A MODEL INTERCOMPARISON STUDY OF CLIMATE CHANGE-SIGNALS IN EXTRATROPICAL CIRCULATION

MONIKA RAUTHE, ANDREAS HENSE, and HEIKO PAETH

a Leibniz Institut für Atmosphärenphysik, Kühlungsborn, Germany
b Meteorologisches Institut, Universität Bonn, 53121 Bonn, Germany

Received 2 July 2003
Revised 6 January 2004
Accepted 17 January 2004

ABSTRACT

Since 1970, the observed time series of various extratropical circulation modes have revealed remarkable trends. In many studies it has been suggested that these trends may be related to global warming due to increasing greenhouse gas (GHG) concentrations. Coupled climate model scenario experiments may give a hint of such a relationship. Here, a large model intercomparison study is presented, incorporating most state-of-the-art models of the international modeller community with GHG and GHG plus sulphate aerosol (SUL) forcing, in order to quantify the signals common to different climate models and to determine the degree of uncertainty. The extratropical circulation candidates are the Arctic oscillation (AO), the North Atlantic oscillation (NAO), the Aleutian low (AL) and the Antarctic oscillation (AAO).

Most climate models agree in predicting positive AO and AAO trends into the 21st century, these being different from the respective results of long-term control experiments. The NAO appears to be less sensitive to radiative forcing, with slightly positive and negative trends occurring in different models. The AL tends to strengthen in several models with GHG + SUL forcing. Projecting the spatial structure of the circulation modes onto the trend patterns of mean sea-level pressure (SLP) indicates that, in particular, the AO and AAO contribute considerably to the simulated long-term trends in SLP. Intermodel variations in Northern Hemisphere SLP trends become predominantly apparent over the mountainous regions and the North Pacific. In the Southern Hemisphere, the Antarctic region is subject to large model uncertainties. The multi-decadal trends of all circulation modes except the NAO are statistically significant in the majority of the climate-change experiments. At the interannual time scale, external radiative forcing does account for a small but statistically significant part of total multi-model variability, but this part is of the same order of magnitude as the systematic differences between the models. At decadal time scales, the external impact clearly stands out from the other sources of variability. Present-day climate models mostly agree in predicting a strengthening of the annular modes AO and AAO. As all models generally project a deepening of SLP over the polar caps, time series indicative of these regions might be a more appropriate measure of the sensitivity of extratropical circulation. Copyright © 2004 Royal Meteorological Society.

KEY WORDS: Arctic oscillation; North Atlantic oscillation; Aleutian low; Antarctic oscillation; climate change; model intercomparison

1. INTRODUCTION

The extratropical circulation modes describe important aspects of coherent climate variability. It is essential to know the past, present and future conditions of their spatial structure and temporal behaviour. In a series of papers, various extratropical circulation modes have been defined like the North Atlantic oscillation (NAO; e.g. Hurrell, 1995), the Arctic oscillation (AO; Thompson and Wallace, 1998), the Antarctic oscillation (AAO; e.g. Fyfe et al., 1999), and the Aleutian low (AL) as one centre of action of the so-called Pacific North America (PNA) pattern (e.g. Wallace and Gutzler, 1981; Hurrell and van Loon, 1997). In this study, we consider the AO and its Southern Hemisphere (SH) counterpart, the AAO, as almost zonally symmetric patterns, as well as the more regional NAO and the AL. The latter has been chosen instead of the PNA pattern because it basically counteracts the Pacific centre of the AO.

* Correspondence to: Heiko Paeth, Meteorologisches Institut, Universität Bonn, 53121 Bonn, Germany; e-mail: hpaeth@uni-bonn.de

Copyright © 2004 Royal Meteorological Society
The AO (AAO) is defined as the leading mode of sea-level pressure (SLP) variability north (south) of 20°N (20°S) (Thompson and Wallace, 1998). The AO pattern describes a tripole with two positive centres over the mid-latitude oceans and an opposing one over the Arctic. A strong positive trend from the 1960s onward, associated with a deepening of pressure over the northern latitudes, is observed (Thompson et al., 2000). If there is a trend in the AO index, than the meridional pressure gradient changes, modifying the orientation of the storm track over the North Atlantic and North Pacific. The result is a modification of moisture and heat transport into the adjacent continental areas (Shindell et al., 1999).

A similar situation is found over the North Atlantic, where the NAO is resident with the Icelandic low (IL) and the Azores high (AH) as anti-correlated centres of action (Hurrell, 1995; Kapula et al., 1998). In this sector, a reliable database exists with long-term observations of SLP. The AO and NAO are quite similar to each other, and some people argue that it may be difficult to distinguish between them (Ambaum et al., 2001). Here, we will demonstrate that, with respect to climate-change sensitivity, the modes are easily distinguishable. The NAO index also reveals a pronounced positive trend during the last decades of the 20th century (Hurrell and van Loon, 1997) which is also simulated under greenhouse gas (GHG) forcing by some selected climate model experiments (Paeth et al., 1999; Ulbrich and Christoph, 1999).

The AL is the North Pacific counterpart of the IL. In addition, it interacts with the Pacific centre of the AO: if the AL strengthens, then the AO index consequently decreases, since a positive AO index is an indicator of positive SLP anomalies over the North Pacific (Wallace and Thompson, 2002). There is some evidence in the observational data that the wintertime AL has intensified since the middle of the 20th century (Graham and Diaz, 2001).

