Climate Dynamics (2001) 17: 61–81

M. Collins · S. F. B. Tett · C. Cooper

The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments

Received: 24 August 1999 / Accepted: 28 April 2000

Abstract We examine the internal climate variability of a 1000 year long integration of the third version of the Hadley Centre coupled model (HadCM3). The model requires no flux adjustment, needs no spin up procedure prior to coupling and has a stable climate in the global mean. The principal aims are (1) to validate the internal climate variability against observed climate variability, (2) to examine the model for any periodic modes of variability, (3) to use the model estimate of internal climate variability to asses the probability of occurrence of observed trends in climate variables, and (4) to compare HadCM3 with the previous version of the Hadley Centre model, HadCM2. The magnitude and frequency characteristics of the variability of the global mean surface temperature of HadCM3 on annual to decadal time scales is in good agreement with the observations. Observed upward trends in temperature over the last 20 years and longer are inconsistent with the internal variability of the model. The simulated spatial pattern of surface temperature variability is qualitatively similar to that observed, although there is an overestimation of the land temperature variability and regional errors in ocean temperature variability. The model simulates an El Niño Southern Oscillation with an irregular 3–4 year cycle, and with a teleconnection pattern which is much more like the observations than was found in HadCM2. The interdecadal variability of the model ocean in the tropical Pacific, North Pacific and North Atlantic is broadly similar to that in the real world with none of the simulated patterns having any periodic behaviour. HadCM3 simulates an Arctic Oscillation/North Atlantic Oscillation (NAO) in Northern Hemisphere winter which has a spatial pattern consistent with the observations in the Atlantic region, but has too much teleconnection with the North Pacific. The recent observed upward trend in the NAO index is inconsistent with the model internal variability. The variability of the simulated zonal mean atmospheric temperature shows some marked differences to the observed zonal mean temperature variability, although the comparison is confounded by the sparse observational network and its possible contamination by a climate change signal.

M. Collins (✉) · S. F. B. Tett · C. Cooper
Hadley Centre for Climate Prediction and Research,
Meteorological Office, London Road, Bracknell, RG12 2SZ, UK
E-mail: matcollins@meto.gov.uk

1 Introduction

Many coupled ocean-atmosphere models have had a tendency to drift when simulating the climate of the present day. Such climate drifts are usually reduced by the use of flux adjustment (e.g. Sausen et al. 1988; Manabe et al. 1991), whereby often large fluxes of heat and salinity are introduced at the interface between the ocean and the atmosphere components of the model. Such adjustments are clearly unphysical and, to this end, there have been considerable efforts to build coupled ocean atmosphere GCMs that do not require these corrections. Such a model has been built at the Hadley Centre at the UK Meteorological Office. HadCM3, the third version of the Hadley Centre coupled model, requires no such flux adjustment and has a stable and realistic present day mean climate (see Gordon et al. 2000 for more details).

We focus on the "internal" variability of a 1000 year simulation of HadCM3 in which all "external" factors, such as increases in greenhouse gases or variations in solar output, are kept fixed. Thus we define internal variability as that which is consequence of the internal dynamics of the coupled ocean-atmosphere system. Climate variability that is forced by anthropogenic emissions of greenhouse gases, solar variability etc. will be dealt with elsewhere. A successful simulation of internal variability in a coupled model can be of importance for several reasons. Variability on interannual to decadal time scales (e.g. the El-Niño Southern Oscillation, ENSO) has a significant impact on society, and the

ability to predict such changes is reliant on a successful simulation of the phenomenon in question. Also the detection and attribution of anthropogenic and natural climate change is crucially reliant on a realistic simulation of the internal unforced variability (e.g. Hasselmann 1993). The prediction of future climate change is now being extended to include not only predictions of changes in the mean climate but also changes in variability, for example, ENSO (e.g. Timmermann et al. 1999; Collins 2000). Such predictions rely on the ability of models to reproduce present day climate variations.

Multi-century simulations of internal variability allow us to answer many questions about the climate system. For example, can we identify any periodic phenomena in the coupled ocean-atmosphere system? We define a periodic phenomena as having a peak in its power spectra which is significantly different from the spectra of an autoregressive process of order 1 (AR(1)). An AR(1) process is defined by the following equation:

$$x_{n+1} = a_1 x_n + a_0 z \ , \qquad (1)$$

where x is a measure of the phenomenon under test (e.g. global mean surface temperature), the subscripts refer to discrete time intervals (e.g. years) and z is random variable (white noise) with unit variance. The coefficient a_1 is the autocorrelation of x. AR(1) processes have more power at low frequencies than at high frequencies (i.e. they have a red spectra) and are a good statistical model for many geophysical systems. Hasselmann (1976) describes a conceptual model of the coupled ocean-atmosphere system in terms of an AR(1) process in which the ocean gives the "memory" of the system and the atmosphere provides the random forcing. Identification of periodic phenomena may lead to possible predictability of weather and climate beyond the usual predictability horizon associated with sensitive dependence on initial conditions. Also we can ask, are recent observed trends or climate variations outside the range of internal climate variability? This detection of climate change then leads on to questions of attribution of climate change to factors such as anthropogenic increases in greenhouse gases (e.g. Tett et al. 1999).

A necessary pre-requisite for answering these, and other, questions about the climate system is the verification of the internal variability of the coupled model. Hence a large part of this work is devoted to the comparison of the model variability with observed variability. Uncovering the mechanisms for variability in the coupled model requires much more detailed analysis than we have space to include here, as well as further sensitivity experiments with the model. It is hoped that mechanisms will be dealt with in more detail in future publications. We focus our attention on the variability on interannual to decadal and longer time scales and to global and large-scale variations in HadCM3. Also, we mainly focus on the surface temperature because of its relevance to the climate change question, and because of the extensive observational records that exist for this variable. However we also briefly examine other variables where it is appropriate to do so. Where observational data exists, we perform the same analysis on both the model and the observations to reduce uncertainties in the comparison. Also we compare the variability of HadCM3 with the previous flux adjusted version of the model, HadCM2 (Johns et al. 1997; Tett et al. 1997), and note any improvements or reductions in skill.

The study is organised as follows. The next section gives a description of the formulation of HadCM3. This is followed by Sect. 3 which includes an assessment of the magnitude and spectral characteristics of the global mean surface temperature, as well as an analysis of linear trends in the model and in the observations. In Sect. 4, the spatial pattern of surface temperature variability is validated against observations and is examined for any periodic phenomena. The variability over the Northern Hemisphere land masses is next examined in Sect. 5, to see if there is any improvement in this over HadCM2, which simulated too much variability in these regions (Tett et al. 1997). In Sect. 6 we examine the ENSO-like variability in HadCM3 and in Sect. 7, the decadal variability of Equatorial Pacific, North Pacific and North Atlantic sea surface temperatures (SSTs) are examined and compared with the observations. In Sects. 8 and 9 the variability of the atmosphere is examined in terms of the Arctic Oscillation/North Atlantic Oscillation and in terms of the variability of the zonal mean atmosphere. Finally the results are summarised in Sect. 10.

2 Model formulation

The atmospheric component of HadCM3 is a version of the UK Meteorological Office Unified Model (Cullen 1993). The model dynamics and physics are solved on a $3.75° \times 2.5°$ longitude-latitude grid with 19 hybrid vertical levels. Significant changes with respect to the previous version (HadCM2, Johns et al. 1997) are the introduction of a new radiation scheme (Edwards and Slingo 1996; Cusack et al. 1998), the representation of convective momentum transport (Gregory et al. 1997), the introduction of a new surface scheme (Cox et al. 1999), better representation of the effects of subgrid scale orography (Milton and Wilson 1996; Gregory et al. 1998), a reformulation of the treatment of clouds in the model (see Gordon et al. 2000 for more details) and a revision of the boundary layer scheme (Smith 1993). More details of the formulation and climate of the atmospheric component of the model can be found in Pope et al. 2000.

The oceanic component of the model is an updated version of that used in HadCM2 (Johns et al. 1997) which is version of the Cox (1984) model. The horizontal resolution is increased to $1.25° \times 1.25°$ in comparison to HadCM2 which used a $3.75° \times 2.5°$ longitude-latitude grid. HadCM3 employs the Gent and McWilliams (1990) scheme for adiabatic diffusion of tracers with variable coefficients (Wright 1997; Visbeck et al. 1997), vertical mixing is parametrised using Kraus-Turner scheme (see Gordon et al. 2000 for more details) and salinity conservation is achieved using a representation of river runoff and by assuming a balance between iceburg calving and accumulation of snow on the ice-sheets.

The model is initialised in September using a previously derived model atmospheric state and the September ocean observations of Levitus and Boyer (1994) with zero ocean currents. The atmosphere and ocean are coupled once a day. We analyse the full 1000 years of a control simulation in which all concentrations of greenhouse gases and aerosols etc. are set as constants representative of the pre-

industrial era. There is a small drift in the model such that the global mean temperature cools by approximately one tenth of a degree in the entire 1000 year period. The spatial pattern of this drift is uniform over much of the globe, although there are positive and negative drifts of order one degree in the Southern Ocean and North Atlantic. These drifts are removed prior to the analysis presented later. Gordon et al. (2000) discuss the stability of the mean climate of the model in considerable detail.

3 Global mean surface temperature variability

Firstly we examine the global mean surface temperature variability of HadCM3 and compare this with the observed variability. The observational data we use is the monthly mean Jones (1994) gridded land surface station temperature data blended with the gridded sea surface temperature observations of Parker et al. (1995) (each data set has been updated to 1998 by colleagues at the Hadley Centre). This we denote this the HadCRUT data set (Fig. 1). The data is binned into a 5° by 5° longitude-latitude grid and has variable coverage in both space and time. Jones et al. (1997b) showed that the standard error, due to incomplete sampling, on interannual time scales for the global and annual mean of the HadCRUT data is ± 0.12 K for the later part of the record, rising to ± 0.18 K for the earlier part.

In order to asses the effect of these uncertainties which result from variable data coverage in the observations, it is important to process the model data in a similar way to the observations i.e. to interpolate the model data onto the HadCRUT grid and to mask areas where HadCRUT has no observations. First we take the monthly average 1.5 m temperatures from the model and remove the mean annual cycle. Then, we split the first 1000 years of the HadCM3 control into ten 100 year segments, and re-grid them from the $3.75° \times 2.5°$ longitude-latitude model grid to the $5° \times 5°$ HadCRUT grid using bilinear interpolation. We then remove grid boxes which have missing data in HadCRUT, in order to have the same changing data coverage in both space and time. Thus we are left with 10 independent realisations of the HadCM3 control global temperatures on the same space-time grid as the HadCRUT data. We denote these realisations as "processed model data" and we compare these with 1898–1998 HadCRUT data and also with the unprocessed model data where appropriate.

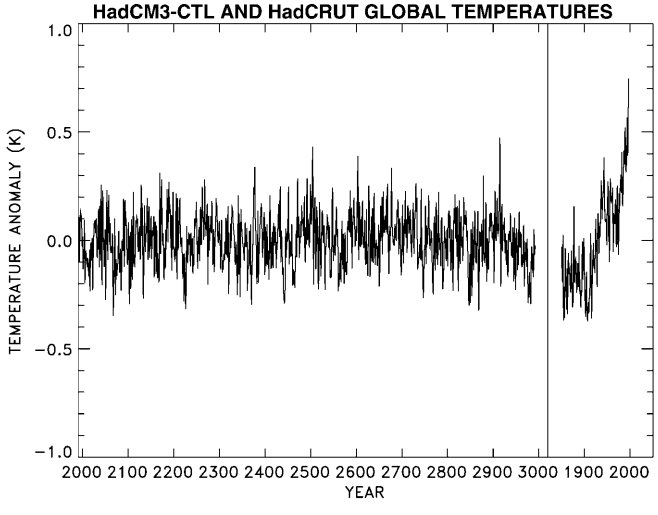

Fig. 1 The global and annual mean 1.5 m temperature from 1000 years of the HadCM3 control experiment (*left curve*) and the global and annual mean historical observed HadCRUT surface temperature from 1851 to 1998 (*right curve*). The time scale marked on the *abscissa* for the model curve is essentially arbitrary although the nominal model start year is 1991. The model time series has had a long term trend of −0.010 K/century removed from it (see Sect. 2)

A simple measure of the amplitude of the global temperature variability is the standard deviation. The standard deviation of the HadCM3 annual and global mean 1.5 m temperature shown in Fig. 1 is 0.12 K for the unprocessed data and 0.11 K for the square root of the average of the variance of the 10 sections of the 1000 years processed on the HadCRUT grid. There is poor observational data coverage in the northern high latitudes and in the southern middle to high latitudes and these are regions where the model variability is generally large compared to the rest of the globe (see Sect. 4). Thus their exclusion from the calculation of global mean temperature variability reduces the standard deviation, although this effect is tempered by the small area covered by the missing data regions.

A strict comparison with the standard deviation of the observed temperatures is complicated by the removal of the obvious trend from the observations (see Fig. 1). The standard deviation computed from the observations with a linear trend removed is 0.14 K and that computed from the observations with a quadratic trend removed is 0.13 K. Hence, subject to uncertainties in the removal of the trend from the observations, HadCM3 has a good simulation of the magnitude of the global mean surface temperature variability.

