pattern for the other simulated data sets, resulting in an artificial suppression of their temporal variance. The alternative, used in the present study, is to use the EOF from each model simulation to define the NAO index for that simulation. This makes allowance for small variations in the EOF patterns between simulations, though the comparison would lose its validity if the pattern differences were large. Using different patterns requires care to be taken with the scaling of the EOF patterns, to ensure that the associated principal component time series can be compared on a like-with-like basis.

In this study, EOF patterns are scaled so that they are unit vectors (i.e. the sum of all squared values is equal to one) and the corresponding PC time series are scaled so

that their value at time $t$ is equal to the dot product of the EOF pattern and the SLP field at time $t$ (and, thus, the original SLP data set can be reconstructed by multiplying each EOF pattern by its PC time series, and summing over all EOFs). This scaling is inappropriate, however, if the number of space points in one data set is higher than in another (i.e. due to different model resolutions, see Table 1) because the more points there are, the smaller the value they will be given to ensure that the sum of squared values still equals one. The smaller the values in the EOF pattern, the larger the values in the PC time series, thus hindering a comparison of the temporal variance of the PC time series without additional reference to the EOF patterns. Scaling could be applied relating to the density of grid points (and therefore the area of their grid boxes), but we instead follow the alternative of re-gridding (using Gaussian-weighted re-sampling) all model output data sets onto the coarser (5° latitude by 10° longitude) grid of the observed data set (which is adequate since we are only interested in the large-scale NAO phenomenon), prior to PCA. Thus a unit deviation in any PC time series corresponds to an SLP anomaly field whose root-mean-squared value is 1 hPa.

While the scaling described allows PC time series to be inter-compared in absolute units, equal PC time series deviations can still be related to different magnitudes of pressure gradient anomalies across the North Atlantic region. This is because, as shall be shown later, some climate models simulate a leading mode of variability (i.e. EOF) that has stronger positive and negative loadings in the Atlantic sector than other models. For example, a unit deviation in the HadCM2 PC time series corresponds to an increase or decrease in the Azores to Iceland pressure difference of 0.20 hPa, compared with 0.15 hPa for a unit deviation in the CGCM1 PC time series. In addition to the standard "EOF scaling", therefore, we also use "pressure difference scaling", scaling each PC time series by the difference between the maximum and minimum values in the Atlantic sector of its corresponding EOF pattern. Under this scaling, a unit deviation in any PC time series corresponds to a maximum pressure difference anomaly of 1 hPa.

### 2.4 Estimating the range of internal variability

For each of the three ways of measuring the NAO (Gibraltar minus Iceland SLP, or principal component time series with either EOF scaling or pressure difference scaling), the observed variations are compared with the range of variability simulated during each of the seven model control integrations. The focus here is on 30-year trends in the NAO index, since Osborn et al. (1999) and others have already suggested that 30-year trends beginning between 1960 and 1967 are unusually strong. From each model control integration, the distribution of 30-year trends is obtained by computing all overlapping trends during the run. From each of these distributions, the 2.5 and 97.5 percentiles are estimated, to obtain a range encompassing 95% of the internally generated NAO trends.

Some of the control simulations are only about 200 years long (Table 1), resulting in considerable uncertainty (and some bias) in the estimates of the 95% range of variability. Monte Carlo simulation has been used to quantify this, generating long white noise time series and attempting to estimate the 2.5 and 97.5 percentile 30-year trends from short sections of them. For a 200-year control run, these estimates could be between 30% lower and 36% higher than the true value, a range that is reduced to 25% lower to 29% higher for the 300-year NCAR PCM run, and to ±12% for the longest simulation (1400 years for HadCM2). The use of short control simulations also produces a bias towards too small a range of variability (similar to the bias in the sample standard deviation statistic when computed using a small sample). For a 229-year control run, 200 overlapping 30-year trends are computed, but because of the overlaps they are not independent. The fifth largest of the 30-year trends will not, therefore, be the fifth largest of 200 independent trends (thus representing an estimate of the 97.5 percentile), but will instead be the largest of some smaller number of effectively independent trends (thus representing something smaller than the 97.5 percentile). The Monte Carlo simulation demonstrates that even for the shortest control run used here, however, the bias is no more than 5% of the true value and it is ignored in this study.

## 3 Recent North Atlantic Oscillation variations

Figure 1 shows the variations in the winter NAO index since 1874, comparing an index based on the Jones et al. (1997) Gibraltar minus Iceland pressure difference with the principal component time series of the leading EOF of the Atlantic SLP field, under pressure difference scaling (see Sect. 2.3). The index presented in Fig. 1a uses the SLP difference computed in absolute values (i.e. the values are not normalised either before or after differencing) and then the mean difference over the 1961–1990 reference period is subtracted. The use of absolute rather than normalised values is preferred here, because normalisation would mask differences in standard deviation between different data sets. For a similar reason, the covariance matrix rather than the correlation matrix was used to generate the series in Fig. 1b.

The two series are well correlated ($r = 0.83$), but are certainly not identical. The difference is most important at the multi-decadal time scale (Fig. 1c): while both exhibit a similar upward trend from the 1960s to the early 1990s, the 1989–1995 period is only moderately higher than the 1903–1913 and 1920–1925 period in the station-based record, yet the recent values are much more prominent in the pattern-based record (which is more similar to the Arctic Oscillation series in this respect, Thompson and Wallace 1998). Ostermeier and Wallace (2003) show that this different behaviour is because large

trends in SLP during the 1920–1969 period are restricted to a smaller region (over the Atlantic) than the large trends that occurred since 1969; the pattern-based index reacts more strongly to the widespread trends of the latter period. This has immediate consequences for the unusualness of the recent period, though note that if the *trend* from the low NAO 1960s to the high NAO 1990s is considered, then this recent trend is indeed unusual in the context of the past 100 years or more, even in the station-based record (Osborn et al. 1999).

The trend from the 1960s to the early 1990s is of particular interest, due both to its magnitude and to the contemporaneous warming of Northern Hemisphere land in winter. The cause of each trend and their possible inter-relatedness are the subject of much current research, initiated in part by the studies of Hurrell (1995, 1996; but see also Fig. 6a of Osborn et al. 1999). It should be pointed out, however, that the NAO trend has not been continued through to the present. The Gibraltar minus Iceland pressure difference has been updated from Jones et al. (1997) by Phil Jones (personal communication, June 2003), to the winter of 2002/3 (Fig. 1a) and shows that the recent winters have not had such strongly positive NAO anomalies as the 1989–1995 period.

## 4 Internally generated variability of the North Atlantic Oscillation

### 4.1 Spatial structure of simulated North Atlantic variability

The first question to be asked is whether the climate models adequately simulate the mean present-day winter circulation in the Northern Hemisphere. Averaging the winter sea level pressure fields over the full length of each control integration, and then across all seven climate models, gives an overall indication of the simulated atmospheric circulation (Fig. 2a) and it can be seen to capture the major observed features (Fig. 2b). Listed in Table 2 are the pattern correlations and root-mean-squared (RMS) errors between the individual simulations and the observations. There is some inter-model variation in mean winter circulation, with HadCM3 showing the lowest pattern correlation ($r = 0.80$, due mainly to a weak Iceland Low and a weak and spatially noisy Siberian High) and CGCM1 showing the worst RMS error (5.4 hPa, due to greater amplitude of spatial variation than observed, with too strong subtropical highs and too deep Aleutian and Iceland Lows). The average of all seven model simulations outperforms any individual model (behaviour that has been noted before for other variables: Lambert and Boer 2001) with a pattern correlation of $r = 0.94$ and RMS error of just 2.0 hPa.

The leading mode of Atlantic-sector inter-annual variability, defined by the leading EOF of the SLP field (see Sect. 2.3 for details) explains between 40 and 66% of the variance in the Atlantic half of the Northern Hemisphere (Table 2), compared to 40% for the observed data. In Fig. 3, these EOF patterns are intercompared and also extended to show associated variations over the North Pacific Ocean using the procedure described in Sect. 2.3. In the Atlantic sector, all models show a north-south dipole of variability in the Atlantic sector, and therefore a North Atlantic Oscillation, though with some variation in the position or amplitude of the centres of action. When extended over the hemisphere, most models show anomalies south of the Aleutian Low that are

**Table 2** Mean circulation and NAO variability statistics from control and GHG simulations and observations

| Model | Mean SLP pattern | | Atlantic sector SLP EOF#1 | | | | Gibraltar – Iceland | |
|---|---|---|---|---|---|---|---|---|
| | $r$[a] | RMSE[b] (hPa) | % variance[c] | $r$[d] | $\sigma$_control[e] (hPa) | $\sigma$_GHG[f] (hPa) | $\sigma$_control[g] (hPa) | $\sigma$_GHG[h] (hPa) |
| Observations | – | – | 40 | – | 6.0 | – | 6.4 | – |
| CCSR/NIES | 0.81 | 4.8 | 66 | 0.88 | 11.9** | 12.2 | 10.2** | 9.5 |
| CGCM1 | 0.81 | 5.4 | 56 | 0.90 | 6.0 | 5.1 | 5.6 | 3.0** |
| CSIRO Mk2 | 0.88 | 3.5 | 40 | 0.94 | 5.2 | 5.5 | 5.5 | 4.8 |
| ECHAM4 | 0.92 | 2.7 | 55 | 0.97 | 8.0** | 6.8 | 8.1** | 6.2** |
| HadCM2 | 0.87 | 3.3 | 41 | 0.91 | 6.4 | 6.8 | 6.8 | 7.3 |
| HadCM3 | 0.80 | 3.6 | 48 | 0.96 | 6.3 | 6.5 | 6.8 | 6.4 |
| NCAR PCM | 0.84 | 4.7 | 53 | 0.90 | 8.9** | 9.4 | 7.2 | 6.6 |
| Mean of 7 models | 0.94 | 2.0 | – | 0.97 | – | – | – | – |