The manifestation of the AAO is even more zonally symmetric than the AO, due to the SH land–sea distribution, and the time series show a remarkable intensification (Fyfe et al., 1999; Kushner et al., 2001; Thompson and Solomon, 2002). The AAO and AO are closely tied to stratospheric dynamics with the polar vortex, which in turn is relevant to ozone depletion (Thompson and Solomon, 2002). This part of the world is characterized by extremely sparse data coverage (Hines et al., 2000). Climate models also have substantial problems in simulating Antarctic climate (Paeth and Hense, 2002).

The remarkable trends in real-climate circulation modes have motivated many studies about the sensitivity of atmospheric circulation to global warming (Corti et al., 1999; Paeth et al., 1999; Gillett et al., 2000; Graham and Diaz, 2001; Houghton et al., 2001; Kwok and Comiso, 2002). It is obvious to consult climate model simulations with scenarios of increasing GHG concentrations, partly involving the effect of sulphate aerosols (SULs), in order to gain insight into the future development of extratropical circulation. So far, such model studies, mainly based on AO and NAO in individual climate model experiments, do not draw a homogeneous picture of the respective responses: the German and Canadian coupled climate models are clearly sensitive to the radiative forcing (Fyfe et al., 1999; Paeth et al., 1999; Ulbrich and Christoph, 1999); but the HADCM2 hardly shows any reaction (Gillett et al., 2000). In some models, the inclusion of stratospheric dynamics is very important to the model response (Shindell et al., 1999). A recent study by Paeth and Hense (2002) has demonstrated that the response of temperature, and particularly rainfall, to the GHG forcing may differ from model to model in many regions over the globe. As it is largely unknown which model’s prediction is most reliable, an objective study of climate-change signals has to consider ensemble simulations of all available climate models of the international modeller community in order to evaluate the signals in terms of internal variability and model uncertainty.

In contrast to previous studies, we enlarge substantially the number of climate models considered and compare the sensitivity of several circulation modes to global warming with each other. The AO, NAO, AL and NAO are analysed in nine state-of-the-art general circulation models (GCMs) with different forcings, i.e. GHG and GHG + SUL. The indices of the four circulation modes are determined based on the monthly mean SLP field. Furthermore, we refer to control runs from four GCMs to give an estimate of natural variability. The climate model predictions are compared with an observed time series for the whole 20th century that represents a merged product of a historical SLP data set (Trenberth and Paolino, 1980; Stein and Hense, 1994) and the National Centers for Environmental Prediction (NCEP) reanalysis data set (Kalnay et al., 1996). This results in centennial-scale descriptions of the observed AO and NAO. Most model data enter as multi-model ensembles with identical forcing but with varied initial conditions, which holds the prospect of separating
external, internal and intermodel contributions to total variability (Krishnamurti et al., 1999; Li, 1999; Kharin and Zwiers, 2002). Moreover, as there is no prior knowledge of the relative development of GHG and SUL into the 21st century (Houghton et al., 2001), the consideration of both scenarios ensures that we follow an objective way. The classical statistical tools to decompose total variability into signal, noise and systematic model differences are one-way (1w) and two-way (2w) analysis of variance (ANOVA).

Section 2 provides some information about the data sets considered and the statistical methods. Section 3 describes the results in detail, which are then summarized and discussed in Section 4.

2. DATA AND METHODS

2.1. Data sets

We use two observational data sets. The first one is a combination of the historical SLP data set (1881–1945; Trenberth and Paolino, 1980) with analyses of the German Weather Service (DWD) during the period 1946–94, describing the long-term behaviour of the NH extratropical circulation modes north of 20°N. The data set is largely homogeneous (Stein and Hense, 1994; Born, 1996) at a monthly time scale. Additionally, we rely on the NCEP–National Center for Atmospheric Research (NCAR) reanalysis product (Kalnay et al., 1996, updated) for the period 1948 to 2001. It has been found that both data sets match each other quite well during the overlapping period, even with respect to individual monthly means. Thus, the indices calculated for both data sets are merged into a long-term time series of the observed NH circulation modes. In the AL time series, gaps occur over periods, where a certain amount of grid-point information is missing in the North Pacific sector. When calculating the AO index, it has also been taken into account that the grid-point information is varying in time.

Nine state-of-the-art coupled climate models, some of them additionally in different versions, are studied. The natural variability of the circulation modes is derived from six control simulations from four different models where the GHG concentrations are kept constant at a late-20th century level (usually 1985). In the forced runs, observed GHG and SUL concentrations are prescribed for the past, whereas the future anthropogenic GHG and SUL emissions are projected with the help of emissions scenarios, corresponding to Intergovernmental Panel on Climate Change (IPCC) ‘business as usual’ (Houghton et al., 2001). The forcings (SA90 versus IS92A) are not exactly the same, but are largely comparable: SA90 amounts to slightly more than 1% per year and is only used in the ECHAM3 experiments. All other climate-change simulations are based on the more recent IS92A scenario, which describes an increasing rate of slightly less than 1% per year into the future. Apart from the GHG and GHG + SUL experiments, there are also three runs that include additional ozone changes. Some details about the exact model version, the integration period and the individual forcings of all climate model simulations considered can be inferred from Table I. All models have been subject to various climatological studies in the past, showing that they provide reasonable representations of extratropical climate. Further information about the model designs and validation can be found for ECHAM3 in Roeckner et al. (1992), for ECHAM4 in Roeckner et al. (1996), for CGCM1 in Flato et al. (1998), for CGCM2 in Boer et al. (2000), for HADCM2 in Johns et al. (1997), for HADCM3 in Pope et al. (2000), for GISS in Shindell et al. (1999), for GFDL in Delworth et al. (2002), for CSIRO in Gordon and O’Farrell (1997), for CCSR in Emori et al. (1999), for ARPEGE in Royer et al. (2002) and for PCM in Washington et al. (2000). It has to be noted that the HADCM2 model is not mass consistent over the poles. This deficiency is eliminated as suggested by Osborn et al. (1999), by removing the hemispheric spatial-mean of SLP at each grid point and time. The ARPEGE/OPA experiments include chemical processes in the stratosphere, once with homogeneous and once with heterogeneous ozone chemistry (Royer et al., 2002). All models except HADCM2 and HADCM3 are spectral models. Given the different horizontal resolutions, all data sets are interpolated to regular a Gaussian \( \lambda \varphi \) grid with \( 10^\circ \times 5^\circ \) in order to achieve a better comparability of the data sets. We have ensured that the interpolation implies only minor differences in the spatio-temporal structure of the large-scale circulation modes.