The standard deviation of global mean temperature from HadCM2 was 0.13 K, hence there is very little difference in the absolute magnitude of the global mean temperature variability between the two models. We might have expected HadCM3 to have greater global mean temperature variability than HadCM2 because of the increased ocean resolution, and the removal of the flux adjustment term, but this is not the case. Also, the ratio of the variance of global mean land and global mean ocean temperature has the same value ($\sim$3) for both models, so there are no compensating factors due to changes in ocean resolution and the land surface scheme. Thus the use of flux adjustment seems to have little impact on the magnitude of global surface temperature variability between HadCM3 and HadCM2.

In comparison with other coupled modes (Covey et al. 2000), both HadCM2 and HadCM3 seem to be at the upper end of the distribution, having larger global temperature variability than other models. While the agreement in the standard deviations of the models and the observations is good, this agreement might not be so

good when we include the effects of external factors such as changes in solar output or volcanic eruptions which may increase the model variability. Also we reiterate the uncertainty due to the removal of the anthropogenic trend from the observations. Hence it may be too soon to conclude that HadCM3 has an adequate simulation of global temperature variability. Work is underway however, to address some of the uncertainties highlighted by including all the major external factors in one simulation.

We next examine the power spectra of global mean temperature, firstly to verify of the model has a similar spectral shape to the observations, and secondly to see if there are is periodic behaviour. To calculate the power spectrum of global mean temperature we first take the Fourier Transform of the autocovariance function of the time series and then apply the Tukey–Hanning window to get a consistent and unbiased estimate of the power spectrum (e.g. Chatfield 1984). We chose a relatively large window width of 25 years in order to retain part of the low frequency component of the variability (up to a period of 50 years). However, there is little change in the results, apart from the loss of the low frequencies, when we take smaller window widths.

The power spectra of 1898–1998 HadCRUT global temperatures (Fig. 2) is red in character, that is there is more power at low frequencies than at high frequencies. A linear trend was first removed from the time series to crudely remove the climate change signal, although it is possible that some residual signal remains after this is done and this residual signal aliases into the lowest frequency of the power spectra which is therefore unreliable. Tett et al. (1999) discuss the non-uniformity of the twentieth century warming in detail. Testing the power spectra of observed temperatures against a null hypothesis of an AR(1) process gives no significant spectral peaks, indicating that there are no periodic modes of the observed global mean surface temperature.

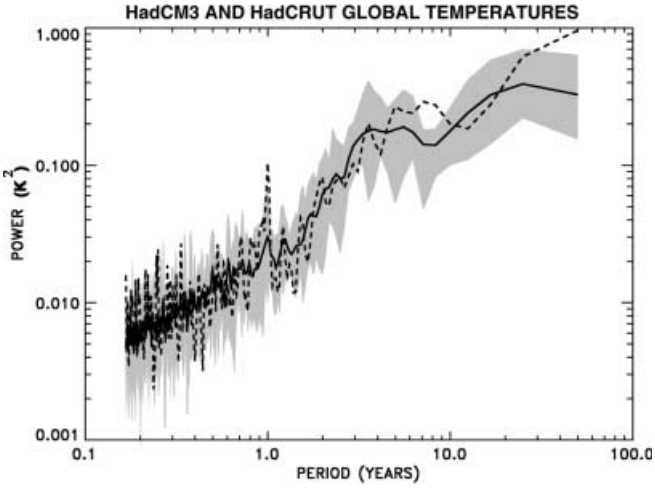

Fig. 2 Power spectra of the global mean surface temperatures from the HadCM3 control run (*solid line*) and from the HadCRUT observations (*dashed line*). The *shading* indicates the maximum and minimum power computed from ten 100 year sections of the model control interpolated to the HadCRUT grid. The *solid line* is the average of the power spectra of these ten sections

The power spectra of the HadCM3 global mean temperatures, computed as an average of the 10 power spectra of the 100 year segments processed on the HadCRUT grid (Fig. 2) is also red in character and has a similar slope and magnitude to the observed spectra. Also shown in Fig. 2 is the range of variability of the model power spectra computed as the maximum and minimum of the 10 power spectra of the 100 year segments. This represents a median estimate of the 93% interval (Gillett et al. 2000) and if the power spectra of the observed HadCRUT global temperatures lies outside this range we may conclude that there is a statistically significant difference (at the 93% level) between the model and the observations. Such deviations occur at periods of 1 year, which is an artifact of the removal of the annual cycle from HadCRUT (calculated over the period 1961–1990 only), and at around the 8 year period which may be put down to chance. There is also some difference at the lowest frequency, which may be due to the aliasing of the residual variability (i.e. that left after the linear trend is removed) into the observed spectra. It may also be the case however, that there may be some multi-decadal to centenial period fluctuation of the climate system (e.g. Schlesinger and Ramankutty 1994) which the model does not simulate. Validation of the model variability on time scales of centuries and greater is limited by shortness of the observed climate record. However we may be able to use proxy indicators of climate, for example tree-rings (e.g. Jones et al. 1998), for such a validation. Nevertheless we conclude that, on time scales of less than a century, HadCM3 reproduces well the observed spectral characteristics of global mean temperature variability.

As was the case for the observations, the model shows no peaks in the spectra which are significantly different from an AR(1) process, indicating that the are no periodic modes of variability in the models' global mean temperature. This however, does not preclude the existence of regional periodic modes of variability, which we look for in the next section.

Having established that HadCM3 global temperatures have similar variability to the observations in terms of both the absolute magnitude (standard deviation) and in terms of the spectral characteristics, it is pertinent to use the model estimate of internal variability to asses how unusual, or not, recent trends in global temperature are. The global and annual mean HadCRUT temperatures show a clear upward trend over the 148 year period (1850–1998, Fig. 1) and linear trends computed over the most recent 20, 50, 100 and 148 year periods are 1.9, 0.9, 0.6 and 0.4 K/century respectively. We can compute the probability of occurrence of observed trends from the 1000 year control run to asses how unusual such trends are. Following Stouffer et al. (1994) we first find the linear trend of the 5 years (1993–1998) of the HadCRUT temperature series. Next we compute

the trends of 995 overlapping segments of the HadCM3 series. We count the occasions on which the trend of the 5 year HadCM3 segment is greater than that of the observed 5 year segment and divide that number by the total number of model segments (995 in this case). This gives a relative occurrence rate for the 5-year trend which is a number between 0 (no model trends are greater than the observed trend) and 1 (all model trends are greater than the observed trend). We then repeat this analysis for the observed and modelled 6 year trends, for 7 year trends etc. up to 148 year trends.

A selection of different length observed global mean temperature trends and their occurrence in the 1000 year HadCM3 control are shown in Table 1. On examination of the trends for all time periods, we find that the observed trends of length 9–12 years have a greater than 10% chance of occurring in the control simulation, and thus we may conclude that they can be easily explained by internally generated variability. Observed trends of length 13–16 years occur less frequently and are therefore less likely to be due to internal variability. Trends of 17–20 years length occur very infrequently (less than 1% chance) in the control simulation and there is only 1 occurrence of a 20 year trend of 1.9 K/century in the entire 1000 years of the model simulation. Such trends are very unlikely to be generated by the internal variability of the climate system. There are no occurrences of any trends of length greater that 20 years in the control, thus, we conclude that it is highly unlikely that observed trends in global temperature from before 1978 to 1998 are a consequence of the internal variability of the climate system. This leaves the question of attributing these temperature trends to various possible external forcing mechanisms such as increases in greenhouse gases. Tett et al. (1999) examine this question in detail, although they focus on periods of 50 years in length. Given the results of the trend analysis, it may now be possible to examine the attribution question over shorter time periods where, e.g., tropospheric temperatures may be better constrained by satellite observations and forcing mechanims, such as variations in solar output, have been observed more quantitatively.

Table 1 Trends in the observed global temperature and corresponding occurrences of trends greater than the observed trends computed from the 1000 year HadCM3 control run. Occurrences are quoted in both actual and relative terms (see text for more details)

Years (inclusive)	Length of trend (years)	Trend (K/century)	Occurrence	Relative occurrence
1994–98	5	7.0	32	0.03
1989–98	10	2.9	64	0.06
1984–98	15	2.8	2	0.002
1979–98	20	1.9	1	0.001
1974–98	25	2.2	0	0
1949–98	50	0.9	0	0
1899–98	100	0.6	0	0
1850–98	148	0.4	0	0

4 The spatial pattern of surface temperature variability

In this section we examine the spatial pattern of surface temperature variability in the observations and in the model. The standard deviation of monthly mean HadCRUT temperatures calculated from 1898–1998 are shown in Fig. 3a. A linear trend plus annual cycle was removed from each grid-box of the HadCRUT data prior to the calculation of the standard deviation. Also, a requirement of having at least 10 years of data in each grid-box in order to estimate the standard deviation was imposed. The general pattern of variability in Fig. 3a is one of greater variability over land areas than over ocean areas. This can be understood by recalling that land has a lower "thermal inertia" than the ocean so its temperature can respond more rapidly to variations in the surface energy balance. The northern high latitudes are the most variable region in terms of temperature because of variations in albedo caused by changes in snow cover. In the ocean there is an area of enhanced temperature variability in the tropical East Pacific associated with ENSO.

The model surface 1.5 m temperatures were processed in the same way as the observations (with the same missing data regions etc., Sect. 3) for the ten 100 year sections. The variance in each grid-box was then averaged over these 10 sections and the square-root taken to find the model average spatial pattern of standard deviation (Fig. 3b). There is broad agreement between the model and the observations with the model having more variability over the land than over the ocean and with the maximum standard deviations in the model occurring over land in the northern high latitudes. Notable differences between model and the observed variability are in the North Atlantic region, where the model overestimates variability and in the Southern Ocean, south of Africa, where the model underestimates variability. Errors in the North Atlantic region are likely to be due to sea ice in the model which moves too far south during winter effecting the baroclinic zone of the storm track. The reasons for errors in the Southern Ocean are less obvious.

In terms of the ratio of model to observed variance (figure not shown), the model tends to overestimate variability over land in the sub-tropics and over Central Europe and underestimate it in the tropics and in polar regions. Over the ocean there is no systematic pattern of variability errors with some regions where the model overestimates variability (the North Atlantic, the sub-tropical North and South Pacific) and other regions where the model underestimates variability (the North Pacific and the tropical and South Atlantic). As was the case for global mean temperatures, the comparison of variance in this way is confounded by the removal of the climate change signal from the observations, so it is difficult to accurately quantify errors in model surface temperature variability.

We have examined the AR(1) coefficients (Eq. 1) of the model and observed temperature in order to assess if

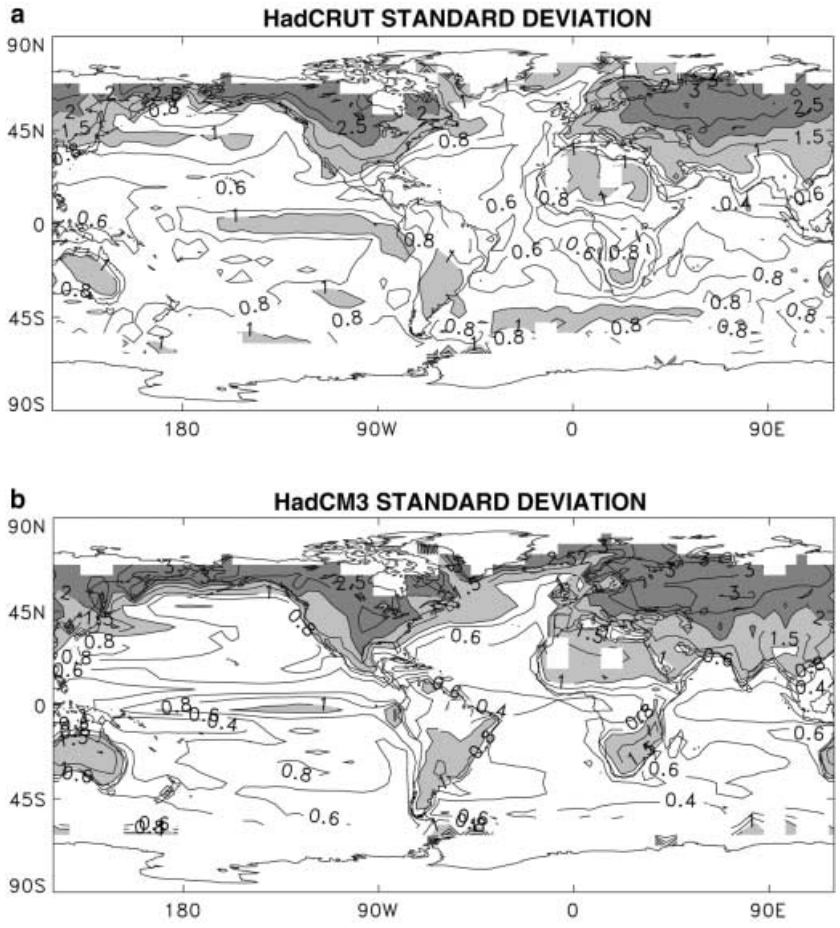

Fig. 3a, b The standard deviation (K) of **a** HadCRUT temperatures and **b** HadCM3 surface temperatures (processed on the HadCRUT grid). A non-linear contour interval is used to emphasise the variability over the oceans. *Light shading* indicates standard deviations between 1 K and 2 K and *dark shading* indicates standard deviations greater than 2 K

errors in the spatial pattern of variability are due to errors in the magnitude of the a_0 coefficient, which can be interpreted as errors in the atmospheric forcing, or the a_1 coefficient, which can be interpreted as errors in the oceanic memory. Examination of the ratios of the model to observed coefficients shows that where the model overestimates the variance, it is the random component which is largely responsible, and where the model underestimates variability, it is the memory component which is largely responsible. Thus areas of the model where the variability is overestimated are likely to be due to errors in the atmosphere model variability (due, for example, to errors in the strength or location of the storm tracks) and areas of the model where the variability is underestimated are likely to be due to low ocean model "memory" (due, for example, to poor representation of the propagation of Rossby waves due to low resolution). Again however, this comparison is confounded by the removal of the climate change signal from the observations, as any residual climate change signal would be likely to lead to an overestimation of the true a_1 coefficient in the case of the observations. Hence we conclude that HadCM3 simulates at least the broad scale features of surface temperature variability.