[a]Pattern correlation between simulated and observed mean SLP pattern, weighted by grid box area
[b]Root-mean-squared error between simulated and observed mean SLP pattern, weighted by grid box area
[c]Percent of Atlantic sector SLP variance captured by the leading EOF
[d]Pattern correlation between simulated and observed EOF pattern, weighted by grid box area
[e]Standard deviation of leading PC under pressure difference scaling; ** indicates significantly different to observations with 95% confidence
[f]Standard deviation of leading PC under pressure difference scaling from 2050–2099 of GHG
[g]Standard deviation of Gibraltar minus Iceland SLP difference; ** indicates significantly different to observations with 95% confidence
[h]Standard deviation of Gibraltar minus Iceland SLP difference from 2050–2099 of GHG; ** indicates significantly different to control with 95% confidence

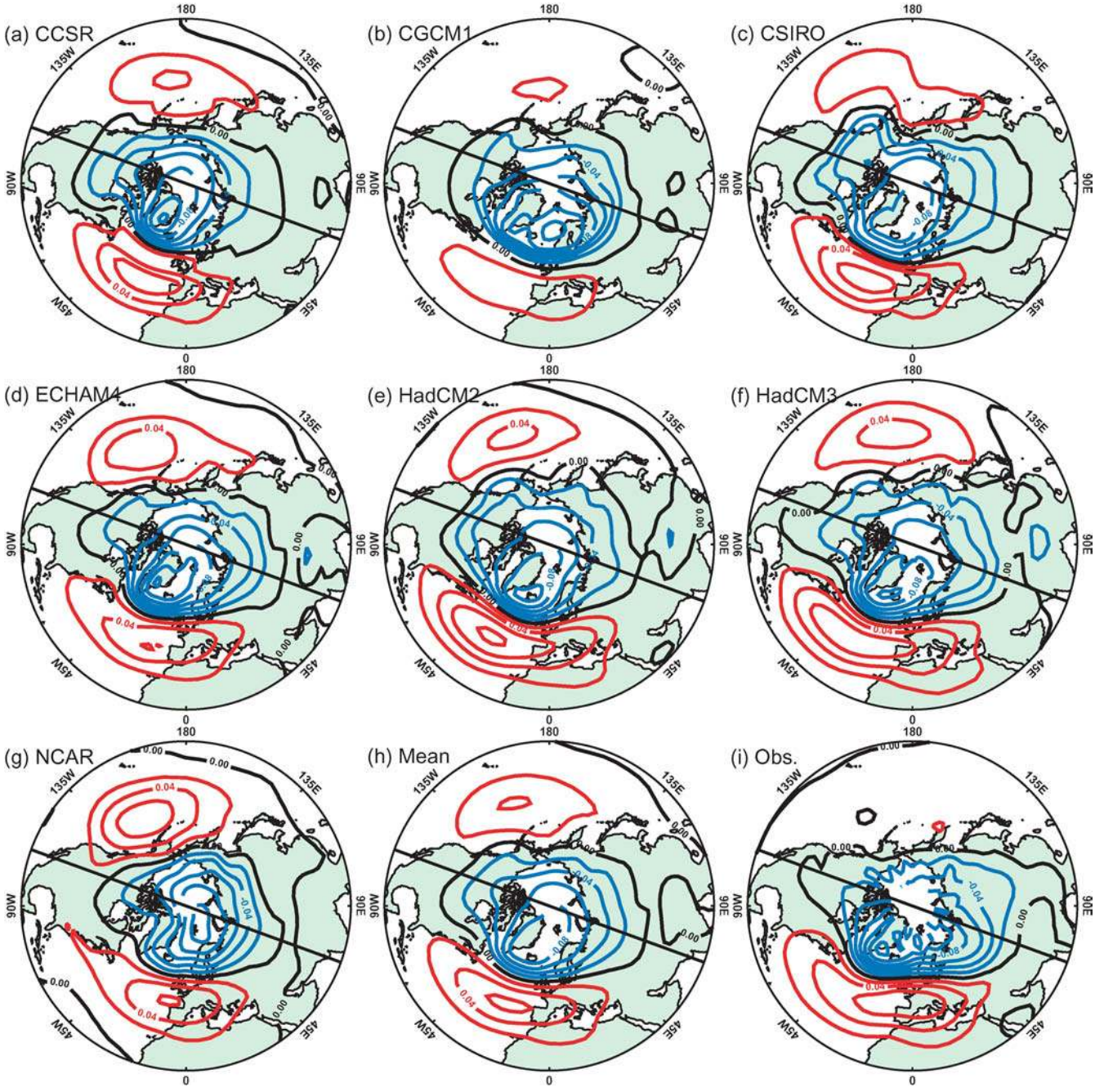


**Fig. 3** SLP patterns associated with the leading PC of Atlantic sector SLP variability in the control simulations of **a** CCSR/NIES, **b** CGCM1, **c** CSIRO Mk2, **d** ECHAM4, **e** HadCM2, **f** HadCM3, and **g** NCAR PCM. The average of patterns **a–f** is shown in **h**, and the observed pattern is shown in **i**. *Blue contours* are negative, zero is *black*, and positive are *red*. The *black diameter line* marks the Atlantic half of the hemisphere from which the PC was defined

out of phase with the Iceland Low SLP. Thus, these models (notably CCSR/NIES, ECHAM4, HadCM2, HadCM3 and NCAR PCM; Figs. 3a, d, e, f, g) show an Arctic Oscillation type pattern (Thompson and Wallace 1998) even though the pattern is defined using only Atlantic-sector SLP. Indeed, the NCAR PCM simulation shows a southern centre of action that is stronger over the Pacific Ocean than over the Atlantic Ocean. The Pacific teleconnection is not present in the observed SLP (Fig. 3i), and represents the largest difference between observations and the mean of the seven model EOF patterns (Fig. 3h).

Differences between the model simulations and the observations, particularly the teleconnection with the North Pacific, but also seemingly more minor differences such as the zonal nature of the zero line in Fig. 3h compared with the southwest–northeast orientation of the observed Atlantic EOF zero line (Fig. 3i), may

reflect important errors in the models' physical and dynamical behaviour. For example, Castanheira and Graf (2003) show that the teleconnection between the North Atlantic and the North Pacific *is* present in the observed winter climate, but only during winter months with a strong northern polar vortex in the lower stratosphere (i.e. strong zonal westerlies). Some models (e.g. ECHAM4) are biased towards a too strong polar vortex and thus their climate nearly always reflects behaviour that is only observed during strong vortices. Castanheira and Graf (2003) point out that this may not be an explanation for all models, and that other explanations for differences in behaviour must be investigated.

The inter-model variation in the EOF patterns, and the differences between these patterns and the observed EOF (Fig. 3i), might raise concern about the comparability of the NAO indices defined by using each model's own EOF pattern. The NAO indices are defined using only the Atlantic sector of the patterns (marked on each panel of Fig. 3), and in this region the pattern correlations between simulated and observed EOF vary from 0.88 to 0.97 (Table 2). To determine the impact of these pattern variations, we compare the NAO index time series obtained by projecting the observed SLP data set onto the observed EOF (Atlantic sector only) with the time series obtained by projecting the same observed SLP data set onto each of the different model EOFs (Atlantic sector only). The temporal correlation between the observed EOF time series (i.e. first PC) and the observed projection onto each model EOF only ranges from 0.94 to 0.97. Thus the use of a slightly different pattern to define the NAO index for each model makes virtually no difference to the shape of resulting time series. The temporal standard deviation of the time series *is* affected by the EOF pattern used to define it, ranging from 88% to 98% of the observed standard deviation when model patterns are used instead of the observed EOF. That these are lower than 100% simply reflects the fact that the model patterns are not optimised to capture the maximum possible variance of the observed SLP, supporting the decision to use the EOF from each model simulation to define the NAO index for that simulation (similarly, the observed EOF is used to define the observed NAO index).

As discussed in Sect. 2.3, the difference between the maximum and minimum values in the Atlantic sector of the EOF patterns (Fig. 3) varies from about 0.17 (CGCM1 and CSIRO Mk2) to 0.21 (CCSR/NIES). The PC time series (i.e. the NAO index) for each model can be scaled (pressure difference scaling) so that a unit deviation corresponds to a maximum pressure difference anomaly of 1 hPa, rather than to an SLP anomaly field whose RMS value is 1 hPa (EOF scaling). The standard deviations of the PC time series under pressure difference scaling are compared in Table 2. The control simulations of CGCM1, CSIRO Mk2, HadCM2 and HadCM3 have similar interannual standard deviations to the observed NAO index, while CCSR/NIES, ECHAM4 and NCAR PCM all exceed the observed NAO standard deviation by a margin that is statistically significant with 95% confidence (using a two-tailed $F$ test). Using the Gibraltar minus Iceland SLP difference as an alternative NAO index yields a very similar result (also Table 2), except that, while the NCAR PCM index still has greater interannual variance than observed, it is no longer a statistically significant difference. The reason for this change is evident from Fig. 3g, because Iceland is not close to the minimum value in the NCAR PCM pattern, and thus the Gibraltar to Iceland pressure difference has less variance than the peak variance of the EOF pattern.