Most climate-change simulations are available in ensemble mode (Li, 1999). This allows one to distinguish between the external part of variance, imposed by the common forcing, and the internal fluctuations, which
Table I. Coupled climate model data sets considered, with integration period, external forcing(s) and number of ensemble members (CONTR: control run; GHG: greenhouse gases; SUL: sulphate aerosols; OZ: ozone)

<table>
<thead>
<tr>
<th>Model</th>
<th>Period</th>
<th>Forcing</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECHAM3/LSG</td>
<td>1880–2049</td>
<td>4 GHG</td>
</tr>
<tr>
<td></td>
<td>1–1749</td>
<td>1 CONTR</td>
</tr>
<tr>
<td>ECHAM4/OPYC</td>
<td>1860–2099</td>
<td>1 CONTR; 1 GHG</td>
</tr>
<tr>
<td></td>
<td>1860–2049</td>
<td>2 GHG + SUL</td>
</tr>
<tr>
<td>CGCM1</td>
<td>1900–2099</td>
<td>1 GHG; 3 GHG + SUL</td>
</tr>
<tr>
<td>CGCM2</td>
<td>1900–2100</td>
<td>1 CONTR; 3 GHG + SUL</td>
</tr>
<tr>
<td>HADCM2</td>
<td>1861–2095</td>
<td>4 GHG; 4 GHG + SUL</td>
</tr>
<tr>
<td>HADCM3</td>
<td>1860–2099</td>
<td>1 GHG; 1 GHG + SUL + OZ</td>
</tr>
<tr>
<td>GISS</td>
<td>1959–2069</td>
<td>3 GHG</td>
</tr>
<tr>
<td>GFDL</td>
<td>1866–2090</td>
<td>3 GHG + SUL</td>
</tr>
<tr>
<td>PCM</td>
<td>1872–2098</td>
<td>5 GHG</td>
</tr>
<tr>
<td>CSIRO</td>
<td>1961–2100</td>
<td>2 GHG + SUL</td>
</tr>
<tr>
<td></td>
<td>1991–2100</td>
<td>3 GHG + SUL</td>
</tr>
<tr>
<td>CCSR/NIES</td>
<td>1890–2099</td>
<td>1 CONTR; 1 GHG; 1 GHG + SUL</td>
</tr>
<tr>
<td>ARPÈGE/OPA</td>
<td>1950–2099</td>
<td>2 CONTR; 2 GHG + SUL + OZ</td>
</tr>
</tbody>
</table>

are induced by varied initial conditions chosen by a random process. Given several ensembles from different climate models, the data constellation is called a multi-model ensemble or superensemble (Krishnamurti et al., 1999; Kharin and Zwiers, 2002). Such multi-model ensembles can be used to quantify the additional effect of intermodel variations in climate-change analysis. In the present study, GHG and GHG + SUL ensemble runs are combined in two superensembles, consisting of four and five climate model ensembles respectively. The non-ensemble simulations are mainly derived from more recent model versions and join this study in order to complete the picture.

2.2. Index definitions and statistical methods

The derivations of the different index time series as a measure of the extratropical circulation modes is as follows. The AO is defined as the first empirical orthogonal function (EOF) of monthly mean SLP north of 20° N. The AAO is the SH counterpart of the AO. For all data sets (models and observations), the EOF analysis is based on the NCEP period 1948–2001, in order to reduce the effect of changes in time of the AO and AAO spatial structure. If this period is not part of the integration period, then the nearest 54 year period has been chosen. It cannot be assumed that the AO and AAO patterns are exactly the same in the observations and various climate models. Therefore, the first five EOFs of each simulation are projected onto the NCEP AO pattern, resulting in a weighting matrix that fits the models’ EOF structure to the observed one. Comparing the weighting factors with each other clearly shows that the leading EOF of each model usually accounts for more than 85% of the spatial variance in the observed pattern (not shown). Apparently, the AO pattern is the predominant characteristic in the SLP field of all climate models. Concerning the long-term observed AO index, the poor data coverage over the Arctic and the North Pacific sector before 1967 has to be taken into consideration. On the one hand, only those grid points that provide continuous data during the 1950–94 period enter the EOF analysis. On the other hand, the varying number of data gaps is taken into account when projecting the historical data onto its leading EOF. Despite these data gaps, it is found that the AO patterns of the historical data set and NCEP are largely identical and the corresponding time series between 1948 and 1994 are in agreement with each other. Thus, for our purpose, the two observational data sets can definitely provide a unified long-term time series of AO and NAO. The AAO index calculation is only based on the 1948–2001 NCEP reanalysis.