We next examine the frequency characteristics of the variability to asses the spectral shape (or colour) of the variability and to see if there are any periodic modes of surface temperature. Maps of model grid-box spectra power averaged into different frequency bands (figures not shown) generally show the same features as the maps of standard deviation Fig. 3b, i.e. less variability over the ocean in comparison with the land, with the maximum variability in northern high latitudes. We illustrate the dependence of spectral power on frequency by plotting the $\log_{10}$ of the ratio of the 20–50 year variability to the 2–8 year variability (Fig. 4). Where the ratio is large the spectral signature is one of a red noise process with more power at lower frequencies than at higher frequencies. Where the ratio is around unity the spectral signal is one of a white noise process with similar power at all frequencies. Where the ratio is small the spectral signature is one of a blue noise process or is indicative of a spectral peak which may be statistically distinct from the null hypothesis of an AR(1) process.

The demarcation of the land and ocean regions is clearly evident in Fig. 4. The ocean areas in HadCM3 are consistent with a red noise process and the land areas are consistent with a white noise process. The only notable exception is in the tropical Pacific which has a maximum at interannual time scales associated with the models' representation of ENSO (see Sect. 6). Plots of the ratios of the other frequency bands show similar features, land areas have similar power at all frequen-

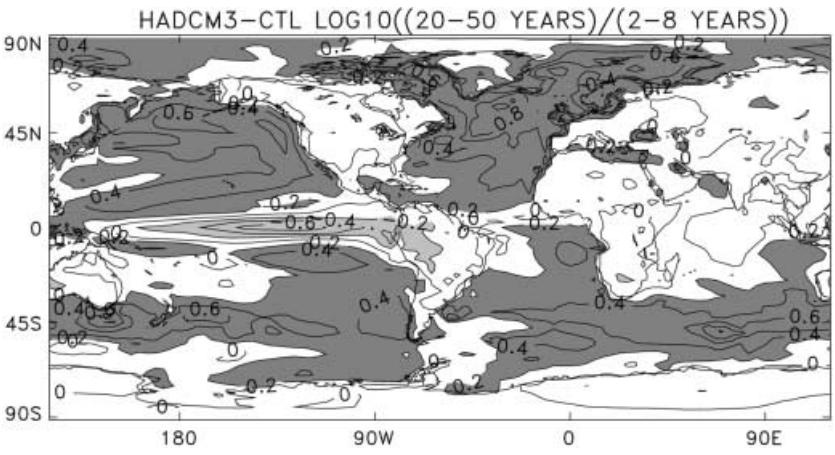

Fig. 4 The logarithm of the ratio of 20–50 year temperature variability to 2–8 year temperature variability (contour interval 0.2). *Dark regions* indicate log ratios greater than 0.2 and *light shading* indicates log ratios less than −0.2

cies, ocean areas have greater power at low frequencies and the tropical Pacific region is the only region in the model that shows enhanced variability at shorter time scales relative to longer time scales.

Because of this fundamental difference between the land and ocean we next analyse the variability of the two components separately.

5 Variability of land surface temperatures

We have seen that surface temperatures over land in HadCM3 are consistent with a white noise process and that the model overestimates this land variability. In HadCM2 Tett et al. (1997) found the variability over the Western European region and over the continental United States was overestimated in comparison with the observations. They speculated that the most likely explanation for this overestimation was excessive dryness of the soil in the model (except during the winter over the Western European region) so that too much of the variation in the net radiation entering the ground resulted in changes in ground temperature rather than changes in the evaporation of soil moisture. In this section we compare the variability of HadCM2 and HadCM3 in these regions to see if modifications to the model, in particular the introduction of the new surface scheme of Cox et al. (1999), have made any beneficial impact on this problem.

The ratio of the standard deviation of 1.5 m summer land temperature of HadCM3 to HadCM2 is shown in Fig. 5 for the Western European and continental United States regions. In both regions the interiors of the continents have less variability in HadCM3 than in HadCM2 (light shading). Cox et al. (1999) show that inclusion of soil water phase changes and the increased root depths in MOSES tends to alleviate the problems of excessive drying in summer and cooling in winter, and this seems to have a small, but beneficial, impact on the variability of HadCM3 compared to HadCM2. However, there remains a tendency for HadCM3 to overestimate the land variability. The global mean standard deviation of the observed land variability is 1.38 K, while it is 1.84 K and 1.81 K for HadCM2 and HadCM3 respectively. Examination of global maps of land variability (not shown) reveal that while there is some improvement in the northern mid-latitude regions and in the tropics, there are other land areas, notably the subtopical belts, in which the overestimation of variability in HadCM3 is increased in comparison with HadCM2.

6 The El-Niño Southern Oscillation

ENSO is the largest mode of interannual variability of global climate system. Coupling between the ocean and atmosphere is crucial in the formation of ENSO (see e.g. Neelin et al. 1998) so a successful simulation of the phenomena is an important test of a coupled model. As was seen in Fig. 3a there is an observed region of enhanced temperature variability to the east of the dateline in the tropical Pacific corresponding to the oceanic component of ENSO. In HadCM3 (Fig. 3b), there is also a region of enhanced temperature variability of a similar magnitude. However, this variability extends west of the dateline into the West Pacific warm pool which is a rather quiescent region in the observations. The extension of the model ENSO variability into the warm pool is a likely consequence of errors in the mean climate of the model in this region which has a cold tongue, and associated shallow thermocline, which extends too far into the west. The ocean surface temperature in this region can thus respond more rapidly to changes in surface heat flux and wind stress anomalies in comparison with the warm deep layer which exists in reality. In the observations there is also a region of variability to the south of the equator along the Pacific coast of South America. This is well simulated by the model and is a consequence of the reasonably high oceanic resolution which can resolve equatorial Kelvin waves which turn into coastally trapped Kelvin waves on reaching the coast.

ENSO has a well-known quasi-periodic nature corresponding to a peak in the power spectra around the 2–8 year frequency band. As was seen in Fig. 4 there is

enhancement of model variability at 2–8 year time scales in comparison with the 20–50 year time scales indicating a peak in the spectra. We take the NINO3 anomalies, which are temperatures averaged in the region 150°W–90°W, 5°S–5°N, as a simple measure of ENSO from the model and the observations. For the observed NINO3 anomalies we also use the reconstruction of sea surface temperatures of Rayner et al. (1996) (the Global Ice and Sea Surface Temperature data version 3, GISST3) as this provides a longer, and more complete, record of ENSO. The standard deviations of NINO3 indices are shown in Table 2. The model standard deviation is larger than the GISST3 estimate, but smaller than the estimate from the HadCRUT data. Hence the amplitude of ENSO, as measured by the standard deviation of NINO3 anomalies, is within the range of uncertainty of the observations. Power spectra of the model and observed NINO3 index are shown in Fig. 6. The spectra of the observed GISST3 NINO3 anomaly has a broad peak at 2–8 year time scales which is distinguishable from an AR(1) process at the 95% level. The broadness of the peak indicates an cycle of rather irregular period. The model also simulates a broad spectral peak, which is again significantly different from AR(1) but at a slightly higher frequency. Hence HadCM3 captures the periodic nature of ENSO with about the right frequency. However the model does simulate too much variability on interannual time scales and too little variability on time scales of less than a year, a feature which is masked by analysing the amplitude of ENSO by simply looking at the standard deviation of NINO3 anomalies.

In comparison with HadCM2 (see Collins 2000), HadCM3 has a similar representation of the amplitude, time scale and phase locking to the annual cycle (figures not shown). One of the main problems with HadCM2 was that SST anomalies associated with the models ENSO in the Tropical Pacific produced tropic wide atmospheric response which was much stronger and more

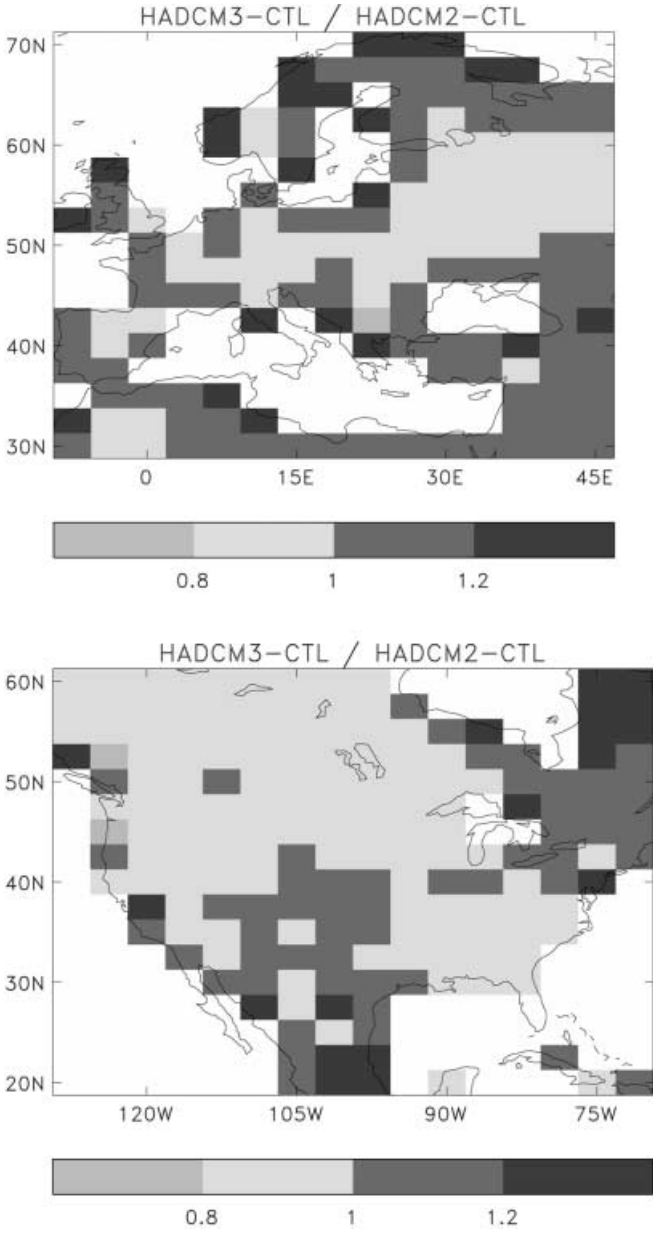

Fig. 5 The ratio of the standard deviation of 1.5 m temperature of HadCM3 to that of HadCM2 over the Western European region and over the continental United States region. *Lighter shading* shows where HadCM3 is less variable than HadCM2 and *darker shading* shows where HadCM3 is more variable than HadCM2

Table 2 The standard deviations of NINO3 temperature anomalies from HadCM3 and from two observational data sets, HadCRUT (Parker et al. 1995) and GISST3 (Rayner et al. 1996)

Data	Standard deviation of NINO3 anomalies
GISST3	0.77
HadCRUT	0.94
HadCM3	0.89
HadCM3 on HadCRUT grid	0.87

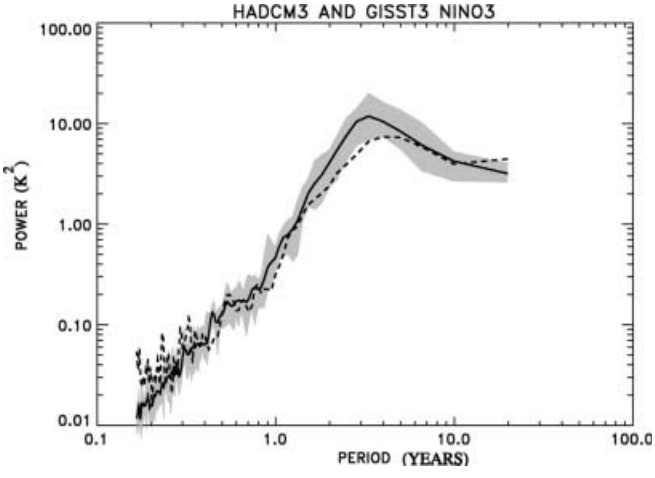

Fig. 6 Power spectra of the NINO3 anomalies from HadCM3 (*solid line*) and from the GISST3 observations (*dashed line*). The *shading* indicates the maximum and minimum power computed from ten 100 year sections of the model control NINO3 anomaly. The *solid line* is the average of the power spectra of these ten sections

wide-spread than is seen in the observations (see Fig. 10 of Tett et al. 1997). This teleconnection pattern dominated the variability of the model in the tropics.

To display the teleconnection patterns we follow Tett et al. (1997) and perform a regression analysis between the annually averaged surface temperatures and the temperature anomalies averaged in the NINO3 region. The regression coefficients (β) and the amount of variance explained (R^2) are shown in Fig. 7 for the HadCRUT temperatures regressed onto the GISST3 NINO3 anomalies. There are large regression coefficients in the tropical regions over the Pacific, over the Indian Ocean, and to a lesser extent over the tropical Atlantic. There are also positive coefficients to the north and south of the tropical Pacific showing the main mid-latitude response to ENSO. The R^2 diagram gives some indication of the significance of these regression coefficients.