### 4.2 Internally generated North Atlantic Oscillation trends

Figure 4 shows trends in the observed NAO indices, computed in 30-year sliding windows, compared with the 2.5 and 97.5 percentiles of the control simulation distributions of 30-year trends. There is a wide scatter amongst model estimates of the percentile values, reflecting the difference in NAO index standard deviations (Table 2) and the uncertainty in estimating distributions of trends from limited samples of data (Sect. 2.4). Nevertheless, both pattern-based indices (under EOF scaling, Fig. 4b, and pressure difference scaling, not shown) have 30-year trends that exceed the 97.5 percentile of all seven models (1965–1994 and 1966–1995). A longer run of trends are significant when compared with the mean of the individual model control run 97.5 percentiles (see Fig. 2 of Gillett et al. 2003). No trend in the station-based index (Gibraltar minus Iceland SLP) exceeds all control run 97.5 percentiles (Fig. 4a), though some exceed all but the CCSR/NIES threshold.

The results in Fig. 4a,b appear to indicate only marginal significance of the recent observed NAO trends at the 95% confidence level. Significance is greatly enhanced, however, by two further considerations. First, later results will indicate that the expected sign of the NAO index trend under greenhouse gas forcing is positive. This applies to the pattern-based NAO indices for which the signal is of consistent sign across all seven models, but does not apply to the station-based NAO index for which the sign of the GHG signal is less consistent. For the pattern-based NAO indices, therefore, it is reasonable to apply a one-tailed test to positive trends rather than the two-tailed test applied in Fig. 4b; under a one-tailed test, the critical thresholds are the 95 percentiles rather than the higher 97.5 percentiles that are shown.

The second consideration is that the two models with the highest trend percentiles are those whose interannual NAO variability is significantly higher (two to three times the variance) than that observed (Table 2). It is likely that this excess variance has resulted in an overestimated frequency of strong trends in these

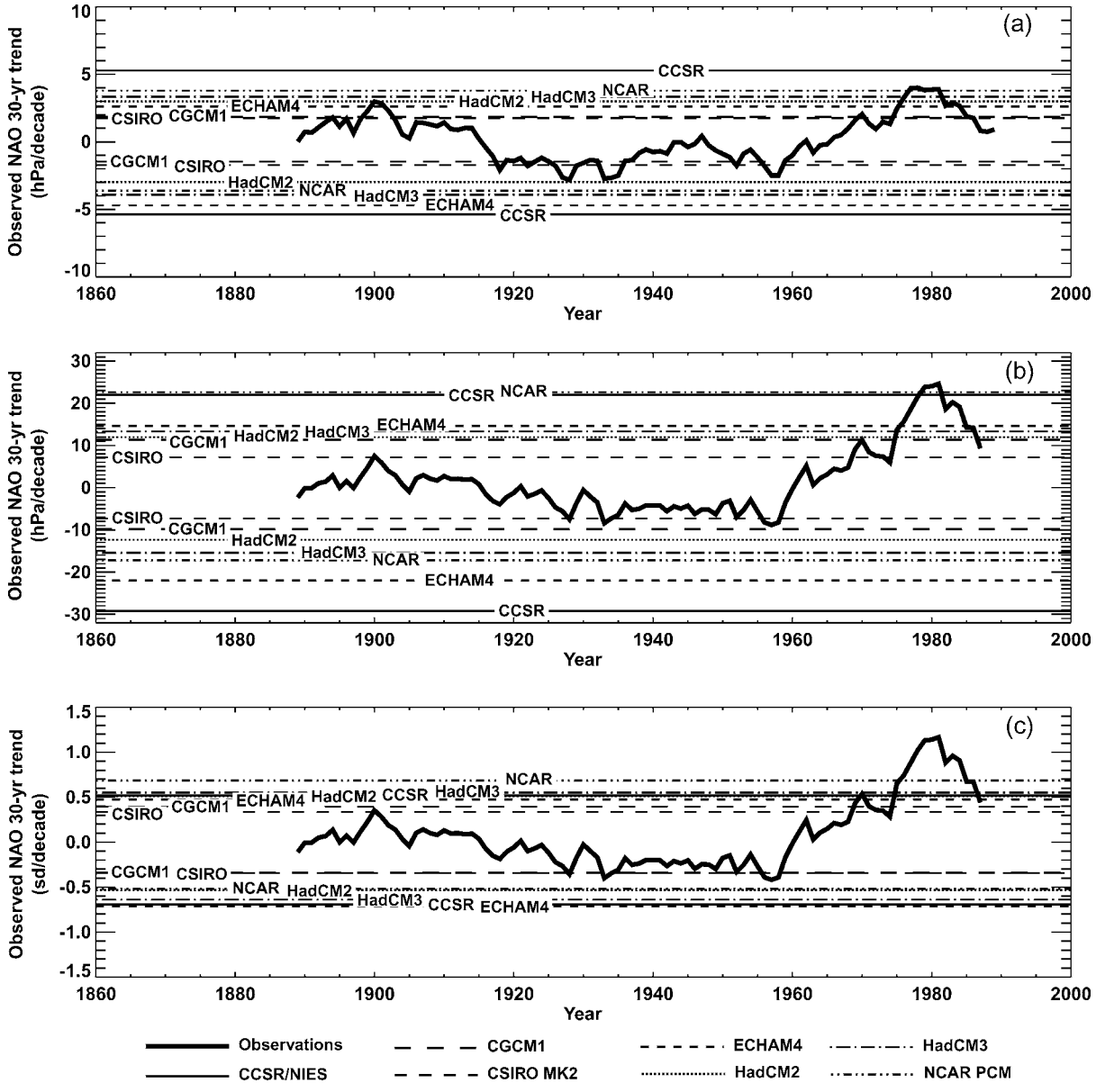


**Fig. 4** The 30-year trends (*thick lines*) in the observed NAO index time series, computed in a sliding window and plotted against the central year of the window, for **a** Gibraltar minus Iceland SLP difference, **b** pattern-based index under EOF scaling, and **c** pattern-based index scaled so that 10-year high-pass filtered series has unit variance. Units are "hPa (decade)$^{-1}$" in **a–b** and "standard deviations (decade)$^{-1}$" in **c**. *Horizontal lines* indicate the 2.5 and 97.5 percentiles of the distributions of 30-year trends of the equivalent indices simulated during the control simulations of CCSR/NIES (*solid*), CGCM1 (*long dashed*), CSIRO Mk2 (*medium dashed*), ECHAM4 (*short dashed*), HadCM2 (*dotted*), HadCM3 (*dot-dashed*) and NCAR PCM (*dot-dot-dashed*)

model simulations. The analysis has been repeated with all NAO index time series from observations and models first scaled so that when they are filtered with a 10-year high-pass filter, they all have unit variance. Fig. 4c shows the result for the pattern-based index. All trends between 1961–1990 and 1970–1999 now exceed all model 97.5 percentiles; similar results are obtained for the station-based index. The results in Figure 4c must be interpreted with care, however, because they are no longer addressing the simple question "are observed trends unusual in comparison to those simulated in model control runs?" By forcing model and observed series to have the same high-frequency variance, we are instead asking whether the spectral shape (e.g. the ratio of 30-year trends to intra-decadal variance) of the control run indices is sufficiently red to explain the occurrence of the observed trends (the observed trends have previously been shown to be significant in comparison with white noise and first-order autoregressive noise models by Stephenson et al. 2000; Thompson et al. 2000; Gillett et al. 2001). Figure 4c demonstrates that the control run spectra are not sufficiently red to explain the observed trends as internally-generated variability.