As mentioned above, the space–time structure of the AO and AAO is derived from EOF analysis, which implies an optimal linear approximation to the data. However, recent studies have suggested that a more
appropriate description of the extratropical circulation modes may be related to nonlinear decomposition methods (Monahan et al., 2001). Although it is conceivable that the circulation response to global warming is of a nonlinear nature, as has been found with respect to El Niño–southern oscillation (Wu et al., 2003), we rely on the classical linear definition of AO and AAO (Thompson and Wallace, 1998; Thompson and Solomon, 2002) in order to focus on the first-order response of the circulation phenomena and to make our results comparable to previous studies.

We use a Lagrangian definition of the NAO index. It is defined as the leading EOF derived from the four time series of the central pressure and the meridional position of the IL and AH (Glowienka-Hense, 1990; Paeth et al., 1999). This index results from a multi-variate statistical method, but it may be more sensitive if a continuous meridional shifting of the centres of action is part of the signal.

The AL is the centre of cyclonic activity in the North Pacific sector. There are two reasons why the AL has been chosen instead of the PNA index. First, the AL directly counteracts the Pacific pole of the AO (Wallace and Thompson, 2002). Second, the continental centre of the PNA pattern is strongly affected by orography. Since climate models usually fail in simulating the barrier effect of the Rocky Mountains correctly, further uncertainty is imposed when comparing to the PNA index. To avoid this problem, we define the AL index as the standardized spatial-mean of SLP over the oceanic region 30–65°N and 160°E–140°W.

Climate-change signals have to be evaluated against the background of internal variability and systematic differences between the climate models. ANOVA is the appropriate statistical tool to quantify the external forcing part of variance with respect to the uncertainty factors and to give an estimation of the statistical significance (von Storch and Zwiers, 1999; Paeth and Hense, 2002). We apply two approaches. First, given several ensemble members of the same climate model, the total variability of a circulation index can be split up by 1W-ANOVA into the external signal, which is common to all ensemble runs, and internal noise. Second, 2W-ANOVA requires a multi-model ensemble with more or less identical forcing. In this case, a model is assumed that splits the total variability of an index time series into the external part (common to all runs and models), the intermodel variations and the interaction variability, which accounts for differences in the forcing between the models.

3. RESULTS

3.1. Observed time series

The observed AO, NAO, AL and AAO time series as a combination of the historic data set and NCEP reanalysis data are shown in Figure 1. All four times series are characterized by large interannual variations. In addition, there are various decadal components. In the AO index (Figure 1(a)), interdecadal variability intensifies from 1950 onward, which might be partly overestimated by the fact that grid-point information is missing before 1948 in the Pacific and Arctic sectors, and culminates in a remarkable positive trend until the mid 1990s (see Thompson et al. (2000)). At the end of the 20th century, the AO decreases but remains mostly in a positive state. The recent positive trend of the NAO index (Figure 1(b)) is not as pronounced as in the AO, possibly because the effect of the IL and AH displacement is not fully taken into account. This aspect is picked up later. The most prominent feature is the positive phase during the 1920s and 1930s, which was coincident with warm winter temperatures in Europe (Fu et al., 1999). The AL index is interrupted several times when the data coverage is not sufficient (Figure 1(c)). In contrast to the other circulation indices, negative AL anomalies denote a strengthening of the phenomenon. During the 1920s, a distinct negative trend prevails, which looks unrealistic given the otherwise small decadal variations. The 1970 to 1990 period is mainly characterized by strong cycloonic activity over the North Pacific, which has also been reported by Hurrell and van Loon (1997) and Graham and Diaz (2001). In recent years, the AL has not exhibited large transient variability, paving the way for an intensification of the AO (Wallace and Thompson, 2002). The most striking trend, which clearly dominates the interannual fluctuations, is revealed by the AAO (Thompson and Solomon, 2002). It is not clear whether the Antarctic observational data entering the reanalyses and the model skill itself are sufficiently reliable to give credence to this remarkable trend (Hines et al., 2000). These recent dynamics of extratropical circulation, especially the AO and AAO, have been interpreted as an indicator of anthropogenic climate change (Hurrell, 1995; Thompson et al., 2000; Thompson and Solomon, 2002). Thus,
it is of major concern to gain insight into the possible future development of NH and SH circulation as predicted by climate model simulations under increasing GHG scenarios.

3.2. Simulated time series

Subsequent figures are composed in the same way: panels (a) and (b) depict the ensemble-mean decadal time series of GHG and GHG + SUL experiments in ensemble mode. Panel (c) presents the single GHG simulations, and the single GHG + SUL runs are shown in panel (d). The grey shaded area denotes the corresponding temporal-mean 90% confidence interval over all control simulations. For this purpose, the 1749-year control run of ECHAM3 has been split up into 240-year time slices, which results in seven
adequate control periods. These are synchronized with the remaining five control runs. For each decadal-mean within the 240 year period the standard deviation (STD) and 90% confidence interval are computed and finally averaged over all years. Thus, the grey rectangle is a measure of the fluctuation amplitudes in the undisturbed case. Note that the scaling varies from panel to panel.