The pattern of regression coefficients for HadCM3 (Fig. 8) is broadly similar to the observations though there is too strong a response to ENSO in the West Pacific warm pool region. This is linked to the enhanced variability in this region in comparison with the observations (see discussion and Fig. 3). The R^2 diagram for HadCM3 is much more like the observations than that for HadCM2 (see Tett et al. 1997, Fig. 10) which was too extensive in its response. Hence the tropic wide mode of variability, which was a feature of HadCM2, is much reduced. The reason for this improvement in ENSO teleconnections is difficult to asses because of the many changes to the physical parametrization schemes between HadCM2 and HadCM3 (see Sect. 2). However, a comparison of the climate change response of the two models (Mitchell et al. 1998) shows that HadCM2 has an enhanced warming in the tropics in comparison with HadCM3 with a tropic-wide pattern which is much the same as that seen during a HadCM2 El-Niño event. The enhanced HadCM2 warming in comparison with HadCM3 is a consequence of subtle changes in the boundary layer scheme and in the specification of the critical relative humidity for cloud formation (Williams et al. 2000), and it appears that these changes lead to an improvement in the response of HadCM3 to ENSO SST anomalies.

7 Interdecadal variability of ocean surface temperatures

In this section we examine variability in three oceanic regions, the tropical Pacific, the North Pacific and the North Atlantic. We focus on decadal to interdecadal time scales because of the possibility that variations on

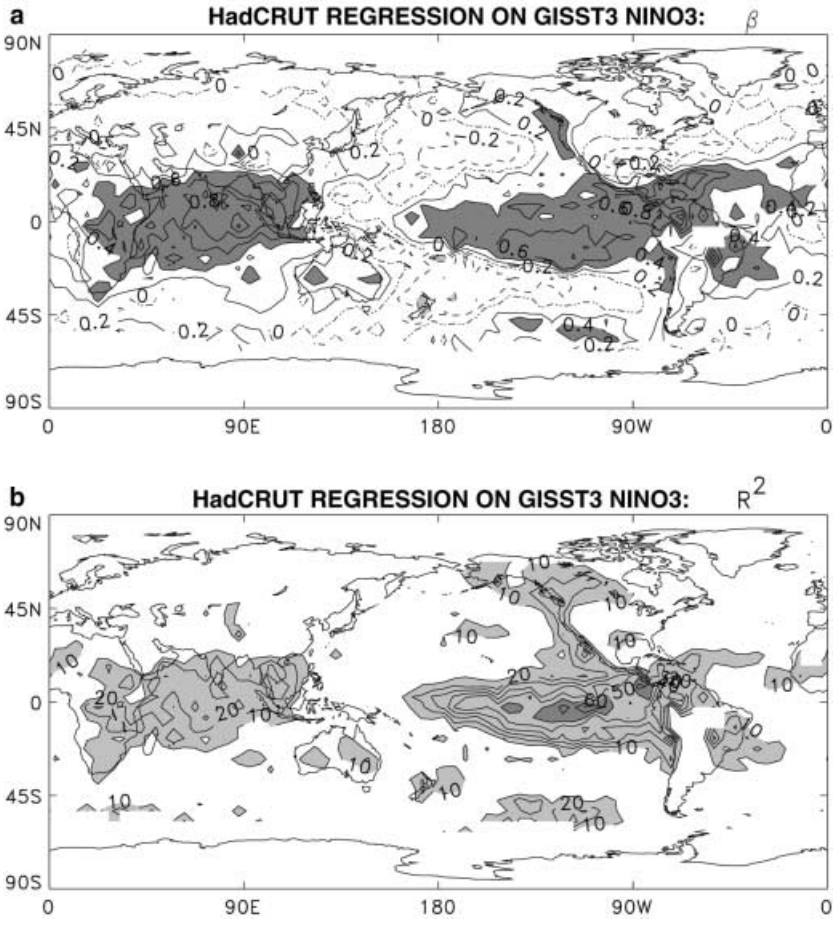

Fig. 7 **a** Regression coefficients (contour interval 0.2, positive contours *solid*, negative contours *dashed*, zero contour *dotted*, values greater that 0.4 *shaded*) and **b** the percentage of variance explained (contour interval 10%, values between 10% and 60% *light shading*, values between 60% and 100% *dark shading*) in a regression analysis of annually averaged HadCRUT temperatures and GISST3 NINO3 anomalies

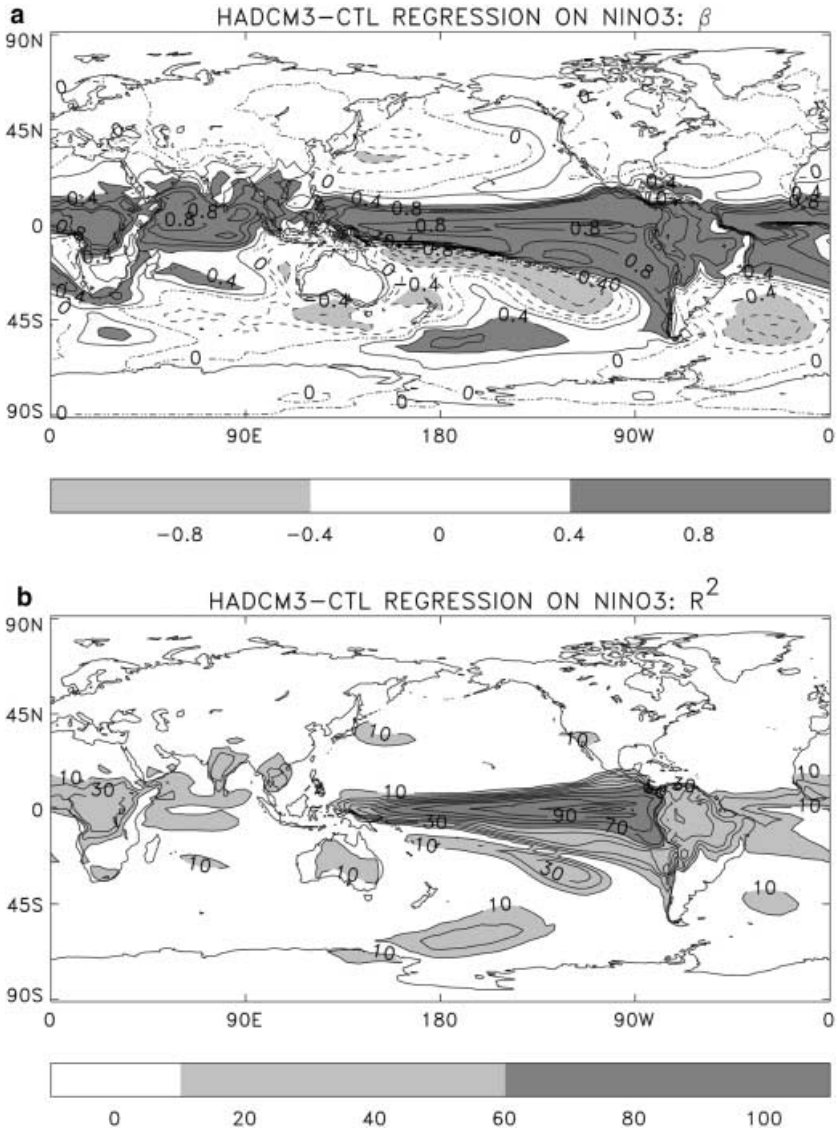

Fig. 8 **a** Regression coefficients (contour interval 0.2, positive contours *solid*, negative contours *dashed*, zero contour *dotted*, values greater that 0.4 *shaded*) and **b** the percentage of variance explained (contour interval 10%, values between 10% and 60% *light shading*, values between 60% and 100% *dark shading*) in a regression analysis of annually averaged HadCM3 temperatures and NINO3 anomalies

these time scales may be predictable and thus may lead to enhanced atmospheric predictability. Our principal aim is to validate the variability of HadCM3 but we also examine the 1000 year control run of the model to see if there are any periodic modes (aside from ENSO described in the previous section) which may lead to enhanced predictability.

To focus on decadal to interdecadal time scales we first average the SSTs from the model ocean $1.25° \times 1.25°$ longitude-latitude grid onto the $3.75° \times 2.5°$ atmosphere grid and then low-pass filter using a Butterworth filter of order 2 and half width of 8 years to remove time scales of less than 10 years. We then perform an EOF analysis of temperatures in the Tropical Pacific, the North Pacific and the North Atlantic. For comparison with observed decadal variability we compute the EOFs of the low-passed GISST3 (Rayner et al. 1996) data. We choose this over the HadCRUT data (Parker et al. 1995) because of its uniform space-time coverage when compared to the HadCRUT data which has variability coverage in both space and time. However, because of the nature of the in-filling algorithm used in the production of the GISST3 data, it is unlikely that the EOF analysis will produce different results. Folland et al. (1998) have performed an EOF analysis of the HadCRUT data and find very similar patterns to those found here. We average the SSTs onto the coarser atmosphere grid in order to make a comparison with the GISST3 observations which are more representative of variability on this scale because of the algorithm used in the in-filling and to insure that there are more time points than space points in EOF algorithm which leads to a more stable estimate of the covariance matrix. We scale each EOF by the standard deviation of its principal component (PC) which gives an estimate of the magnitude of SST variability associated with the EOF.

Studies of decadal and longer time scales variability in the tropical Pacific highlight a pattern of variability (sometimes called the Pacific Decadal Oscillation, PDO)

which is similar to the ENSO pattern, but with a broader meridional extent (see e.g. Zhang et al. 1997; Knutson and Manabe 1998). The first EOF of the low pass filtered GISST3 data in the region 120°E to 60°W and 30°S to 25°N (Fig. 9a) explains 49% of the low-passed variance and is similar to that found by Zhang et al. (1997) and Knutson and Manabe (1998) who use different analysis techniques and different observational data. The corresponding first EOF of the 1000 years of the HadCM3 control (Fig. 9b) explains 43% of the variance. The familiar triangular shape pointed out by Knutson and Manabe (1998) is evident in the model EOF, but there are considerable differences in the detail. The model has more positive loading in the West Pacific on the equator compared to the observed EOF. A similar feature was seen in the models interannual ENSO variability and is likely to be due to errors in the mean climate which has a cold tongue which extends too far into the west. There is also an anomalous maxima to the north off the equator in the model EOF and the model East Pacific maxima is displaced to the south in comparison with the observed East Pacific maxima. Nevertheless it appears that the model simulates a form of PDO.

The 1000-year long simulation of HadCM3 allows us to assess if there is any cyclic or periodic behaviour of the PDO. We first form a time series by projecting the model EOF onto the unfiltered model SST fields and then compute the power spectra and compare this with an AR(1) process. This power spectra has a significant peak at interannual time scales, corresponding to the projection of ENSO onto the PDO, but no significant peaks at any other time scales. Thus the models PDO is consistent with an AR(1) process at the 95% level and is not an interdecadal periodic phenomena.

We next examine decadal variability in the North Pacific region defined as 120°E to 120°W and 25°N to 60°N. Latif and Barnett (1996) found a pattern of interdecadal variability in the North Pacific both in observations, and in a coupled model, which could be described by the two highest ranking EOFs. These are shown for the GISST3 data in Figs. 11a and 10a. In combination these EOFs explain 73% of the low-passed variance and are similar to those found by Latif and Barnett (1996). The leading two model EOFs (Figs. 11b and 10b) have their order swapped in comparison to the observations, a feature which was also found by Latif and Barnett (1996) with their model patterns. The model EOFs explain 55% of the low-pass filtered variance, somewhat less than the observations. The model EOF1 and observed EOF2 (Fig. 10) are very similar in pattern but with the maximum loading in the model displaced slightly to the west in comparison with the observations. The model EOF2 and observed EOF1 (Fig. 11) are less similar though there is a general pattern of positive loading throughout the North Pacific. The model EOF2 has only small negative loading in the sub-tropical western region.

Latif and Barnett (1996) described their model decadal variability in terms of a cyclic mode with a period of approximately 20 years. Power spectra (not shown) of the time coefficients formed by projecting the leading two North Pacific EOFs from HadCM3 onto the unfiltered model fields show no peaks in the spectra that are significantly different from an AR(1) process at the 95% level. This indicates that, although HadCM3 has a similar spatial structure of variability in comparison with other models and with the observations, there is no indication of a periodic or cyclic mode of variability.