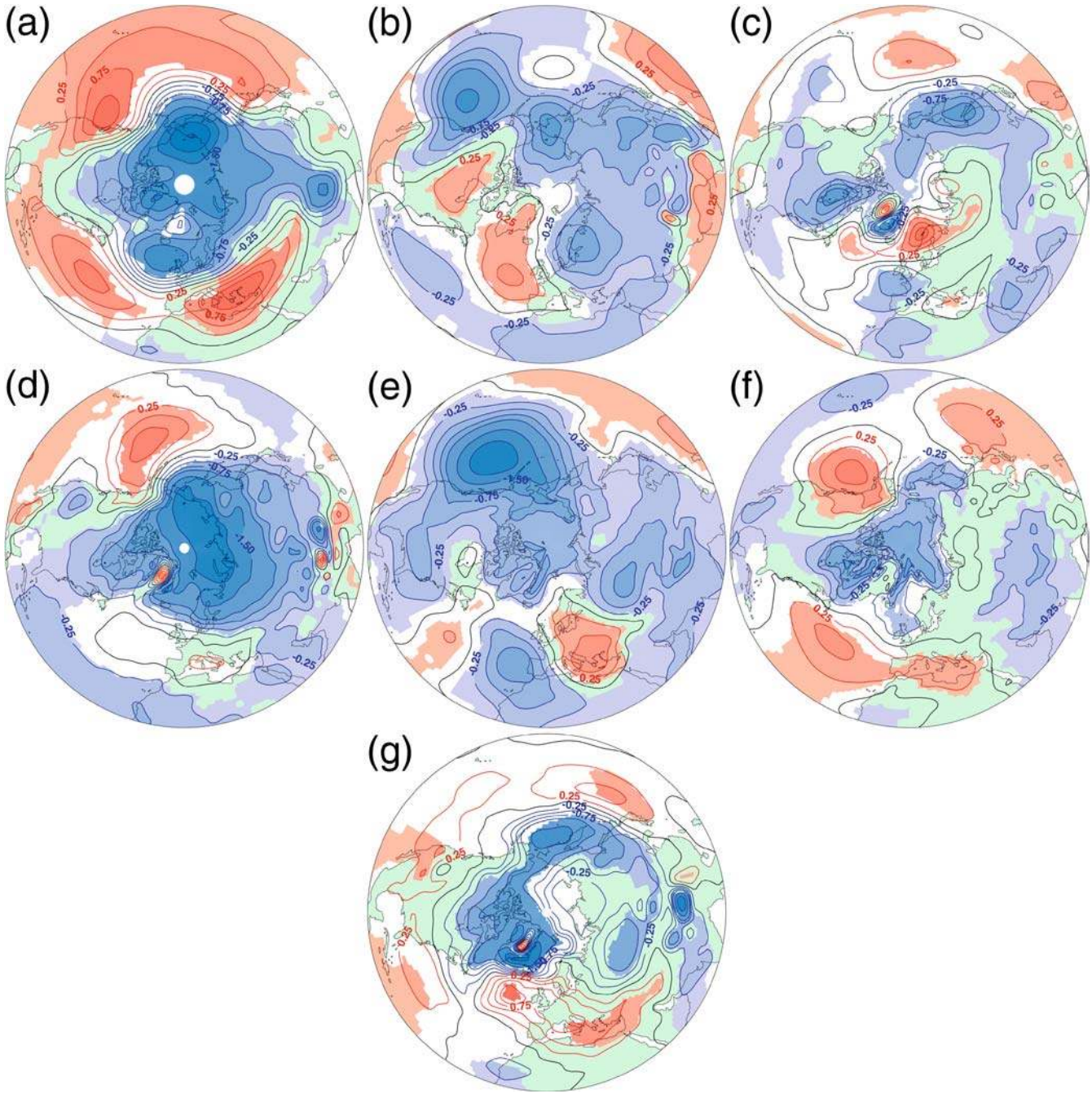


**Fig. 5** Trend in SLP under increasing greenhouse gas forcing, expressed as hPa per K of global-mean temperature increase, simulated by **a** CCSR/NIES, **b** CGCM1, **c** CSIRO Mk2, **d** ECHAM4, **e** HadCM2, **f** HadCM3, and **g** NCAR PCM. *Blue contours* are negative, zero is black, and positive are *red*; *blue* and *red shading* indicates regions with statistically-significant trends (95% confidence)

## 5 Simulated response of the North Atlantic Oscillation to greenhouse gas forcing

### 5.1 Trends in atmospheric circulation and the North Atlantic Oscillation

Under enhanced greenhouse gas forcing, all models simulate some changes in atmospheric circulation, represented in Fig. 5 by the trend in winter SLP fields over the Northern Hemisphere. The trends are expressed as changes per K of global and annual mean temperature increase. The patterns are almost identical to the *temporal* trend in SLP, except that the magnitudes are different because Fig. 5 removes the influence of different amounts of warming due to different model climate sensitivities. None of the simulated trends exhibit a pattern identical to the NAO pattern, so it would be a simplification to say that there is a shift in

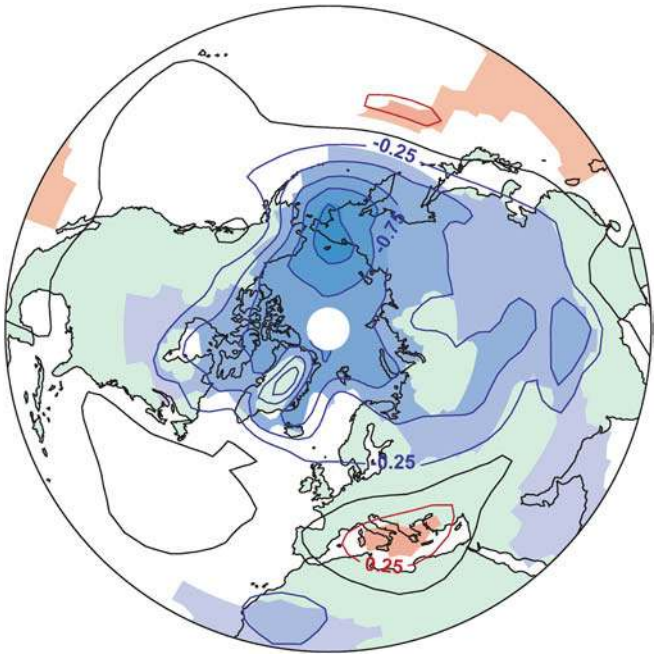


**Fig. 6** Contours show the mean of the seven simulated SLP trend patterns (hPa $K^{-1}$) shown in Fig. 5, induced by increasing greenhouse gas forcing. *Blue contours* are negative, zero is *black*, and positive are *red*; *blue* and *red shading* indicates regions where at least six of the seven models show a decrease or an increase, respectively

the phase of the NAO with increasing greenhouse gas concentrations. There is, however, a clear correspondence between the SLP trends and the NAO patterns (represented by the EOF of SLP) for CCSR/NIES (Figs. 5a and 3a) and ECHAM4 (Figs. 5d and 3d), and over the Atlantic half of the hemisphere for HadCM3 (Figs. 5f and 3f). The pattern correlations between trend pattern and EOF pattern over the Atlantic half of the hemisphere are about 0.75 for these three models. Even for the other models, with weaker detailed agreement between NAO pattern and climate change pattern, there are decreases in SLP over most of the Arctic Ocean.

There is clearly little agreement between the models concerning the detailed regional-scale changes in winter atmospheric circulation. This limits the certainty with which regional-scale climate change can be estimated, whether directly from climate models or via statistical downscaling models driven by these uncertain changes in regional circulation. At the largest scale, however, there is some semblance of inter-model agreement (Fig. 6), with the mean of all seven model patterns showing a decrease in SLP over the Arctic Ocean of 0.5 to 1.0 hPa $K^{-1}$ and an increase in SLP over the Mediterranean Sea of 0.25 hPa $K^{-1}$ . Both these changes are consistent across all seven model simulations. All show in increase over at least part of the Mediterranean Sea (Fig. 5), with six out of seven models agreeing on the sign of SLP change over the shaded regions of Fig. 6. At least six, and mostly all seven, models exhibit negative SLP trends over each Arctic Ocean grid box under increasing greenhouse-gas-induced warming.

Over the Atlantic half of the hemisphere, the seven-model-mean SLP trend pattern (Fig. 6) is highly positively correlated ($r = 0.86$) with the mean of the seven model EOF (i.e. NAO) patterns (Fig. 3h). If the EOF patterns are used to define the NAO index, therefore, it is not surprising that there is a general tendency for a shift towards the positive phase of the NAO (Fig. 7a). This is regardless of the scaling used (EOF scaling or pressure difference scaling). The strongest increases are in the ECHAM4 and CCSR/NIES simulations, and the weakest are CGCM1 and CSIRO Mk2. There is, therefore, inter-model agreement on the sign of the NAO change under greenhouse forcing, though little agreement on the magnitude of the shift towards the positive phase.

If the station-based NAO index is used instead, then there is not even agreement on the sign of the NAO response (Fig. 7b). Using this measure, the NAO index decreases significantly in HadCM2 (as found by Osborn et al. 1999) and slightly in CGCM1 and CSIRO Mk2. Overall, the station-based NAO index (Fig. 7b) gives a much wider scatter of results than does the pattern-based NAO index (Fig. 7a), and the pattern-based index is, therefore, preferred. The reasons for the wider scatter can be identified by inspection of the SLP trend patterns in Fig. 5. In HadCM2 (Fig. 5e), for example, SLP decreases slightly over Iceland (–0.25 hPa $K^{-1}$) but more strongly over Gibraltar (–0.75 hPa $K^{-1}$), leading to a reduction in the Gibraltar to Iceland pressure gradient and a reduction in that measure of the NAO index. Thus, regional-scale variations in SLP can have more influence over the selected stations than the "common" model signal (Fig. 6) that corresponds more closely with the NAO pattern.