Figure 2 shows the simulated decadal-mean AO index time series under radiative forcing into the 21st century. At first sight, most ensemble means and single runs predict a strengthening of the AO. In general, the GHG-only experiments reveal an earlier increase from around 1980 onward. The additional SUL forcing seems to delay the trend appearance by a couple of decades. In detail, the model response is far away from being homogeneous: some models like ECHAM3/LSG (Figure 2(a)) and ECHAM4/OPYC (Figure 2(c)) are remarkably sensitive to the enhanced greenhouse effect. Others, like HADCM2, do not show up with long-term changes in the AO (Gillett et al., 2000). In most climate models, the AO intensification is superposed

Figure 2. Simulated 10-year running mean AO index of all climate-change experiments considered: (a) ensemble means of GHG runs; (b) ensemble means of GHG + SUL runs; (c) individual GHG runs; (d) individual GHG + SUL runs. The grey shading represents the corresponding mean 90% confidence interval of all control simulations considered. Note that the scaling is different.
by a strong decadal component with the possibility of blurring out of the trend for several years. This is reminiscent of the drop in the observed AO during the late 1990s. The fact that the ensemble-mean time series are subject to weaker trends than the single runs is due to the averaging over several ensemble members. Actually, Figure 2(a) and (b) is more meaningful than Figure 2(c) and (d), although the latter partly refers to more recent model versions. Evaluating the GHG-induced AO trend against the background of natural variability (grey shading), the disturbed time series usually emerge from behind the control climate very late in the 21st century. However, this picture is somewhat misleading: all index time series are standardized, which implies that the respective long-term mean is zero. In the case of nonstationary time series this leads to the basic structure in Figure 2: mainly negative (positive) anomalies during the first (second) half of the time series. A direct comparison of the AO state amplitudes with the range of control variability is, therefore, difficult. A solution would be to standardize all time series with respect to a common mean value derived from the observational data. However, this is not possible without making prior assumptions, as it is fairly unknown which observed period is actually representative of natural variability: e.g. the NAO was in a remarkably strong phase during the 1920s and 1930s, likely due to solar and volcanic forcing (Stott et al., 2000). Therefore, we are more interested in the temporal dynamics of the circulation indices like the linear long-term trends rather than the decadal-mean state. Thus, the aspect of statistical significance is picked up later when studying the trends. Nonetheless, there is one point to be inferred from the comparison with the control confidence interval: obviously, the AO exhibits large decadal changes even without external forcing. The radiative heating only enhanced the amplitude of the variations in the direction of more frequent and stronger positive departures from the long-term mean.

The AAO response presents a similar picture (Figure 3). The discrepancy between GHG-only and GHG + SUL forcing is less pronounced and the model agreement in projecting a positive trend is more distinct than in the AO case. Again, HADCM2 appears to be less sensitive than other climate models. The single experiments reveal even stronger decadal fluctuations overlying the intensification tendency, and the 90% confidence interval is broader than in Figure 2.

The sensitivity of the NAO, as depicted in Figure 4, is less clear. Although most climate models simulate a slight dominance of positive decadal-mean anomalies into the future, the spectrum of responses varies from markedly positive trends in ECHAM3 to barely any long-term changes in HADCM2 (Osborn et al., 1999). Apparently, the impact of increasing GHG concentrations in the North Atlantic sector is less homogeneous than in the annular modes, incorporating the NH and SH high latitudes. These results are based on the Lagrangian NAO index described in Section 2. Comparing this index with the usual Eulerian one (e.g. Hurrell, 1995) reveals that the trends are generally stronger if the systematic displacement of the centres of action is taken into account in the Lagrangian index (not shown). Thus, an important part of the NAO signal consists of the northward shift of the IL and AH (Paeth et al., 1999; Ulbrich and Christoph, 1999).

The AL response is also quite different from model to model (Figure 5). There is only one climate model, the CSSR/NIES, that predicts a weakening (positive trend of the AL time series) of the cyclonic activity in the North Pacific sector outside the range of the control confidence interval, favouring a positive AO trend. The models, which are characterized by a pronounced AO trend, do not present any considerable changes in the long-term AL time series. However, several models suggest an intensification of the AL during the second half of the 21st century. In particular, this holds for ensemble simulations with GHG + SUL forcing. It is not clear why these ensemble-mean time series differ from the individual GHG + SUL experiments in the bottom panel of Figure 5. In addition, the AL is strengthening in some models, which do not show a prominent transient change in the AO, as well as in HADCM2, although it is subject to a positive trend in the AO. Thus, the AL trend is partly counteracting the AO signal in some models.

The most robust model-comprehensive signal is found in the annular modes, which include the SLP trends over the NH and SH polar caps. The response of the more regionally confined NAO is most heterogeneous. The AL long-term behaviour is partly competing with the AO centre over the North Pacific. The simulated differences in the extratropical circulation response to radiative forcing are not simply a function of the individual model resolutions, different parameterization schemes, the inclusion of ozone dynamics or other known systematic differences between the model designs. Even different versions of the same model lead to slightly different responses (e.g. HADCM2 versus HADCM3 and CGCM1 versus CGCM2). In the following,
the aspect of model intercomparison with respect to the basic regions of model uncertainty and the relative contributions of the circulation modes to the SLP trend pattern are analysed in more detail.