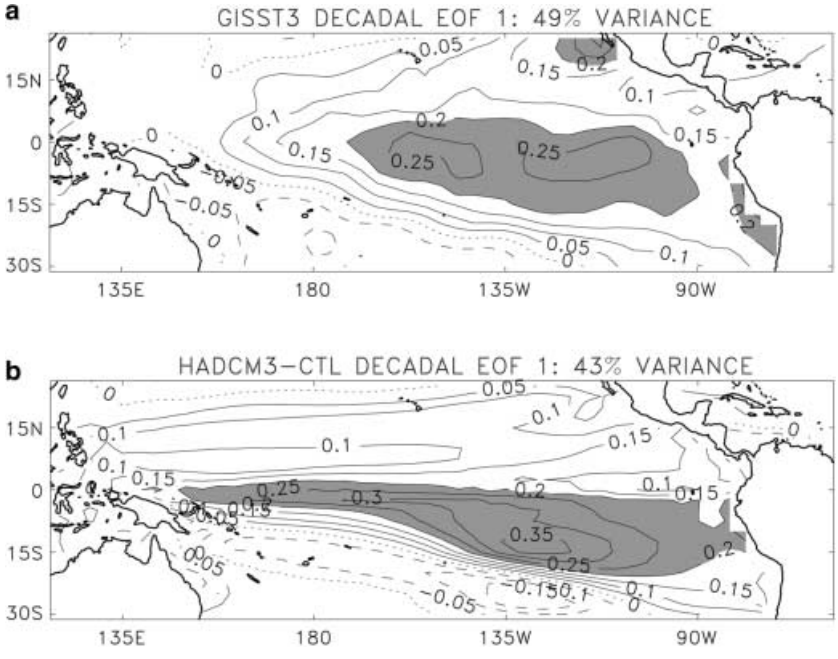

Fig. 9 **a** The first EOF of the low-pass filtered GISST3 data in the Tropical Pacific region and **b** the first EOF of the low-pass filtered HadCM3 control run. The contour interval is 0.05 and *solid lines* indicate positive values, *dashed lines* indicate negative values and the zero contour is shown as the *dotted lines*. Values greater than 0.2 are *shaded*

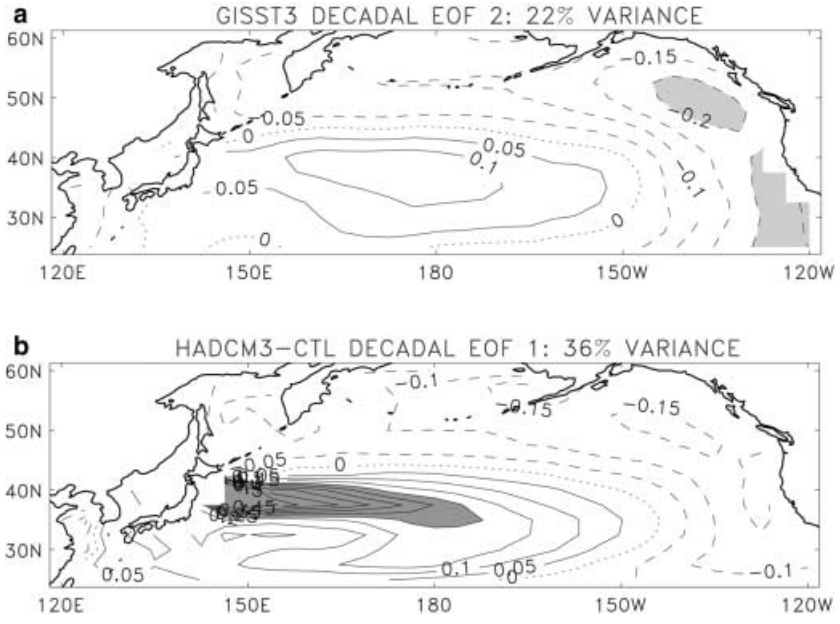

Fig. 10 **a** The second EOF of the low-pass filtered GISST3 data in the North Pacific region and **b** the first EOF of the low-pass filtered HadCM3 control run. The contour interval is 0.05 and *solid lines* indicate positive values, *dashed lines* indicate negative values and the zero contour is shown as the *dotted lines*. *Dark shading* is used for values greater than 0.2 and *light shading* is used for values less than −0.2

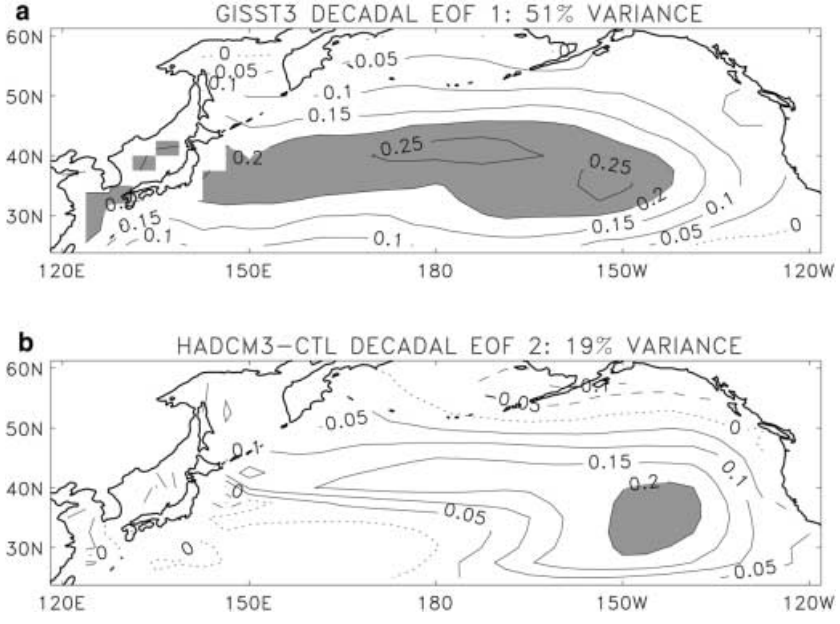

Fig. 11 **a** The first EOF of the low-pass filtered GISST3 data in the North Pacific region and **b** the second EOF of the low-pass filtered HadCM3 control run. The contour interval is 0.05 and *solid lines* indicate positive values, *dashed lines* indicate negative values and the zero contour is shown as the *dotted lines*. Values greater than 0.2 are *shaded*

Latif and Barnett (1996) suggested that the decadal North Pacific mode of variability was decoupled from the tropical regions in both the model and in the observations. Other authors (e.g. Trenberth and Hurrell 1994) have suggested though that there may be some link between North Pacific and tropical Pacific decadal variability. To assess this we simply calculate the correlation coefficients of the time coefficients of the leading EOFs from the tropical Pacific (Fig. 9) and from the North Pacific (Figs. 11 and 10). Large correlations are found between the tropical Pacific PC1 and the North Pacific PC2 (0.75) for the observations and between the corresponding tropical Pacific PC1 and the North Pacific PC1 (0.62) for the model (recall the order of the North Pacific EOFs are swapped with respect to the observed EOFs). Instantaneous correlations between other principal components are much less and there are no large correlations at lags of up to 20 years that would indicate propagating structures. This suggests that, for the model and for the observations, there is one pattern of variability, with a spatial structure like Fig. 11 that exists independently in the North Pacific and another pattern, with a spatial structure like Fig. 10 which is coupled to the tropical Pacific. A EOF analysis of the Pacific Ocean basin from 30°S to 60°N (not shown) confirms this, giving a leading EOF pattern that is very similar to the combined tropical Pacific EOF1 (Fig. 9a) and North Pacific EOF2 (Fig. 10a) for the

observations and the combined tropical Pacific EOF1 (Fig. 9b) and North Pacific EOF1 (Fig. 10b) for the model.

Grötzner et al. (1998) use a canonical correlation analysis (CCA, see Barnett and Preisendorfer 1987) of winter season SST and surface pressure to identify patterns of possible coupled variability in the North Atlantic region in observations and in a coupled model. CCA finds the spatial patterns from two data sets such that the correlations between the time coefficients of the patterns is maximised. CCA patterns do not necessarily correspond to EOF patterns so rather than using the EOF method to investigate low-frequency variability of the North Atlantic, we adopt the CCA approach. The leading CCA patterns of unfiltered December–February mean SSTs and mean sea level pressure (MSLP) in the North Atlantic region (90°W–0°W, 10°N–80°N) are shown in Fig. 12. The SST pattern explains 11% of the SST variance and the MSLP pattern explains 40% of the MSLP variance. The correlation between the CCA mode time coefficients is 0.64 which is similar to that found by Grötzner et al. (1998). The SST pattern is a "tripole" and the MSLP pattern is the models representation of the North Atlantic Oscillation (NAO) which we examine in detail in the next section. Both patterns are very similar to those found by Grötzner et al. (1998) for the observations (see their Fig. 3) although we find no preferred time scale for the patterns, the time coefficients being consistent with an AR(1) process i.e. there are no significant spectra peaks.

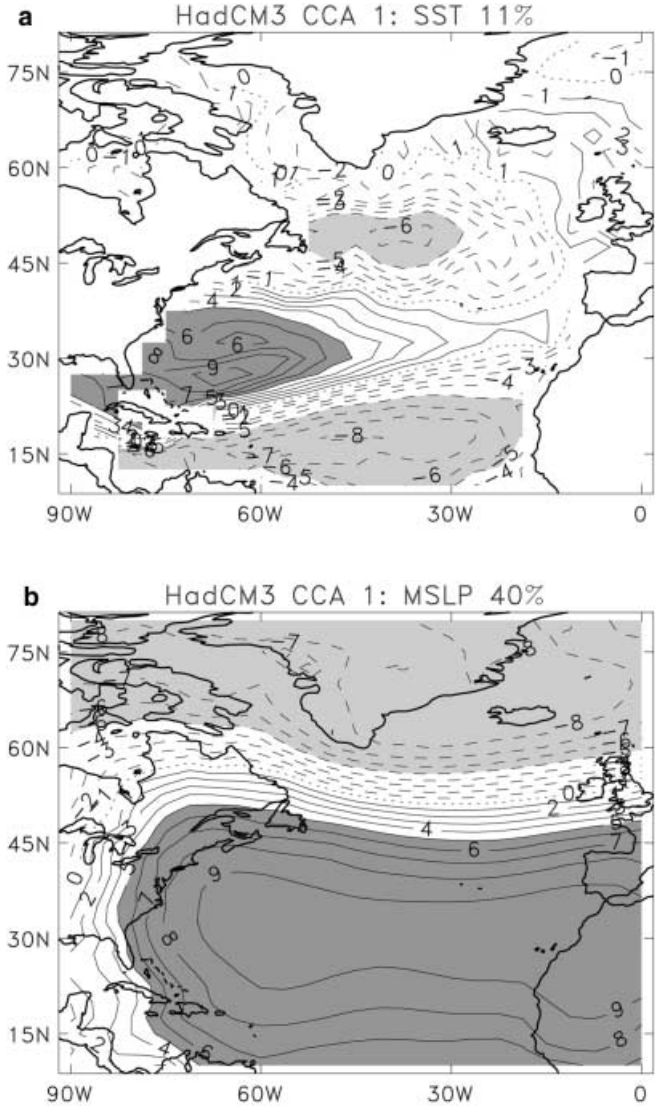

Fig. 12 **a** The leading CCA pattern HadCM3 SSTs and **b** MSLP. The contour interval is 1 and *solid lines* indicate positive values, *dashed lines* indicate negative values and the zero contour is shown as the *dotted lines*. *Dark shading* is used for values greater than 5 and *light shading* is used for values less than 5

The existence, or not, of a coupled "mode" of variability in the North Atlantic has been a subject of much recent debate (see e.g. Rodwell et al. 1999; Bretherton and Battisti 2000). The CCA analysis suggests that there may be some ocean-atmosphere interaction in the model but we cannot be definite about the existence of a coupled mode (in the sense of ENSO being a coupled mode) without further analysis and possible experiments with the model (e.g. decoupling the components). We simply state the fact here that the model reproduces the main CCA patterns of the observations but there is no preferred time scale for those patterns. In the next section we examine the atmospheric component of the CCA pattern, the NAO.

8 The North Atlantic Oscillation

The dominant atmospheric modes of interannual to decadal variability in the Northern Hemisphere winter are the North Atlantic Oscillation (NAO, e.g. Hurrell 1995) and the related hemispheric wide Arctic Oscillation (AO, e.g. Thompson and Wallace 1998). Signatures of the NAO and AO can be found in the leading EOFs of, for example, surface pressure and 500 hPa geopotential height as well as in various indices computed from station data, and the two modes seem to be intimately linked. The pattern of variability associated with both involves an alternation of relatively low and high pressure between the mid-latitudes and polar regions. Associated with these changes in pressure there are changes in storminess, precipitation and surface temperature (Hurrel 1995). It has been suggested that the AO and NAO may be coupled to SSTs and that they may predictable on seasonal to decadal time scales (e.g. Rodwell et al. 1999). In this section we examine the variability of Northern Hemisphere winter in HadCM3 to see if it simulates AO and NAO behaviour and to see if this variability is in anyway periodic (i.e. statistically distinguishable from AR1) suggesting possible predictability. Also the observed NAO index shows a marked increase from the mid 1960s to the mid 1990s and it has been suggested (e.g. Shindell et al. 1999) that this may be a signature of climate change, hence we look at trends in

the observed and modelled NAO to see if this increase in NAO index can be explained by natural internal climate variability.

We first examine the variability of 500 hPa geopotential height which characterises the large-scale tropospheric circulation. The standard deviation of Northern Hemisphere (20°N–90°N) DJF 500 hPa height simulated by HadCM3 is shown in Fig. 13a. For comparison the standard deviation of the DJF 500 hPa height from NCEP reanalysis (Kalnay et al. 1996) for 1958–1996 is shown in Fig. 13b. Both show maxima over the North Atlantic and North Pacific coincident with the two major Northern Hemisphere storm-tracks. Overall the model captures the observed pattern of variability well but the amplitude is slightly underestimated.