The observed NAO indices are also overlaid on Fig. 7, and indicate absolute values that fall outside the group of model simulations for the pattern-based index (Fig. 7a) though the ECHAM4 and CCSR/NIES models do simulate values as high as those observed under the enhanced greenhouse gas forcing prescribed in the second half of the twenty first century. For the station-based index (Fig. 7b) the observations are less clearly separated from the group of model simulations. These comparisons of absolute NAO index values are dependent upon what the "zero" index value represents (i.e. what is the baseline from which these deviations are measured?). In all simulations, the SLP data that were used to compute the NAO indices were expressed as anomalies from their respective control run means. Thus, a simulated zero index value would represent normal control run conditions, which are effectively "pre-industrial" conditions because the increased greenhouse gas concentrations are applied as increases from the control run forcing (though, of course, other forcing factors which are neglected in this study complicate the issue of assigning the baseline to be truly

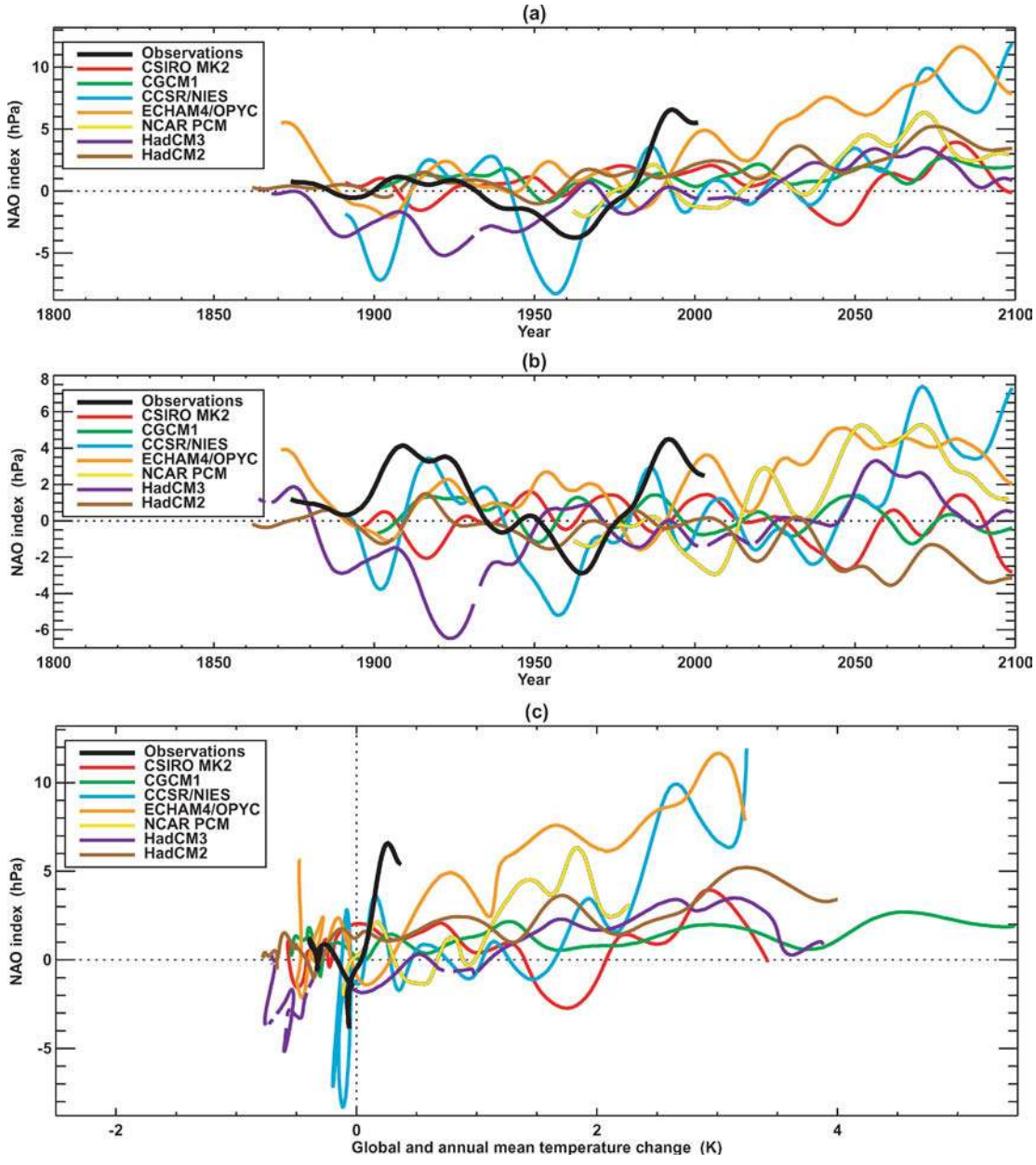


**Fig. 7a–c** Time series of NAO indices (hPa) as observed (*black*) and simulated under increasing greenhouse gas concentrations (CCSR/NIES: *blue*, CGCM1: *green*, CSIRO Mk2: *red*, ECHAM4: *orange*, HadCM2: *brown*, HadCM3: *purple*, NCAR PCM: *yellow*). **a** Pattern-based index under pressure difference scaling; **b** Gibraltar minus Iceland SLP difference; and **c** same as **a**, but plotted against the observed or simulated global and annual mean temperature anomaly (with respect to the 1961–1990 mean). All series have been smoothed with a 30-year low-pass filter. Gaps in the HadCM3 series are due to missing years in the data archives

"pre-industrial" climate). The observed SLP data are expressed as anomalies from their long-term mean, because we do not have "pre-industrial" observations. Given the relative weakness of the climate change signals shown in Fig. 7, the difference between the 1874–2001 or 1874–2003 means and the "pre-industrial" mean might be expected to be rather small. Nevertheless, it is this uncertainty in determining appropriate baseline values that led us to attempt detection of unusual *trends* (Sects. 2.4 and 4.2) rather than detection of unusual mean values.

Figure 7c addresses whether the different magnitudes of NAO response shown in Fig. 7a are due to different rates of climate change in the model simulations (although the forcing is very similar in each experiment, the temperature response, i.e. sensitivity, of these models ranges from 1.7 K to 4.3 K for a doubling of $CO_2$ concentration, Cubasch et al. 2001). Plotting the smoothed NAO index changes against smoothed global and annual mean temperature change ($\Delta T$) shows that the scatter is at least as wide (Fig. 7c) as when they are plotted against simulation year (Fig. 7a). The model

**Table 3** NAO index trends (per K of global-mean temperature change) from observations and from simulations under increasing greenhouse gas concentrations

| Model | Trend (hPa $K^{-1}$) | |
|---|---|---|
| | Pattern[a] | Station[b] |
| Observations | 5.0 | –0.1 |
| CCSR/NIES | 3.0 | 1.6 |
| CGCM1 | 0.2 | –0.1 |
| CSIRO Mk2 | 0.2 | –0.2 |
| ECHAM4 | 2.7 | 1.0 |
| HadCM2 | 0.8 | –0.6 |
| HadCM3 | 1.2 | 0.8 |
| NCAR PCM | 2.3 | 2.3 |

[a]Pattern-based NAO index under pressure difference scaling[b] Gibraltar minus Iceland SLP difference

that warms most (CGCM1) is not one of the models with the greatest NAO change and the model that warms least (NCAR PCM) is not one of the models with the smallest NAO change.

Linear trends fitted to the NAO–$\Delta T$ relationships are given in Table 3 (note that both records are smoothed before constructing Fig. 7c and fitting the trends, and note that the trends do not demonstrate a cause–effect relationship between NAO and global-mean temperature, but rather are expressing the NAO as a change per unit measure of climate response). All trends are positive when the pattern-based NAO index is used. Statistical significance is difficult to assess because of the smoothed series and uneven sampling of the temperature range, but providing there are at least ten effectively independent samples in each smoothed series, then all are significantly correlated with $\Delta T$ except for CSIRO Mk2, and this seems reasonable from inspection of Fig. 7c. The observed trend is likely to be only marginally significant because of the limited sampling of the temperature range, and the strength of the observed NAO–temperature relationship is not significantly different to any of the model-based trends. Repeating this exercise with the station-based NAO index (data not shown, but trends are given in Table 3), yields three negative and four positive relationships amongst the models, though neither the CGCM1 nor the CSIRO Mk2 trends are likely to be significant (again, assuming about ten degrees of freedom in the smoothed series). The negative trend for HadCM2 is statistically significant (Osborn et al. 1999).

### 5.2 Changes to the pattern and variability of the North Atlantic Oscillation

In addition to the changes in the mean winter atmospheric circulation under enhanced greenhouse gas forcing, it is possible that the nature of the inter-annual variability of winter circulation may change. In this section, the 2050–2099 periods of the simulations under increasing greenhouse gas forcing are detrended (to focus on inter-annual variability rather than climate trends) and analysed to determine whether systematic changes in the spatial pattern or temporal variance of the leading mode of Atlantic-sector variability occur. The leading EOFs of the detrended SLP data, subsequently extended to show the covarying pattern over the full hemisphere, are shown in Fig. 8 and can be compared with those obtained from the control simulations (Fig. 3).

For CCSR/NIES (Figs. 3a, 8a) there is an eastward shift and strengthening of the Azores centre of action and a northward shift of the zero line. For CGCM1 (Figs. 3b, 8b) there is a complete disappearance of the Azores centre of action (which was already much weaker than observed) and a dramatic eastward shift (and slight strengthening) of the Iceland centre of action to northern Siberia. CSIRO Mk2 (Figs. 3c, 8c) shows smaller changes, a north-eastward movement of the Azores centre of action, a northward displacement of the zero line, and a slight strengthening of the Iceland centre of action (over Greenland). The changes in ECHAM4 (Figs. 3d and 8d) have previously been noted by Ulbrich and Christoph (1999), i.e. an eastward shift of both centres of action associated with an eastward extension of the Atlantic storm track into western Europe. HadCM2 (Figs. 3e and 8e) and HadCM3 (Figs. 3f and 8f) both simulate a moderate eastward shift in the Azores centre of action, and a clear reduction in covariance with North Pacific SLP (bringing them closer to the observed EOF pattern, Fig. 3i), and HadCM3 also shows a slight northward shift of the zero line. NCAR PCM (Figs. 3g, 8g) simulates a strengthening of the Azores centre of action and a weakening of the Iceland centre of action.