3.3. Model intercomparison

Trends in the circulation modes do not necessarily correspond to the basic pattern of SLP change over the extratropics, because the AO, AAO, NAO and AL only account for limited portions of total SLP variance (mainly less than 30%; Thompson and Wallace, 1998; Paeth et al., 1999). Projecting the spatial structure of the circulation modes, which in the case of the NAO and AL is inferred from regressing the SLP field onto the index time series, gives insight into the relative role of the phenomena in the overall SLP trend pattern (Fyfe et al., 1999). Figure 6 depicts the portion of total spatial variance explained by the four circulation modes as a function of the two different forcings. Concerning the GHG-driven experiments (Figure 6(a)), it
is seen that, on average (crosses), the AO agrees best with the NH SLP trend patterns of the individual model simulations. However, the explained portion of the total trend varies strongly, from 3% to almost 80%. This number largely depends on whether the North Pacific is subject to negative or positive SLP trends. In some models, the main plot is over the Atlantic sector, with the NAO accounting for almost 65% of the SLP trend pattern. Also, in the SH, the AAO pattern is not solely responsible for the SLP changes. GHG + SUL-induced circulation modes project to a lesser extent onto the pattern of SLP trends (Figure 6(b)). In particular, the AO pattern is clearly less relevant than for the pure GHG case, because the additional SUL forcing implies negative SLP trends over the North Pacific. In contrast, the AAO seems to play a more important role over the SH.

The superensemble-mean NH SLP trends and the corresponding STD between the superensemble members are illustrated in Figure 7. The overlapping period of all climate-change experiments, GHG and GHG + SUL,
is 1991–2049. Thus, the SLP changes in Figure 7 denote hectopascals per 59 years. Concerning the GHG experiments (Figure 7(a)), the models seem to agree that SLP is steadily decreasing north of 60°N by about 1.5 hPa. Over the mid-latitude oceans, where the positive centres of the AO pattern are located, the trends are positive but the amplitudes are much less pronounced. The grey shading reveals that the intermodel variations are very strong over the North Pacific and the mountainous regions. The high values of STD over the Arctic are partly explainable by the high absolute values of the trends. As mentioned above, the basic difference in the SLP trend pattern of various climate model simulations with GHG forcing stems from the future behaviour of the AL: some models predict a strengthening and others a weakening. Over the North Atlantic sector, the multi-model mean trend pattern indicates an NAO-like structure with increasing SLP gradient between the AH and IL. However, in the NAO time series in Figure 4 this is not clearly reflected, probably due to the fact that the AH trend does not really prevail in all models. Actually, the highest trend amplitude in the AH region is located northward of the climatological-mean
Figure 6. Contribution of the spatial structure of the circulation modes to the overall trend pattern of SLP, mean (crosses) and minimum–maximum range (bars) over all climate-change experiments considered: (a) GHG runs; (b) GHG + SUL runs.

Figure 7. Mean (contour lines) and STD (grey shading) of SLP trends 1991–2049 over all climate change experiments north of 20°N: (a) GHG runs; (b) GHG + SUL runs.

position of the AH. In the GHG + SUL simulations (Figure 7(b)), the trend pattern is also characterized by decreasing SLP over the Arctic, albeit of lower amplitude. In contrast to the GHG forcing, the AL is clearly deepening, as shown in Figure 5, and the strengthening of the SLP gradient over the North Atlantic occurs much more to the north. For the experiments presenting a positive AO trend in Figure 2, this tendency is mainly supported by the Atlantic centre and the Arctic rather than by the North Pacific sector. We do not have a conclusive explanation for why the additional SUL forcing leads to an inverse

Copyright © 2004 Royal Meteorological Society

reaction of the climate models in the North Pacific compared with the GHG-only forcing. Hypothetically, anthropogenic sulphur emissions from East Asian sources might be advected by westerlies to the AL region, modifying the radiative effect of increasing GHG concentrations. Comparing the superensemble-mean patterns with the observed NH trend pattern of SLP during the second half of the 20th century (Thompson et al., 2000: Figure 1), reveals that the observed trend pattern appears to be a mixture of the two patterns in Figure 7: the polar cap is subject to remarkably negative trends, whereas positive anomalies occur over the mid-latitude North Atlantic basin. This is in perfect agreement with the GHG-induced response pattern. However, lower SLP is developing over the northern Pacific, which is rather a characteristic of the GHG + SUL forcing.

In the SH, the superensemble-mean trend pattern describes a more or less zonally symmetric increase of the meridional SLP gradient, indicative of a stronger AAO (Figure 8). Despite the distinctly negative SLP trends over Antarctica, amounting to 1.5 hPa per 59 years on average, model uncertainty is considerable in this region (Paeth and Hense, 2002). The trend patterns of both radiative forcing types correspond to a large degree, although the amplitudes are generally weaker in the GHG + SUL case. The SH observed SLP trend pattern can be found in Thompson and Solomon (2002: Figure 3). It looks highly similar to the superensemble-mean patterns in Figure 8, enhancing our belief in the spatial structure of the multi-model response.

Altogether, it is found that state-of-the-art climate-change experiments mainly agree in predicting a deepening of surface pressure over the NH and SH polar caps, whereas the dynamics over the mid latitudes are subject to large model uncertainty. Thus, these results suggest that index time series, representative of these high-latitude regions, might be a more appropriate indicator of the sensitivity of extratropical circulation to global warming.