We next examine the leading EOFs of the model and NCEP reanalysed 500 hPa geopotential height. We scale the EOFs by the standard deviation of their corresponding principal components to give some idea of the typical magnitude of anomaly associated with the EOF. The model and reanalysis EOFs (Fig. 13c, d respectively) both show an anti-correlation between the Atlantic and the Arctic regions, although the model does not show the positive loadings in the Atlantic extending over western Europe found in the reanalysis. This model error is likely to be a consequence of the underestimation of blocking events at the end of the Atlantic storm tracks (Pope et al. 2000). A more striking difference is the enhanced correlation seen over the North Pacific in the simulation in comparison to the reanalysis. The leading mode of DJF 500 hPa height has a more AO than NAO pattern with greater teleconnection between the Pacific and the Atlantic. A similar pattern was also found by Osborn et al. (1999) in HadCM2. As was the case for the MSLP component of the CCA mode in the North Atlantic (see previous section), the PC of leading EOF has a power spectra which is consistent, at the 95% level, with an AR(1) process so there is no preferred time scale or periodic behaviour in the model AO.

We also examine correlations of the PC of the leading EOFs with DJF near-surface temperatures (Fig. 14). In both the model and the reanalysis there are no significant correlations with the tropical Pacific region, suggesting that ENSO plays little role in the leading mode

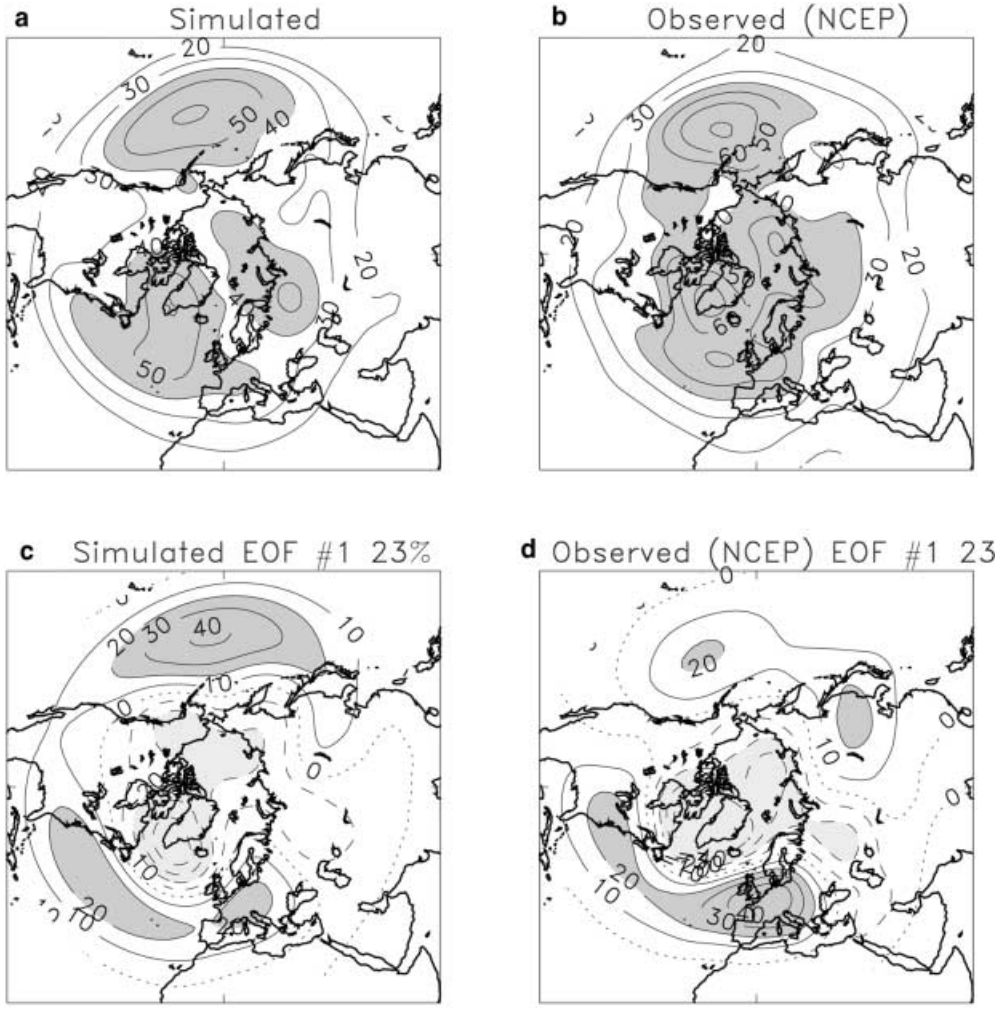

Fig. 13 **a** Simulated standard deviation of DJF 500 hPa height anomalies (contour interval is 10 m with *shading* above 40 m); **b** As **a** but for the NCEP reanalysis data. **c** Leading EOF of simulated 500 hPa height (contour interval is 10 m with *dark shading* above 20 m and *light shading* below −20 m). **d** As **c** but for NCEP reanalysis

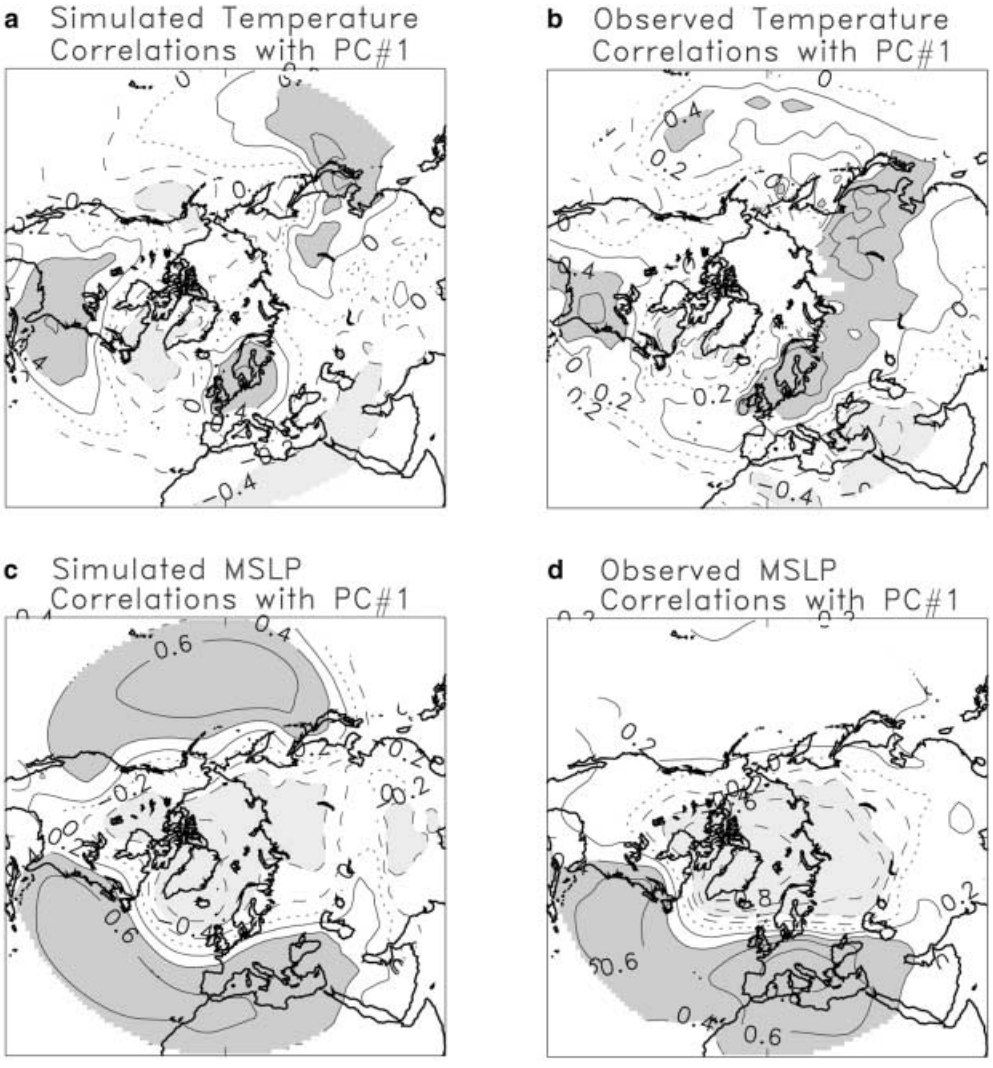

Fig. 14 **a** Simulated correlation of DJF 1.5 m temperature with leading simulated PC of 500 hPa height. **b** As **a** but for observed data. **c** Simulated correlation of DJF mean sea level pressure leading simulated PC of 500 hPa height. **d** As **c** but for observed data. The contour interval is 0.2 with *dark shading* for values greater than 0.4 and *light shading* for values less than −0.4

of both simulated and observed 500 hPa height (a similar result was found for HadCM2, see Osborn et al. 1999). The correlation structure does appear to be similar in both the model and the observations with positive correlations in the southwest USA, Western Europe and East Asia with negative correlations apparent over the Labrador Sea and Africa/Arabia. However the model does not show a region of positive correlations throughout northern Eurasia which is clearly evident in the observations (correlating the time series found by projecting the model fields onto the observed EOF with the model surface temperature gives slightly higher correlations in this region, but the model still falls short of the observed correlations). Correlations of the PC of the leading EOFs with DJF mean sea level pressure (MSLP, Fig. 14c, d) show similar patterns to the leading EOFs (Fig. 13c, d), with negative correlations over the polar regions and positive correlations to the south. As was the case for the EOFs, the model has a positive correlation in the Pacific regions which is absent in the observations indicating an enhanced Atlantic-Pacific teleconnection in the model.

A index of the NAO was defined by Jones et al. (1997a) in terms of the absolute pressure difference between stations in Gibraltar and southwest Iceland for the December–March season. This index is highly correlated (coefficients around 0.9) with the leading principal component of the NAO mode and with other station based indices (Osborn et al. 1999) and, because of the long observational records, allows the reconstruction of the NAO from the present-day back to 1824. The index is shown Fig. 15 (right hand curve). The dotted line is the winter index and the thick black line is the index filtered with the 10 year low pass filter described in Sect. 7. The rapid rise in the NAO index from the mid 1960s to the mid 1990s which, in HadCM2, could not be explained by natural variability alone (Osborn et al. 1999), is of considerable interest and it has been suggested (e.g. Shindell et al. 1999) that it may be a signal of climate change.

The standard deviations of the HadCM3 and observed NAO series compare well being 6.1 and 6.4 hPa respectively for the annual indices and 2.5 and 2.6 hPa for the low-pass filtered data. The power spectra of the

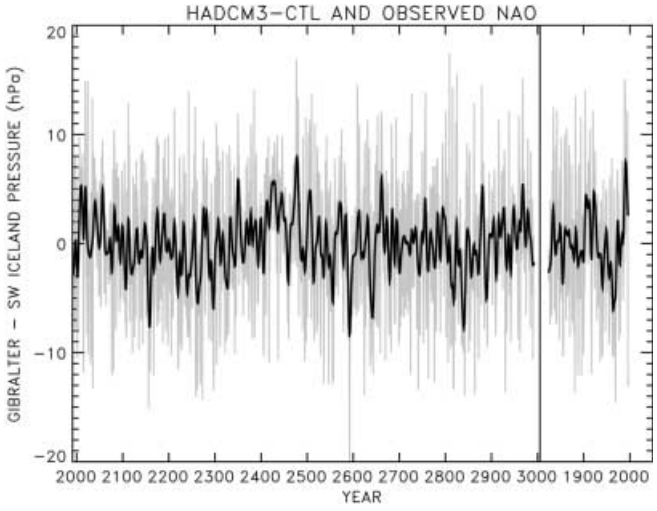

Fig. 15 North Atlantic Oscillation index (Gibraltar minus southwest Iceland pressure anomaly for the December–March season) from the HadCM3 control (*left curve*) and from the observations of Jones et al. (1997a) (*right curve*). The winter indices are shown as the *grey lines* and the 10 year low-pass filtered indices are shown as the *thick black line*

model and the observations (not shown) are both consistent with a white noise process and have no significant spectral peaks. The power spectra of the observed NAO is also consistent with the model power spectra using the test described in Sect. 3. This suggests that the observed NAO is consistent with the model and that the recent trend in the NAO index can be explained by natural variability alone. Indeed it seems, from Fig. 15, that the recent large values of NAO index are within the range of variability of the model. However, the key aspect is not the absolute magnitude of the NAO, but the rate of change of NAO with time. An analysis of the trends in the observed and modelled NAO indices (using the method described in Sect. 3) shows that the recent observed rapid rise in NAO is highly unusual in comparison with the 1000 years of HadCM3 NAO index. Observed trends of 10 to 70 years in length occur very infrequently in the control (less than 1% chance), with trends of 30, 50 and 60 years having only one occurrence in the 1000 year control and trends of 40 years have no occurrence (Table 3). Hence we can conclude that, while the absolute value of the observed NAO is within the range of variability of HadCM3, the rate of change of NAO in recent years is inconsistent with the internal variability in the model.

9 Variability of zonal mean atmospheric temperatures

We next turn our attention to variations in zonal mean temperatures in the atmosphere. Various studies have proposed that observed changes in the temperature of the zonal mean atmosphere indicate a strong likelihood of an anthropogenic influence on climate (Santer et al. 1995; Tett et al. 1996; Allen and Tett 1999). Such claims rely heaviliy on a good model simulation of the internal climate variability of the free atmosphere and are furthermore confounded by the sparse observational radiosonde network. Here we compare the simulated zonal mean atmospheric variability from HadCM3 with the observed variability from the HadRT2.1 gridded radiosonde dataset, an updated version of that described in Parker et al. (1997). The radiosonde profiles are binned into monthly means on a 10° longitude by 5° latitude grid with eight levels in the vertical from 850 hPa to 50 hPa and are for the 40-year period 1958 to 1997.