Assessing the statistical significance of changes in EOF patterns between one data set and another is not a straightforward task. A qualitative assessment has been undertaken, however, by visually inspecting the leading EOF patterns computed from many different (though partly overlapping) 50-year segments of the control simulations. Of the changes between the control (Fig. 3) and GHG (Fig. 8) runs described above, very similar differences occurred during some sub-periods of the control runs for CCSR/NIES, HadCM2 and HadCM3, but not for CGCM1, CSIRO Mk2, ECHAM4 or NCAR PCM.

Considering the seven simulations as a multi-model ensemble, the mean of the control run EOF patterns (Fig. 9a, repeated from Fig. 3h with half the contour interval) can be compared with the mean of the increasing GHG simulation EOF patterns (Fig. 9b), with differences being shown in Fig. 9c. The Azores centre of action shifts eastward and northward (the zero line in the vicinity of the UK moves northwards). While the position of the Iceland centre of action minimum does not change very much, the "trough" over Russia does extend and deepen due to the eastward extension of this centre of action in some models.

An alternative view of these changes is presented in Fig. 10. Given that it is the NAO's relationships with

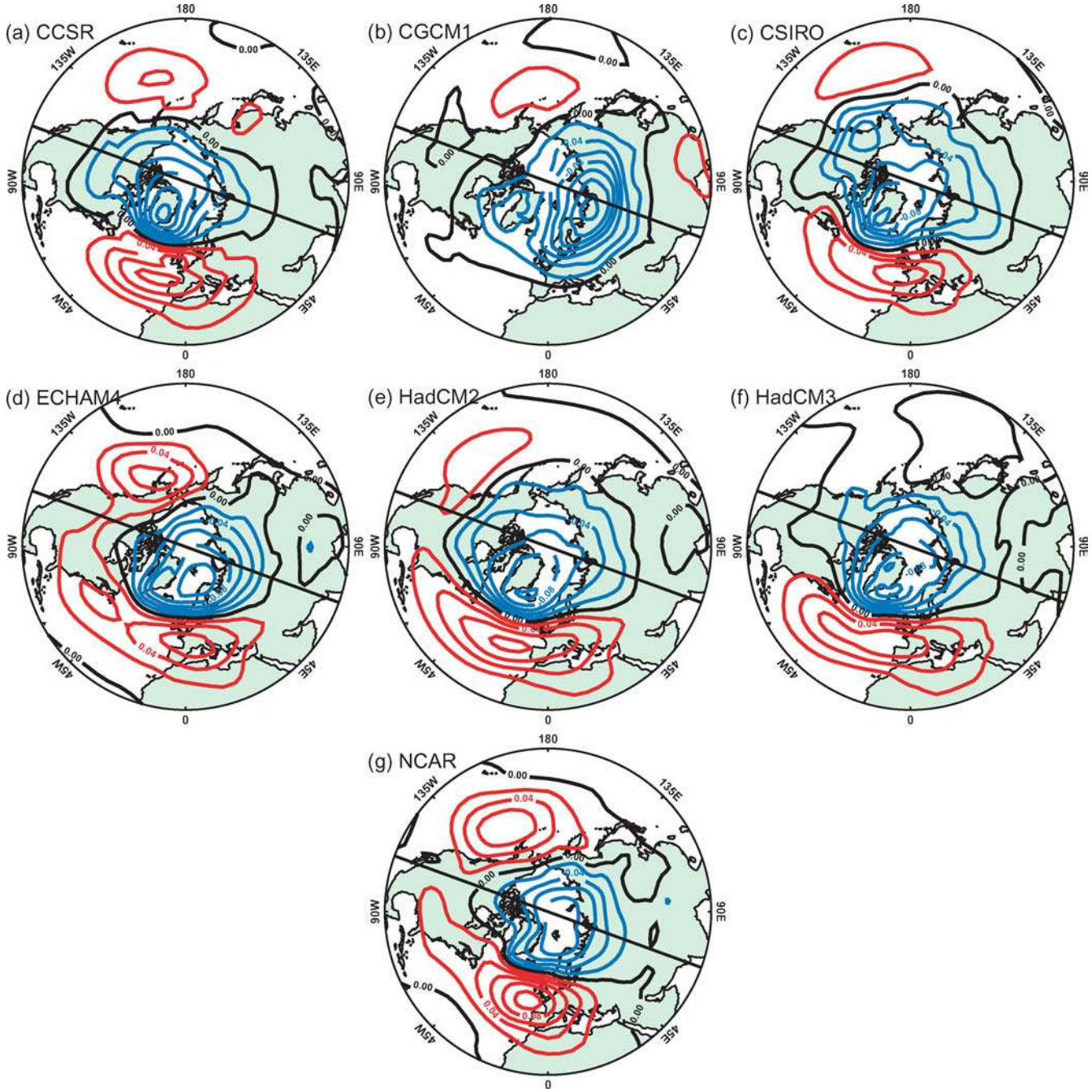


**Fig. 8a–g** As Fig. 3a–g, but using linearly detrended SLP from 2050–2099 of the increasing GHG simulations

the storm track and the mean westerly circulation over the Atlantic/European sector that give the NAO its climatic importance, the change in the meridional pressure gradient might be more important than the change in the minima and maxima of the EOF pattern. The two are, of course, related, but are not necessarily identical. The meridional pressure gradient at every point in each EOF pattern has been computed, then divided by the sine of the latitude (because the geostrophic wind also depends on the inverse of the Coriolis parameter, which varies with the sine of the latitude) to obtain the zonal geostrophic wind anomalies associated with each pattern. These are then averaged over all seven control runs (Fig. 10a) and all seven increasing GHG runs (Fig. 10b), and the composites are differenced (Fig. 10c). The zero line of these patterns follows the peak of the Azores high, with negative values to the south where the NAO influences the strength of the north-easterly trade winds and positive values northwards as far as the peak of the Iceland Low centre of action. The pattern peaks (Fig. 10a) south of Iceland, where the influence of the NAO pattern on zonal wind

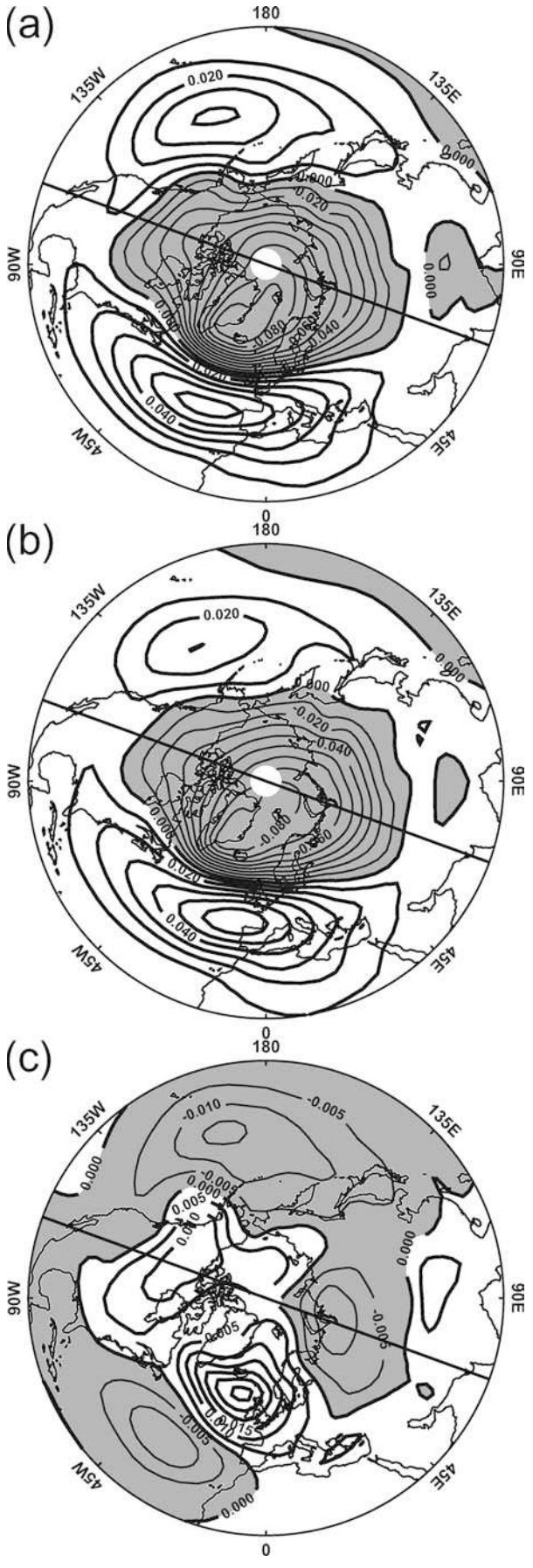


**Fig. 9** SLP patterns, averaged across all seven model simulations, associated with the leading PC of detrended Atlantic sector SLP variability for **a** control, **b** increasing GHG, and **c** the difference **b**–**a**. Negative values are *shaded*; contour interval is 0.01 in **a** and **b**, 0.005 in **c**. The *black diameter line* marks the Atlantic half of the hemisphere from which the PC was defined

will be greatest. This alternative view of the NAO pattern may be useful in other applications too.