Copyright © 2004 Royal Meteorological Society

3.4. Trend significance

In order to gain insight into the statistical properties of the long-term trends in Figures 2–5, the linear trend significance of all index time series has been determined for different time windows. As the signals mostly occur from the 1980s onward, the longest time window spans the period from 1980 until the end of each model integration. Then this window is successively shortened by one year with respect to the starting year until, finally, the period of trend calculation extends over the last 40 years of each climate-change experiment. These results are compared with the trend significance over corresponding time slices (≥40 years), running through the control time series. Thus, we can infer at which time scales significant long-term trends in the climate-change experiments are not consistent with low-frequency internal fluctuations in the undisturbed model runs. As described above, the majority of circulation trends are positive. Negative trends are denoted by triangles. The trend significance of the ensemble-mean AO time series is displayed in Figure 9. Note that the different length of the bars relate to the different lengths of the individual ensembles’ integration periods. Multi-decadal AO trends are predominantly significant at a high level (1%) after 1980. Among the GHG-only and GHG + SUL experiments, HADCM2 represents an exception, with almost no significant changes. The signal deterioration towards the end of the integration period is partly due to the decreasing number of degrees of freedom. However, negative trends suddenly occur in the parallel climate model (PCM) during the second half of the 21st century. Paeth et al. (1999) have also reported that the NAO signal-to-noise ratio is weakening at the end of the integration period. It is hypothesized that negative feedbacks with the world’s oceans may stabilize the circulation modes under radiative forcing. The observed AO trends after 1950 are also statistically significant at least at the 5% level (not shown). The control simulations also reveal some significant multi-decadal changes, indicating that low-frequency AO variations are intrinsic to the coupled climate system. However, the number of such trends is small relative to the total number of control time slices, and the highest level of significance is not less than 10% (not shown). This is true for all circulation indices considered. Thus, the long-term trends in Figure 2 mostly emerge from behind natural variability in the control runs.

Figure 10 depicts the trend significance of the AL index time series in Figure 5. As mentioned above, negative trends, equivalent to an enhancement of the cyclonic activity over the North Pacific, prevail especially under the GHG-induced forcing. Among all ensemble simulations, ECHAM3, the top proclaimer of a positive AO trend, is exceptional in prognosticating a weakening of the AL. In general, the future AL dynamics are less affected by the GHG-only conditions than by the additional SUL forcing.

The significance of the AAO trends is similar to the AO, but the intermodel variations are more pronounced (not shown). Given the stronger decadal component in the AAO time series in Figure 3, no significant trends are found over periods of less than 60 years. The trend observed after 1950 is highly significant at the 1% level. Thus, climate models seem to underestimate the dynamics of SH near-surface pressure, provided that these are correctly modelled in the NCEP reanalysis. The NAO does not exhibit significant long-term changes in the various climate models, except in ECHAM3, even when the Lagrangian index is taken into consideration (not shown). Obviously, the North Atlantic sector is not the Achilles heel of extratropical circulation in present-day climate models. Rather, those circulation modes that incorporate the polar regions undergo significant changes.

3.5. Analysis of variance

The data constellation of this study is suitable for carrying out 1W-ANOVA and 2W-ANOVA (see Section 2.1). Separating the impact of the external radiative signal from internal noise, 1W-ANOVA is applied to the annual-mean time series of each climate-change ensemble. The results are quite heterogeneous and analogous to the inferences made about the time series in Figure 2–5: in some models, like ECHAM3, GISS, CGCM1 and CSIRO, the GHG-induced part of variance amounts to more than 30% and is highly significant. The remaining model ensembles do not show up with a consistent signal (not shown).

More insight into the signal strength is revealed by 2W-ANOVA, which includes the impact of model uncertainties. Figure 11 shows the contributions of the radiative forcing, GHG-only, GHG + SUL and of the systematic differences between the climate model ensembles and total interannual variability of the AO,
Figure 9. Statistical significance of AO trends over varying time windows from 1980 until 40 years up to the end of the integration period; ensemble means of (a) GHG runs and (b) GHG + SUL runs. The triangles indicate negative trends.

NAO, AL, and AAO. The grey bar indicates where the impacts exceed the 5% significance level. Note that the overlapping periods of the multi-model ensembles are different for GHG and GHG + SUL, the ensembles of which cover the periods 1959–2049 and 1991–2090 respectively (see Table I and Figures 2–5). The method is based on running 60-year time slices, with the x-axis referring to the central year of each period considered. Thus, the 2W-ANOVA results are presented in the time window 1989–2020 for the GHG forcing and 2021–2061 for the combined forcing. The fact that the two windows do not overlap is simply due to the uncoordinated simulation periods of the various climate models. Most circulation modes are subject to a significant radiative forcing influence common to all multi-model ensemble members (Figure 11, right-hand-side panels). However, the magnitude is very weak, varying from 5% in the NAO to 10% in the AO. The main player is internal variability, imposed by changing initial conditions. Thus, the enhanced greenhouse effect slightly modifies the otherwise strong interannual fluctuations of extratropical circulation towards more frequent positive (AL: negative) anomalies. Assuming a constant increase of anthropogenic trace gases and aerosols into the 21st century as prescribed by the IPCC scenarios, it is expected that consistent climate-change signals continuously stand out from internal variability. Except for the AO signal under GHG + SUL forcing, the right-hand-side panels in Figure 11 do not agree with this idea: the externally explained variance does not increase steadily. Rather, it remains at the same level, or even decreases, with rising GHG concentrations. This is mainly due to the fact that the future projections of various climate models increasingly diverge in time, as shown in Figures 2–5. Consequently, the influence of intermodel variations is also increasing in time (Figure 11, left-hand-side panels). Comparing the two columns in Figure 11 with each other demonstrates that the common external signal is mainly of the same order of magnitude as the model uncertainty. Thus, it is hardly possible to detect a climate-change signal in extratropical circulation against the background of the remarkable year-to-year changes and differing model predictions. The impact of the interaction component (see Section 2.2) is negligible, since the forcing scenarios are largely comparable to each other. At first sight, these findings appear somewhat inconsistent with the previous results, revealing distinct and significant long-term...
trends, especially in terms of the AO. However, note that the 2W-ANOVA is based on the annual-mean time series and the year-to-year changes are largely dominated by internal atmospheric variability, deteriorating the signal-to-noise ratio. Indeed, much larger impacts (22%) are inferred from decadal-mean time series; but, in this case, the degrees of freedom are not sufficient to test the statistical significance properly (not shown).