Table 3 10–80 year trends in the observed NAO and corresponding occurrences of trends greater than the observed trends computed from the 1000 year HadCM3 control run NAO. Occurrences are quoted in both actual and relative terms (see text for more details)

Years (inclusive)	Length of trend (years)	Trend (hPa/year)	Occurrence	Relative occurrence
1989–98	10	1.3	58	0.06
1979–98	20	0.7	4	0.004
1969–98	30	0.4	1	0.001
1959–98	40	0.3	0	0
1949–98	50	0.2	1	0.001
1939–98	60	0.1	1	0.001
1929–98	70	0.07	12	0.01
1919–98	80	0.04	219	0.2

Annual data was computed from the gridded (latitude, longitude, pressure) radiosonde temperatures by averaging the monthly data in each grid box with the requirement that there be eight months with data in the year. Grid boxes without eight months of data were flagged as "missing". Non-overlapping 40-year segments of annual mean data were taken from the control simulation, tri-linearly (pressure, longitude and latitude) interpolated to the observational grid and where the observations were missing the simulated values were also set to missing. The equivalent 20-year average was removed from each 40-year segment to be consistent with observations which have the 1971–1990 average removed. Both simulated and observed data were then zonally averaged using the criterion that if there was at least one point in each latitude band then a zonal value was computed. Each 40-year segment then had the 40-year average removed with the further requirement that at each longitude-pressure point there be 20 values otherwise the resultant value was again flagged as missing. Standard deviations from the observations and the model data were then computed from these processed datasets. Following Gillett et al. (2000) the maximum and minimum standard deviation from the 19 40-year segments of simulated data were retained. These values correspond to approximately the 5% levels for the control to be consistent with the observations i.e. at any point, we can say, with 95% confidence, that the model has too much variability if the minimum standard deviation is greater than that observed, and too little variability if the maximum is less than that observed.

For comparison the standard deviation of the zonal-mean temperatures were also computed from the full model grid with no interpolation or masking of missing data.

The unprocessed model fields (Fig. 16a) show a maximum in variability in the polar stratosphere regions and at the tropical tropopause, and minimas in the mid-latitude troposphere in both hemispheres. Masking the simulated data by the observational mask has some effect on the magnitude of the variability although the overall pattern remains the same. The model has too little variability in the stratosphere and throughout most of the extra-tropical atmosphere. This is confirmed by computing the ratio of the model to observed standard deviations and shading the regions of statistical significance as defined already (Fig. 16d). There are a few regions where the model has too much variability (25°N and 25°S at approximately 300 hPa) but the model has too little variability in the entire stratosphere (standard deviations are 40–80% of those observed) and in the extra-tropical troposphere (standard deviations are 60–80% of those observed).

The poor simulation of stratospheric variability in HadCM3 is likely to be due to the lack of quasi-biennial oscillation (e.g. Andrews et al. 1987) in the model, and to the poor simulation of major sudden warmings (e.g. Swinbank et al. 1998). The lack of mid-latitude tropospheric variability is of more concern, although it is consistent with the findings of the previous section that the 500 hPa height variability in the Northern Hemisphere region is somewhat underestimated in HadCM3 (see Fig. 13a, b). Other candiates such as anthropogenic or natural climate variations not simulated in the control climate are also possible.

Earlier work (Gillett et al. 2000) has shown that the control simulation of HadCM2 seemed to show too much coherence between changes in the surface and in the free atmosphere. Several authors (e.g. Hurrell and Trenberth 1997) have claimed that there are differences in the observed trends of surface temperature and of temperature in the free troposphere. We investigate this by computing the correlation between the zonal mean temperatures and the global averaged surface temperature for both the model and the observations, taking into account the missing data mask (Fig. 17). Significance is assessed by comparing the maximum and minimum value from the non-overlapping 40-year segments of the control with the observations.

The observations (Fig. 17a) show large positive correlations with the surface throughout most of the tro-

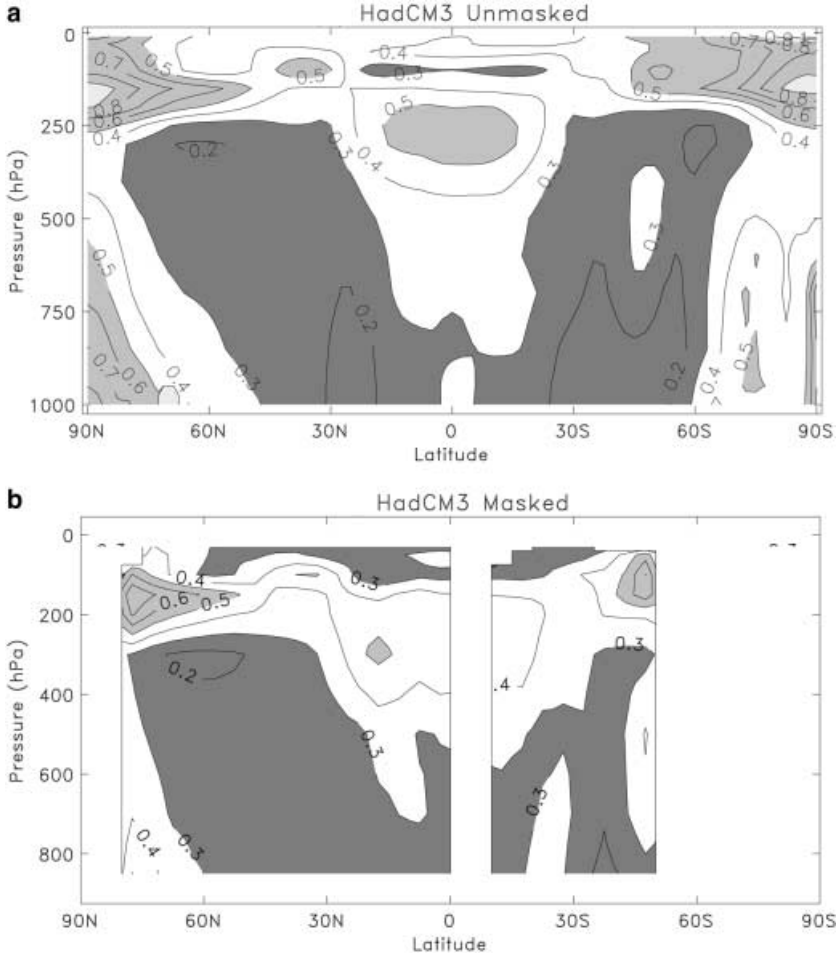

Fig. 16 **a** Standard deviation of annual zonal mean temperature from the control simulation of HadCM3. A contour interval of 0.1 K is used with *dark shading* for values less than 0.3 K and *light shading* for values greater than 0.5 K. **b** As **a** but for masked model. **c** As **b** but for observations (HadRT2.1). **d** Ratio between the model standard deviation and the observed standard deviation. A contour interval of 0.2 is used. *Dark* (*light*) *shading* shows where the observed standard deviations are significantly smaller (larger) the model value

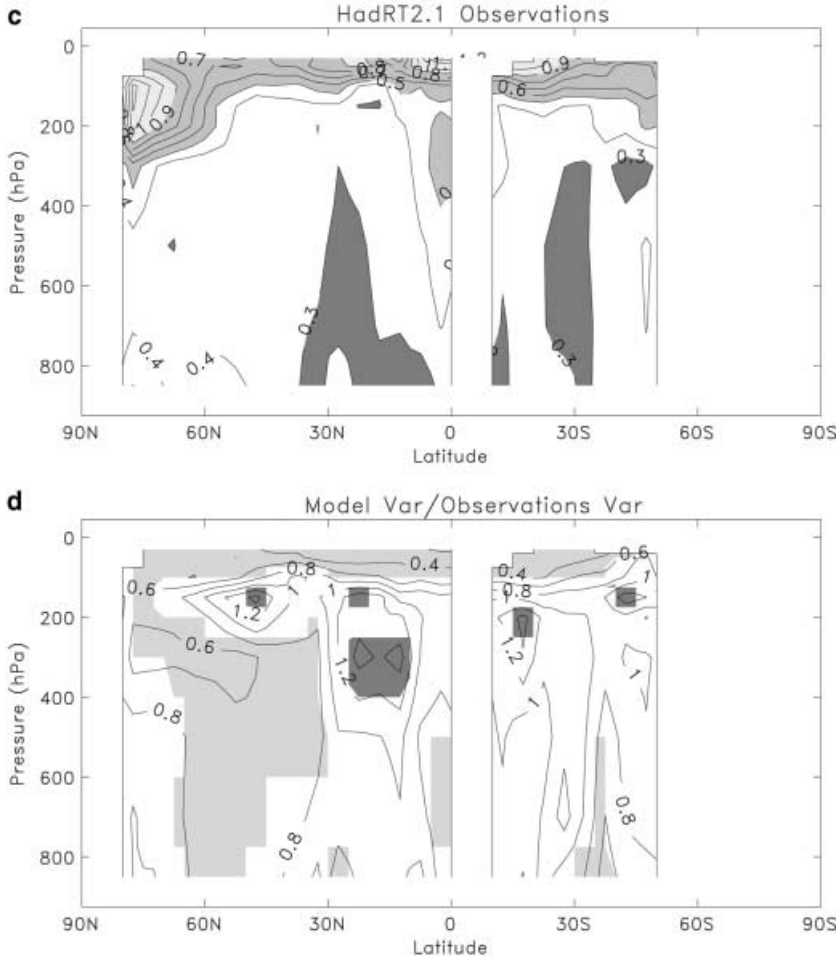

Fig. 16. Contd.

posphere and negative correlations in the stratosphere. The model (Fig. 17b) shows large positive correlations only in the equatorial troposphere with only a small region of large negative correlations in the equatorial stratosphere. There are large regions of significant differences between the model and observations: The model has correlations that are too large in the equatorial upper troposphere and throughout most of the stratosphere (where the observations show large negative correlations and the simulation does not). The difference between the model and observations in the Northern Hemisphere, where the observations show large positive correlations and the model does not, are also significant.

Part of the explanation for these differences between the model and observed correlation structures may be due to external forcings acting on the climate system such as changes on greenhouse gases, stratospheric ozone or explosive volcanic eruptions, though this does not rule out the possibility of model error.

10 Summary

We have examined the internal climate variability of a 1000 year long integration of the third Hadley Centre coupled model, HadCM3. Our aims were to validate the internal climate variability of the model, to examine the model for any periodic modes of variability, to assess the probability of recent observed trends in climate variables, and to compare the model with the previous version of the Hadley Centre model, HadCM2.

The standard deviation of global mean 1.5 m temperature from HadCM3 is in good agreement with the standard deviation of the detrended observed surface temperature record of Jones (1994) and Parker et al. (1995), and with the standard deviation of the global mean 1.5 m temperature of HadCM2. Also, the power spectrum of the global mean surface temperature of the model is consistent with the power spectrum of the observed temperature on interannual to decadal time scales, and the model has no periodic modes of global temperature. Linear trends computed from the observations over the most recent 20 years and longer, are highly unlikely in the context of the variability of HadCM3, indicating that the observed changes in mean climate are inconsistent with internally generated climate variability.

The spatial pattern of surface temperature variability in HadCM3 is qualitatively similar to that of the observations, with there being greater variability over the land than over the ocean areas and with the greatest

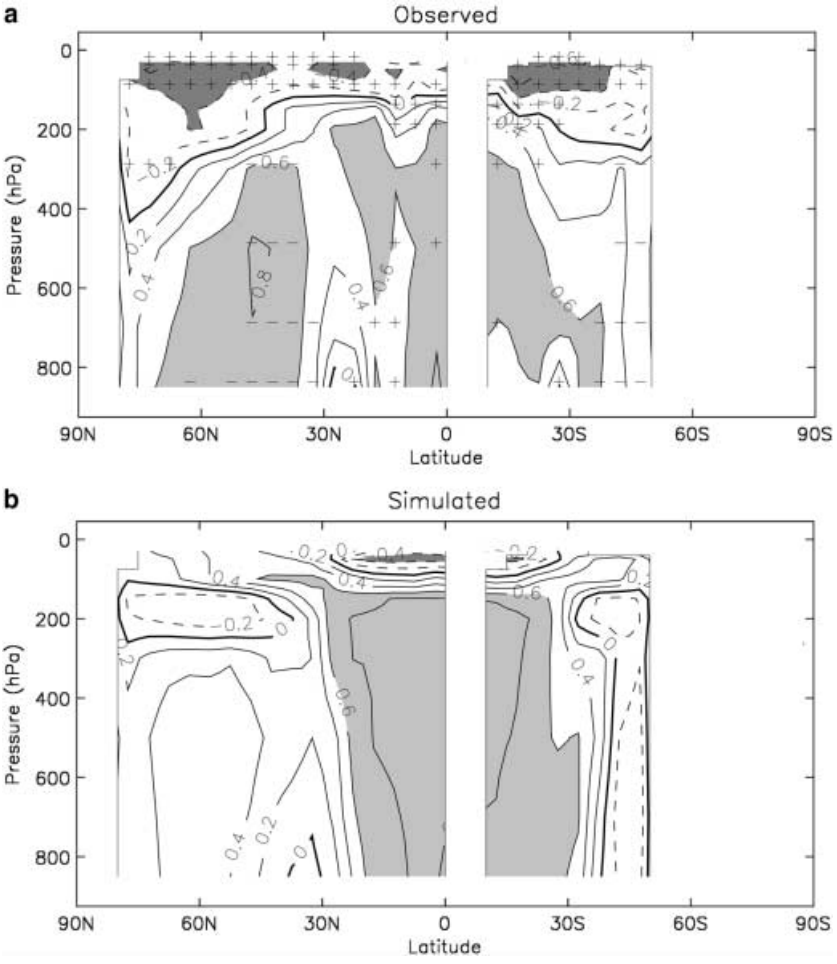

Fig. 17 **a** Observed correlations between zonal-mean annual-mean temperature (computed as in Fig. 16) and global-mean near-surface temperature for the 40 year period 1958–1997. A contour interval of 0.2 is used with negative contours shown *dashed* and the zero contour drawn *bold*. *Light shading* shows correlations greater than 0.6 and *dark shading* shows correlations less that −0.6. The + and − symbols show where control has significantly larger or smaller correlations than the observations (see text for details). **b** As **a** except for the simulated zonal mean temperatures (also computed as in Fig. 16)

variability over the Northern Hemisphere continents. However, there are some regional differences between the model and the observations and there is a tendency for the model to overestimate the variability over land. In comparison with HadCM2, the introduction of the MOSES surface scheme (Cox et al. 1999) alleviates the variability bias over land in Western Europe and North America to a certain extent, but HadCM3 still seems to overestimate the variability in these, and other regions. Decomposing the variability into spectral bands shows that the land areas generally have similar levels of variability on different time scales (white noise) and that ocean areas generally have more variability at longer time scales (red noise).