Under increased GHG concentrations, the change in the pattern (Fig. 10c) indicates a northward rather than an eastward shift. Positive changes over the southern part of the trade wind region indicate a northward retreat. The negative change covers the northern part of the trade wind region (extending it northwards) and the southern part of the westerlies (retreating northward). The positive change over Iceland indicates a northward extension of the NAO influence on the westerlies. The alternative possibility of an *eastward* shift in the dominant mode of inter-annual winter variability is discussed further in Sect. 6.

The percent of Atlantic sector SLP variance captured by the leading EOF indicates the importance of the NAO mode of variability relative to other modes during winter. The values for the control runs are given in Table 2. For the detrended 2050–2099 period of the increasing GHG simulations, five of the models simulate similar values to their control simulations, but two models simulate a much reduced relative importance of the NAO: the SLP variance captured by the leading EOF decreases from 56% to 42% in CGCM1 and from 55% to 47% in ECHAM4. The former is clearly related to the movement of the leading EOF towards the edge of the Atlantic sector (Fig. 8b) in CGCM1. Comparison with the leading EOF from many different (though partly overlapping) 50-year segments of the model control simulations indicates that the CGCM1 and ECHAM4 changes are significant (i.e. the null hypothesis that the lower EOF variance is due to sampling variability is rejected).

These results give the *relative* importance of the NAO only, but the actual magnitude of the inter-annual variability is better measured by the standard deviation of the NAO indices from each model (Table 2). The standard deviation of the Gibraltar minus Iceland SLP difference is lower during 2050–2099 than during the control simulations for six out of the seven models (the exception is HadCM2), though the difference is only statistically significant with 95% confidence for CGCM1 and ECHAM4 (the latter agrees with the results of Paeth et al. 1999). That the reductions are greatest for these two models is not surprising, given the eastward movement of the northern centre of action away from Iceland, and the reduction in relative importance of the leading EOF.

The NAO indices obtained using the control run EOF patterns of each model (under pressure difference scaling), applied to the 2050–2099 period of the increasing GHG simulations, also exhibit lower variance

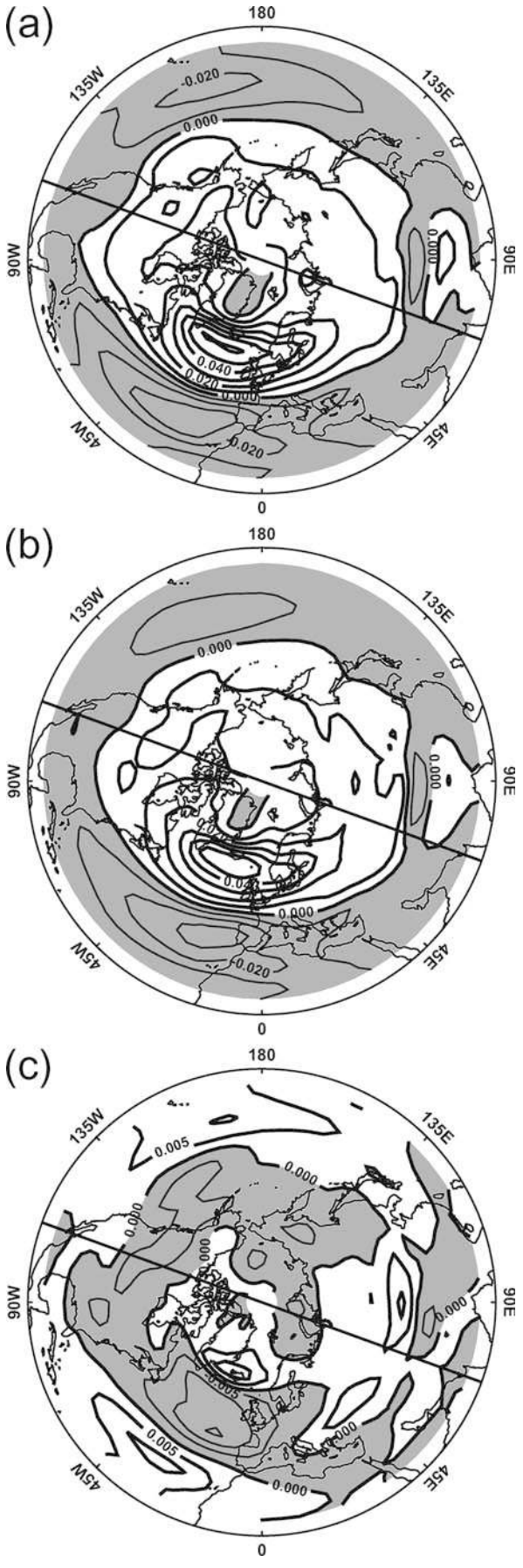


**Fig. 10a–c** As Fig. 9, except that the zonal geostrophic wind anomalies (arbitrary units) are calculated from the meridional SLP gradient divided by the Coriolis parameter, prior to averaging across the seven model simulations

(significant only for CGCM1) compared with the control simulations for all models except HadCM2. That is to be expected for purely statistical reasons, however, because the EOFs were defined to capture maximum variance during the control simulations and may not be optimal for the GHG simulations. The values given in Table 2 show instead the standard deviations when the NAO indices are obtained using the GHG-defined EOFs (Fig. 8), using pressure difference scaling. Comparing control standard deviations obtained with control EOFs against GHG standard deviations obtained with GHG EOFs, gives a more complex result: two models (CGCM1 and ECHAM4, again) show reductions in variance and the other five all show increases. None of the changes are statistically significant. Thus, though the simulations suggest that the Gibraltar to Iceland SLP difference will become less variable in future, the variance associated with the leading EOF of inter-annual SLP variations may behave differently.

## 6 Discussion and conclusions

This study has obtained several important findings and has addressed the two issues raised in the introduction. The simulations of the Northern Hemisphere mean winter circulation and its inter-annual variability are considered good enough to use the seven climate models considered here (Table 1) for addressing these NAO and climate change issues. The patterns of Atlantic-sector variability are very similar to the observations except that the connection with the North Pacific SLP is too strong in some models (Fig. 9a); in some cases (e.g. ECHAM4) the too strong connection with the Pacific may be explained by errors in the stratospheric simulation (Castanheira and Graf 2003). It is interesting that the simulated variability is closer in pattern to the Arctic Oscillation than the NAO, even though the study was deliberately biased towards the NAO by using only the Atlantic half of the Northern Hemisphere to define the inter-annual variability. Despite the simulated links with the North Pacific, the pressure gradients that drive zonal wind anomalies are much stronger in the Atlantic sector than the Pacific (Fig. 10a), so even in the models there is very much an Atlantic-focus to the variability considered here.

The inter-annual NAO variance is too great in some models (Table 2), a bias that probably influences inter-decadal variability too. Nevertheless, the variability exhibited by the seven climate model control runs analysed here is unlikely to be strong enough to explain the recent NAO trend (observed from the 1960s to the 1990s) as a simple internally generated climate signal

(Fig. 4). This result is in agreement with, but extends, the findings of Osborn et al. (1999) and Gillett et al. (2000). Though there is some dependence on which definition of the NAO index is used, the implication is that either the models are deficient in their simulation of inter-decadal NAO variability or that the observed variations have been driven, in part at least, by external forcings.

Only greenhouse gas forcing was considered here, using simulations from seven models forced by a compounded 1% per year increase in effective carbon dioxide. The response is much more model-dependent when a station-based NAO index is used, compared with a pattern-based (EOF) index. For the latter, all models simulate a shift towards a higher winter NAO index (Fig. 7a), though its magnitude varies from 0.2 to 3.0 hPa per K of global-mean temperature rise (the units represent increasing pressure difference between the locations with maximum positive and negative weightings in the EOF pattern). Note that here the changes are expressed per unit surface warming, whereas others (e.g. Gillett et al. 2002) express them per unit radiative forcing. The latter is more suitable if a direct atmospheric mechanism drives the NAO changes (perhaps via the stratosphere, Shindell et al. 1999), while the former is more appropriate if the surface boundary conditions (e.g. temperature) drive the NAO changes. Which form of expression is preferable remains debatable, because the driving mechanism for NAO changes is still uncertain (Gillett et al. 2003).

The analysis of NAO response to greenhouse gas forcing presented here represents a considerable advance, because inter-model comparisons are greatly enhanced by the application of the same methods to a number of model simulations. They are generally in agreement with previously published work: Ulbrich and Christoph (1999) identified an NAO increase in ECHAM4 with increasing GHGs; Fyfe et al. (1999) report an Arctic Oscillation increase in CGCM1 with increasing GHGs and sulfate aerosols; and Gillett et al. (2002) report an Arctic Oscillation increase in HadCM3 with increasing GHGs. Results obtained using models that are not considered here are summarised by Gillett et al. (2003). The HadCM2 response is very dependent upon the NAO index used: Osborn et al. (1999) report a decrease when using the Gibraltar to Iceland SLP difference, whereas here the Atlantic part of the EOF pattern results in an increase, while the full hemisphere EOF pattern (i.e. the Arctic Oscillation) results in a decrease (Zorita and Gonzalez-Rouco 2000) due to the deepening of the Aleutian Low shown in Fig. 5e. This sensitivity is due to the fact that the change in mean circulation does not closely resemble the NAO pattern for HadCM2 (nor for some of the other models). Only for ECHAM4 and CCSR/NIES do the patterns of climate change and the patterns of inter-annual variability match closely. The only signal of winter circulation change that is robust across all models is a reduction in SLP over the Arctic, which is the cause of increased NAO index in all these models. Increased SLP occurs over the Mediterranean Sea in six of the models, contributing to the strengthened meridional pressure gradient.