4. CONCLUSIONS

In this study, we present a broad model intercomparison of climate-change signals in extratropical circulation. State-of-the-art coupled climate models do not present a homogeneous picture of circulation trends into the 21st century. Whereas the AO and AAO are supposed to intensify in the majority of climate-change experiments, the NAO is barely sensitive in most models. In detail, it is found that a large part of the NAO changes is related to a northward shift of the centres of action rather than to an intensification of the pressure gradient. The future development of the AL is subject to large model uncertainty and depends on whether or not additional SUL forcing is imposed. In particular, the annular modes, AO and AAO, are closely tied to the trend pattern of SLP under global warming. All models, without exception, predict a deepening of surface pressure over the NH and SH polar caps, whereas strong intermodel variations prevail over the NH mid latitudes. The multi-decadal AO and AAO trends, and in part the AL trends, are statistically significant and not consistent with low-frequency fluctuations in the control runs. At the interannual time scale, the radiative forcing impact is statistically significant, but it barely stands out from model uncertainty and internal variability, as revealed by 2W-ANOVA. A much clearer signal occurs if the method is applied to decadal-mean time series, with the external forcing accounting for up to 22% of total variability. Thus, climate-change signals in simulated extratropical circulation will only be detectable at longer time scales from decadal to centennial.

On evaluating the intermodel variations with respect to the main characteristics of each climate model, we find no systematic relationships between the model sensitivity on the one hand and model properties like
horizontal and vertical resolution, inclusion of ozone dynamics (Shindell et al., 1999), parameterizations, etc. on the other hand. The amplitude and direction of the response are even functions of the model version, like CGCM2 versus CGCM1 and HADCM3 versus HADCM2. In recent years, the Atmospheric Model Intercomparison Project hindcast studies have provided some insight into the relative realism of various models in terms of 20th century climate (Mao and Robock, 1998). However, as it is not clear which model’s prediction is most reliable in terms of future climate, the present-day variety of climate models is essential to assess the uncertainties in climate-change projections. Paeth and Hense (2002) have pointed out that climate-change
signals in near-surface temperature, and especially rainfall, are blurred by model discrepancies in many regions over the globe. This is also true for extratropical circulation. Thus, it is not useful to study single climate-change simulations or even Monte Carlo experiments derived from one climate model. Selecting ECHAM3 would result in completely different implications for decision makers than would choosing HADCM2.

Index time series that describe the SLP dynamics over the polar regions might be the most appropriate measure of climate-change sensitivity in extratropical circulation. The main model disagreement concerns the mid-latitude centres of the extratropical circulation modes, particularly in the North Pacific and North Atlantic basins. This implies that there is no consistent climate-change signal in the simulated NAO time series. On the other hand, the SLP gradient between the mid and high latitudes, as depicted by the conventional circulation indices used here, is more relevant to circulation, advection processes and regional climate than to an isolated deepening of SLP over the polar caps. Other useful indices might be constructed from Figures 7 and 8, especially since it has been found that, on average, the circulation modes considered only account for a relatively small part of the overall long-term SLP changes. Thus, additional processes and phenomena must be at work.

According to the general idea of the SUL counteracting the greenhouse forcing, the inclusion of SUL usually reduces the simulated response of the circulation modes to radiative heating. However, a basic difference between the two anthropogenic scenarios is found over the North Pacific sector. Increasing SUL concentrations induce a strengthening of the AL into the 21st century, whereas GHG forcing rather presents positive SLP anomalies. If this discrepancy is not a climate-model artefact, related to a deficient presentation of the sulphur cycle and transport, then one physical explanation might be that sulphur emissions over East Asia, representing one of the main source regions, are advected to the North Pacific region where they may fundamentally affect the cyclogenesis and the strength of the AL. Given that all models, except ECHAM3, predict this different AL response to GHG and GHG + SUL forcing, this provides some evidence that, indeed, such a physical mechanism is active.

The AO has been found to provide the clearest climate-change signal in the multi-model ensemble data. Returning to the observational time series in Figure 1, the question arises as to why the observed AO has decreased in recent years whereas most climate models agree in predicting a strengthening under enhanced greenhouse conditions. There are three explanations: (1) the models are completely wrong and there is no AO sensitivity to global warming in the real climate; (2) the externally induced trend is superimposed by low-frequency random variability at time scales less than 5 years; (3) the AO is subject to additional forcing mechanisms other than radiative heating, e.g. solar variability and volcanic aerosols, which counteract each other and may even induce decadal-scale interruptions of the multi-decadal to centennial changes (Stott et al., 2000). Moreover, this conclusion holds the risk that the climate models considered, which do not account for such mechanisms, are not able to draw a realistic picture of future AO changes. To the same extent, the control experiments might underestimate the magnitude of internal variability. Anyhow, the AO behaviour in the near future will provide further evidence of its sensitivity.

ACKNOWLEDGEMENTS

We thank Julie Arblaster, Tom Delworth, Ruth McDonald, Tracey Elliot, John Fyfe, Jean-François Royer and Drew Shindell for making the various climate model data sets available to us and for their technical support. In addition, we are grateful to the two anonymous reviewers whose remarks helped to improve the paper.

REFERENCES