The only notable exception to this rule is in the tropical Pacific ocean of the model which has a spectral peak at interannual time scales associated with the ENSO. HadCM3 has a reasonable simulation of ENSO having an amplitude that is within the range of uncertainty of the observations, and a broad spectral peak centred on a period of 3–4 years which is significantly different from the spectra of an AR(1) process at the 95% level. There is too much ENSO related variability in the Western Pacific warm pool where the mean climate of the model is too cold and the thermocline too shallow. Hence the model can respond more rapidly to changes in surface fluxes and wind stress than in reality. The excessive teleconnection of ENSO to the rest of the globe, that was a feature of HadCM2, is much improved in HadCM3.

The interdecadal variability of the ocean surface temperatures was examined in three regions, the tropical Pacific, the North Pacific and the North Atlantic. The dominant pattern of interdecadal tropical Pacific variability resembles the observed PDO but with too much variability in the west associated with errors in the mean climate in the region. The model PDO shows no periodic behaviour in contrast to the model ENSO which has a quasi-regular oscillation with a period of 3–4 years. The two leading patterns of North Pacific interdecadal variability in the model are similar to those computed from the observations, but with some differences in the region of the Kuroshiro current. As was the case for the tropical Pacific, there was is no preferred time scale for the variability but there is some indication of a link between the tropical Pacific and the North Pacific variability in the model. Joint analysis of SST and MSLP in the North Atlantic region of the model shows a similar pattern to that seen in the observations indicating the possible importance of ocean-atmosphere

coupling in this region. Again though, there are no significant spectral peaks in the time coefficients of the patterns.

HadCM3 simulates AO/NAO variability in Northern Hemisphere winter. The leading EOF of 500 hPa height in the model was found to be in reasonable agreement with the observed leading EOF showing an anti-correlation between high latitudes and mid-latitudes. However, the model shows excessive teleconnection between the North Atlantic and the North Pacific (which was also the case for HadCM2). An index of the NAO computed from the model shows good agreement with the observations in terms of the magnitude of the standard deviation and the spectral characteristics. However, the observed upward trends in the NAO index are highly unusual when compared to the model NAO trends.

The variability of the zonal mean temperature of the atmosphere is generally underestimated in HadCM3, particularly in the stratosphere and in both the northern and southern mid-latitude troposphere. Also the correlations between the global surface temperature and the zonal mean atmosphere are poorly simulated by the model. Although some of these errors may be due to external forcing (e.g. stratospheric ozone depletion) not represented by in the control, this is an area where the model simulation could be improved. We note recent success in an atmosphere only version of HadCM3, with enhanced vertical resolution in the stratosphere and a parametrisation of non-orographic gravity waves, in simulating a QBO (Scaife et al. 2000). Hence it seems that corrections to the models simulation of zonal mean temperature variability are possible.

The removal of flux adjustment has been one of the major goals of climate modelling in recent times. We have shown that internal climate variability can be reasonably well simulated by a coupled model without flux adjustments, although we note that there are several improvements which can, and should, be made. If one assumes that the internal climate variability of HadCM3 can be used as a surrogate for the internal variability of the real climate system, then we may draw the following conclusions:

1. Observed trends in global annual mean surface temperature from all years prior to 1979 to 1998 cannot be adequately explained by internal climate variability and thus must be forced by external factors.
2. Observed trends in the NAO index over the past 20–70 years cannot be adequately explained by internal climate variability and thus must also be forced by external factors.
3. ENSO is the only periodic mode of climate variability, where periodic is defined as having a spectral peak which is statistically distinguishable from an AR(1) process.

Acknowledgements Thanks go to Tim Johns who ran the HadCM3 control integration and to John Mitchell and Tim Osborn who provided comments on an earlier version of the manuscript. The work was supported by UK Department of the Environment, Transport and the Regions (PECD/7/12/37) and by the Public Meteorological Service Research and Development Programme.

References

Allen MR, Tett SFB (1999) Checking for model consistency in optimal finger printing. Clim Dyn 15: 419–434

Andrews DG, Holton JR, Leovy CB (1987) Middle atmosphere dynamics. Academic Press, San Francisco, 489 pp

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

Bretherton CS, Battisti DS (2000) An interpretation of the results from atmospheric general circulation models forced by the time history of the observed sea surface temperature distribution. Geophys Res Lett 27: 767–770

Chatfield C (1984) The analysis of time series. Chapman and Hall, London, 3rd edn

Collins M (2000) The El-Niño Southern Oscillation in the second Hadley Centre coupled model and its response to greenhouse warming. J Clim 13: 1299–1312

Covey C, AchutaRao KM, Lambert SJ, Taylor KE (2000) Intercomparison of present and future climates simulated by coupled ocean-atmosphere GCMs. PCMDI Report

Cox PM, Betts RA, Bunton CB, Essery RLH, Rowntree PR, Smith J (1999) The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. Clim Dyn 15: 183–203

Cox MD (1984) A primitive equation, three dimensional model of the ocean. Ocean Group Technical Report 1, GFDL Princeton

Cullen MJP (1993) The unified forecast/climate model. Meteorol Mag 122: 81–94

Cusack S, Slingo A, Edwards J, Wild M (1998) The radiative impact of a simple aerosol climatology on the Hadley Centre Atmospheric GCM. Q J R Meteorol Soc 124: 2517–2526

Edwards JM, Slingo A (1996) Studies with a flexible new radiation code. I: choosing a configuration for a large-scale model. Q J R Meteorol Soc 122: 689–720

Folland CK, Sexton DMH, Karoly DJK, Johnson CE, Rowell DP, Parker DE (1998) Influences of anthropogenic and oceanic forcing on recent climate change. Geophys Res Lett 25(3): 353–356

Gent PR, McWilliams JC (1990) Isopycnal mixing in ocean circulation models. J Phys Oceanogr 20: 150–155

Gillett NP, Allen MR, Tett SFB (2000) Modelled and observed variability in atmospheric vertical temperature structure. Clim Dyn 16(1): 49–61

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 transport in a version of the Hadley Centre coupled model without flux adjustments. Clim Dyn 16: 147–168

Gregory D, Kershaw R, Inness PM (1997) Parametrization of momentum transport by convection II: tests in single column and general circulation models. Q J R Meteorol Soc 123: 1153–1183

Gregory D, Shutts G, Mitchell J (1998) A new gravity wave drag scheme incorporating anisotropic orography and low level wave breaking; impact upon the climate of the UK Meteorological Office Unified Model. Q J R Meteorol Soc 124: 463–493

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

Hasselmann K (1976) Stochastic climate models. Part I: theory. Tellus 28: 473–485

Hasselmann K (1993) Optimal fingerprints for the detection of time dependent climate change. J Clim 6: 1957–1971

Hurrell J (1995) Decadal trends in the north atlantic oscillation: regional temperatures and precipitation. Science 269(5224): 676–679

Hurrell J, Trenberth KE (1997) Spurious trends in satellite MSU temperatures from merging different satellite records. Nature 386: 164–167

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, Jonsson T, Wheeler D (1997a) Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibralta and southwest Iceland. Int J Clim 17: 1433–1450

Jones PD, Osborn TJ, Briffa KR (1997b) Estimating sampling errors in large-scale temperature averages. J Clim 10: 2548–2568

Jones PD, Briffa KR, Barnett TP, Tett SFB (1998) High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with general circulation model control-run temperatures. The Holocene 8(4): 455–471

Jones PD (1994) Hemispheric surface air temperature variations: a reanalysis and an update to 1993. J Clim 7(11): 1794–1802

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. Bull Am Meteorol Soc 77(3): 437–471

Knutson TR, Manabe S (1998) Model assessment of decadal variability and trends in the tropical pacific. J Clim 11: 2273–2296

Latif M, Barnett TP (1996) Decadal climate variability over the North Pacific and North America: dynamics and predictability. J Clim 9: 2407–2423

Levitus S, Boyer TP (1994) World ocean atlas 1994, vol 4: temperature. NOAA/NESDIS E/OC21. US Department of Commerce, Washington, DC, 117 pp

Manabe S, Stouffer RJ, Spelman MJ, Bryan K (1991) Transient responses of a coupled ocean-atmosphere model to gradual changes of atmospheric CO_2. Part I: annual mean response. J Clim 4(8): 785–818

Milton S, Wilson C (1996) The impact of parametrized sub-grid scale orographic forcing on systematic errors in a global NWP model. Mon Weather Rev 124: 2023–2045

Mitchell JFB, Johns TC, Senior CA (1998) Transient response to increasing greenhouse gases using models with and without flux adjustment. HCTN 2, Hadley Centre, Bracknell, UK

Neelin JD, Battisti DS, Hirst AC, Jin FF, Wakata Y, Yamagata T, Zebiak SE (1998) ENSO theory. J Geophys Res 103(C7): 14 261–14 290

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

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

Parker D, Gordon M, Cullum D, Sexton D, Folland C, Rayner N (1997) A new global gridded radiosonde temperature data base and recent temperature trends. Geophys Res Lett 24: 1499–1502

Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parametrizations in the Hadley Centre climate model – HadAM3. Clim Dyn 16: 123–146

Rayner NA, Horton EB, Parker DE, Folland CK, Hackett RB (1996) Version 2.2 of the global sea-ice and sea surface temperature data set, 1903–1994. CRTN 74, Hadley Centre for Climate Prediction and Research Meteorological Office, Bracknell, RG12 2SY, UK

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

Santer BD, Taylor KE, Wigley TML, Penner JE, Jones PD, Cubasch U (1995) Towards the detection and attribution of an anthropogenic effect on climate. Clim Dyn 12: 77–100

Sausen R, Barthel K, Hasselmann K (1988) Coupled ocean-atmosphere models with flux corrections. Clim Dyn 2: 154–163

Scaife AA, Butchart N, Warner CD, Stainforth D, Norton WA, Austin J (2000) Realistic Quasi-Biennial Oscillations in a simulation of the global climate. Nature (in press)

Schlesinger ME, Ramankutty N (1994) An oscillation of the global climate system of period 65–70 years. Nature 367: 723–726

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

Smith RNB (1993) Experience and developments with the layer cloud and boundary layer mixing schemes in the UK Meteorological Office Unified Model. In: ECMWF/GCSS workshop on parametrization of the cloud-topped boundary layer. ECMWF, Reading, UK

Stouffer RJ, Manabe S, Vinnikov KY (1994) Model assessment of the role of natural variability in recent global warming. Nature 367(64): 634–636

Swinbank R, Lahoz WA, O'Neill A, Douglas CS, Heaps A, Pod D (1998) Middle atmosphere variability in the UK Meteorological Office Unified Model. Q J R Meteorol Soc 124(549): 1485–1526

Tett SFB, Mitchell JFB, Parker DE, Allen MR (1996) Human influence on the atmospheric vertical temperature structure: detection and observations. Science 247: 1170–1173

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

Tett SFB, Stott PA, Allen MR, Ingram WJ, Mitchell JFB (1999) Causes of twentieth century temperature change. Nature 399: 569–572

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

Timmermann A, Oberhuber J, Bacher A, Esch M, Latif M, Roeckner E (1999) Increased El-Niño frequency in a climate model forced by future greenhouse warming. Nature 398: 694–696

Trenberth KE, Hurrell JW (1994) Decadal atmosphere-ocean variations in the Pacific. Clim Dyn 9: 303–319

Visbeck M, Marshall J, Haine T, Spall M (1997) On the specification of eddy transfer coefficients in coarse resolution ocean circulation models. J Phys Oceanogr 27: 381–402

Williams KD, Senior CA, Mitchell JFB (2000) Transient climate change in the Hadley Centre models: the role of physical processes. J Clim (in press)

Wright DK (1997) A new eddy mixing parametrization and ocean general circulation model. Int WOCE Newslett 29: 27–29

Zhang Y, Wallace JM, Battisti DS (1997) ENSO-like interdecadal variability: 1900–93. J Clim 10: 1004–1020