The model simulations imply, therefore, that greenhouse gas forcing has contributed to the observed NAO increase from the 1960s to the 1990s. But the magnitude of the simulated signal (Table 3) is much less than the observed change, which peaked at 16.2 hPa over a 30-year period (using the pattern-based index under pressure-difference scaling) during which global temperature increased by less than 0.5 K. Combining the strongest of the simulated signals (3.0 hPa $K^{-1} \times 0.5$ K = 1.5 hPa) with the strongest of the internally-generated 97.5 percentile trends (5.2 hPa decade$^{-1} \times 3$ decade = 15.6 hPa) gives 17.1 hPa over a 30-year period, which exceeds the largest observed trend. It is possible, therefore, that the observed record can be explained as a combination of internally generated variability and a small greenhouse-gas-induced positive trend. This is supported by the downturn in the NAO index after the mid-1990s (Fig. 1), which might be a reversal of an internally generated variation.

It was pointed out in Sect. 4.2, however, that the models that simulate the strongest internally-generated trends have significantly too much inter-annual variability compared with the observed record. If the model simulations are scaled so that their high-frequency variability matches the observed (e.g. similar to Fig. 4c, but expressed in pressure-difference units Sect. (2.3), then the strongest 97.5 percentile trends are only 9.4 hPa over a 30-year period, which gives less than 11 hPa when added to the GHG signal. The distinct possibility remains, therefore, that other external climate forcings have contributed to the observed change (though errors in the simulation of internal inter-decadal variability or of the response to GHG forcing are possible alternative explanations). Other forcings are not studied here, but they were discussed by Gillett et al. (2003), who concluded that the response of the NAO to these forcings was uncertain. It is likely that stronger volcanic forcing would lead to a higher winter index; it is possible that enhanced solar irradiation leads to a higher index and possible that reduced stratospheric ozone results in a higher index (though more so in spring than in winter). All three of these forcings may have changed, over the past 40 years, in the direction likely to increase the NAO index. Simulations using HadCM3 under these combined forcings (Stott et al. 2000) have not, however, reproduced the observed NAO trend.

The final analysis section of this study addressed the possibility that the spatial pattern and/or temporal variance of inter-annual winter circulation variability might change under increased greenhouse gas forcing. Ulbrich and Christoph (1999) have previously identified a northeastward shift in the dominant pattern of variability in the ECHAM4 simulation, associated with the extension of the Atlantic storm track into Western Europe. Similar changes have been noted in the observations (e.g. Hilmer and Jung 2000). Collectively, the

seven models analysed here show only a small north-eastward shift in the centres of action of the dominant mode of inter-annual variability, and only a northward shift in their influence on zonal winds. Peterson et al. (2003) suggest that an eastward shift in the pattern would occur during a period of high NAO index, so there might be a link between those models showing an increase in the mean NAO index and those showing an eastward shift in the pattern. Of the two models that simulate the strongest increase in NAO index under increasing GHG forcing, CCSR/NIES simulates an eastward shift of the southern (Azores) centre of action only, while ECHAM4 simulates an eastward shift of both centres of action.

**Acknowledgements** This study was supported by the Commission of the European Communities (SO&P, EVK2-CT-2002-00160), with additional support from the Tyndall Centre for Climate Change Research (IT1.15). All modelling centres are gratefully acknowledged for making their data available, and David Viner is thanked for his assistance in obtaining the data. Updated Met Office SLP data were provided by British Atmospheric Data Centre (BADC). Simon Tett, Keith Briffa, Phil Jones and David Stephenson provided useful guidance at various stages of this work. The constructive comments provided by two reviewers helped to improve this paper.

## References

Bacher A, Oberhuber JM, Roeckner E (1998) ENSO dynamics and seasonal cycle in the tropical Pacific as simulated by the ECHAM4/OPYC3 coupled general circulation model. Clim Dyn 14: 1659–1672

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

Castanheira JM, Graf H-F (2003) North Pacific–North Atlantic relationships under stratospheric control. J Geophys Res 108: 4036 DOI 10.1029/2002JD002754

Cubasch U, Meehl GA, Boer GJ, Stouffer RJ, Dix M, Noda A, Senior CA, Raper S, Yap KS (2001) Projections of future climate change. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) Climate change 2001: the scientific basis. Contribution of Working Group I to the third assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, USA, pp 525–582

Emori S, Nozawa T, Abe-Ouchi A, Numaguti A, Kimoto M, Nakajima T (1999) Coupled ocean-atmosphere model experiments of future climate change with an explicit representation of sulfate aerosol scattering. J Meteorol Soc Japan 77: 1299–1307

Flato GM, Boer GJ, Lee WG, McFarlane NA, Ramsden D, Reader MC, Weaver AJ (2000) The Canadian Centre for Climate Modelling and Analysis global coupled model and its climate. Clim Dyn 16: 451–467

Fyfe JC, Boer GJ, Flato GM (1999) The Arctic and Antarctic Oscillations and their projected changes under global warming. Geophys Res Lett 26: 1601–1604

Gillett NP, Hegerl GC, Allen MR, Stott PA (2000) Implications of changes in the Northern Hemisphere circulation for the detection of anthropogenic climate change. Geophys Res Lett 27: 993–996

Gillett NP, Baldwin MP, Allen MR (2001) Evidence for nonlinearity in observed stratospheric circulation changes. J Geophys Res 106: 7891–7901

Gillett NP, Allen MR, McDonald RE, Senior CA, Shindell DT, Schmidt GA (2002) How linear is the Arctic Oscillation response to greenhouse gases. J Geophys Res 107: DOI 10.1029/2001JD000589

Gillett NP, Graf HF, Osborn TJ (2003) Climate change and the North Atlantic Oscillation. In: Hurrell JW, Kushnir Y, Ottersen G, Visbeck M (eds) North Atlantic Oscillation: climatic significance and environmental impact. (Geophysical Monograph 134) American Geophysical Union, Washington, pp 193–209 DOI 10.1029/134GM09

Gordon HB, O'Farrell SP (1997) Transient climate change in the CSIRO coupled model with dynamic sea ice. Mon Weather Rev 125: 875–907

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjusments. Clim Dyn 16: 147–168

Hilmer M, Jung T (2000) Evidence for a recent change in the link between the North Atlantic Oscillation and Arctic sea ice export. Geophys Res Lett 27: 989–992

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: 655–668

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, Jonsson T, Wheeler D (1997) Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland. Int J Climatol 17: 1433–1450

Lambert SJ, Boer GJ (2001) CMIP1 evaluation and intercomparison of coupled climate models. Clim Dyn 17: 83–106

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

Osborn TJ, Briffa KR, Tett SFB, Jones PD, Trigo RM (1999) Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model. Clim Dyn 15: 685–702

Ostermeier GM, Wallace JM (2003) Trends in the North Atlantic Oscillation–Northern Hemisphere annular mode during the twentieth century. J Clim 16: 336–341

Paeth H, Hense A, Glowienka-Hense R, Voss R, Cubasch U (1999) The North Atlantic Oscillation as an indicator for greenhouse-gas induced regional climate change. Clim Dyn 15: 953–960

Peterson KA, Lu J, Greatbatch RJ (2003) Evidence of nonlinear dynamics in the eastward shift of the NAO. Geophys Res Lett 30: 1030 DOI 10.1029/2002GL015585

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

Shindell DT, Miller RL, Schmidt GA, Pandolfo L (1999) Simulation of recent northern winter climate trends by greenhouse-gas forcing. Nature 399: 452–455

Stephenson DB, Pavan V (2003) The North Atlantic Oscillation in coupled climate models: a CMIP1 evaluation. Clim Dyn 20: 381–399

Stephenson DB, Pavan V, Bojariu R (2000) Is the North Atlantic Oscillation a random walk? Int J Climatol 20: 1–18

Stott PA, Tett SFB, Jones GS, Allen MR, Mitchell JFB, Jenkins GJ (2000) External control of 20th century temperature by natural and anthropogenic forcings. Science 290: 2133–2137

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

Thompson DWJ, Wallace JM, Hegerl GC (2000) Annular modes in the extratropical circulation. Part II: trends. J Clim 13: 1018–1036

Ulbrich U, Christoph M (1999) A shift of the NAO and increasing storm track activity over Europe due to anthropogenic greenhouse gas forcing. Clim Dyn 15: 551–559

van Loon H, Rogers J (1978) The seesaw in winter temperature 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

Wanner H, Bronnimann S, Casty C, Gyalistras D, Luterbacher J, Schmutz C, Stephenson DB, Xoplaki E (2001) North Atlantic Oscillation – concepts and studies. Surv Geophys 22: 321–382

Washington WM, Weatherly JW, Meehl GA, Semtner Jr AJ, Bettge TW, Craig AP, Strand Jr WG, Arblaster JM, Wayland VB, James R, Zhang Y (2000) Parallel climate model (PCM) control and transient simulations. Clim Dyn 16 755–774

Zorita E, Gonzalez-Rouco F (2000) Disagreement between predictions of the future behaviour of the Arctic Oscillation as simulated in two different climate models: implications for global warming. Geophys Res Lett 27: 1755–1758