Climate Dynamics

DOI 10.1007/s00382-002-0281-5

The North Atlantic Oscillation in coupled climate models: a CMIP1 evaluation

D. B. Stephenson(✉) · V. Pavan

participating CMIP1 modelling groups

D.B. Stephenson
Department of Meteorology, University of Reading, PO Box 243, Earley Gate, Reading RG6 6BB, UK

V. Pavan
ARPA-SMR, Viale Silvani 6, 40122 Bologna, Italy

Dr. G.J. Boer, Canadian Centre For Climate Modelling And Analysis, Victoria, Canada

Dr. Byron A. Boville, National Center for Atmospheric Research, Boulder, USA

Dr. Ulrich Cubasch, German Climate Computing Centre, Hamburg, Germany

Dr. Laurent Fairhead, Laboratoire de Meteorologie Dynamique du CNRS, Paris, France

Dr. Hal Gordon, CSIRO Division of Atmospheric Research, Victoria, Australia

Dr. Tim Johns, Hadley Centre, United Kingdom Meteorological Office, Bracknell, UK

Dr. Masahide Kimoto, Center for Climate System Research, Tokyo, Japan

Dr. Jerry Meehl, National Center for Atmospheric Research, Boulder, USA

Dr. Ron Miller, Goddard Institute for Space Studies, New York, USA

Dr. Akira Noda, Meteorological Research Institute, Ibaraki, Japan

Dr. Josef M. Oberhuber, German Climate Computing Centre, Hamburg, Germany

Dr. S. Power, Bureau of Meteorology Research Centre, Melbourne, Australia

Dr. Gary Russell, Goddard Institute for Space Studies, New York, USA

Dr. Ed Schneider, Center for Ocean-Land-Atmosphere Studies, Calverton, USA

Dr. Ron Stouffer, Geophysical Fluid Dynamics Laboratory, Princeton, USA

Dr. Laurent Terray, CERFACS - Global Change Team, Toulouse, France

Dr. Xue-Hong Zhang, LASG/IAP, Chinese Academy of Sciences, Beijing, China

E-mail: D.B.stephenson@reading.ac.uk

Received: 15 February 2002 / **Accepted:** 7 August 2002

Abstract. This study investigates the North Atlantic Oscillation (NAO) simulated by 17 global coupled ocean-atmosphere models participating in the Coupled Model Intercomparison Project (CMIP). Robust NAO indices are defined by calculating the leading principal components of winter time mean surface temperatures (land and sea) in the North Atlantic region (120°W-60°E, 20-80°N). Encouragingly, 13 out of 17 of the models capture the NAO surface temperature quadrupole pattern with centres of action over Northwest Europe, the northwest Atlantic, the southeastern USA, and the Middle East. The northern dipole is better captured than the southern dipole which is often simulated too far eastwards over the Atlantic Ocean. Out of the 17 models, ten models produce NAO indices that vary similar to the observations as stationary "weakly red noise" with only small correlations between successive winters ($r < 0.3$). Another five models drift monotonically towards warmer conditions, and two models exhibit long-term stochastic trends. Several of the models significantly overestimate the teleconnection between NAO and the tropical ENSO phenomenon.

1

Introduction

Mid-latitude surface temperature variations are strongly related to large-scale atmospheric flow patterns such as the North Atlantic Oscillation (Wallace et al. *1996*). The North Atlantic Oscillation (NAO) consists of a north-south alternation of pressure in the subtropical Azores high and the Iceland/polar low and represents the most dominant mode of climate variability in the northern extratropics (see Wanner et al. *2001* for a comprehensive review of the phenomena). In this study, we will focus on the more consistent and robust features in the Atlantic sector rather than on the currently popular hemispheric annular mode pattern, see Ambaum et al. (*2001*, *2002*) for a discussion of the difficulties in interpreting the annular mode. The NAO accounts for $R^2 = 31\%$ of the total variance in wintertime mean surface temperatures averaged from 20°N to 90°N for the period 1935-94 (Hurrell *1996*). Furthermore, the increasing trend in the North Atlantic Oscillation from the mid-1960s until the early 1990s helps account for many of the recent temperature trends over the North Atlantic region and surrounding continents (Hurrell *1995*; Wallace et al. *1996*). This important source of "*natural*" climate variability in the North Atlantic region urgently needs to be better understood and isolated, in order for us to be able to discriminate any possible "*unnatural*" changes that might start to occur due to global warming (Paeth et al. *1999*).

Coupled models are invaluable tools for understanding and forecasting climatic variations. However, the impossibility of having frequent global warming events prohibits posterior verification of coupled model projections of future climate. A less "*frequentist*" and more "*Bayesian*" approach to verification therefore needs to be adopted in which posterior uncertainty in forecasts of future events is estimated from prior confidence in how well each model performs in simulating natural climate variations. For example, if a particular model is not capable of simulating natural modes of variability in present-day climate, then it is reasonable to have less confidence in its projections about future climate change.

In contrast to the many El Nino/Southern Oscillation (ENSO) modelling studies, very few published studies have specifically evaluated the ability of climate models to simulate the NAO. Glowienka-Hense (*1990*) showed that the NAO was realistically simulated by an atmosphere-only general circulation model confirming that NAO is primarily related to atmospheric processes. Several studies have also shown that NAO can be captured quite realistically by coupled ocean-atmosphere models (Pittalwala and Hameed *1991*; Osborn et al. *1999*). By comparing coupled simulations made using different ocean models, Ulbrich and Christof (*1999*) demonstrated that the NAO in sea-level pressure variations simulated by ECHAM4 was not sensitive to the choice of ocean model. However, this does not preclude the possibility that the surface temperature signature in NAO can depend on the choice of ocean model and coupling scheme. Zhou et al. (*2000*) found that the IAP/LASG global coupled model was capable of simulating the main spatial patterns in sea-level pressure, surface temperature, and precipitation associated with the NAO. Despite its importance, the recent Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC *2001*) contained only one summary paragraph (8.7.5) on modelling of the NAO. The paragraph claims that "coupled climate models simulate the NAO quite well, although there are some differences in amplitude". The paragraph points out that "atmospheric models with prescribed SST also simulate the spatial pattern of NAO variability fairly well". These claims will be investigated in more depth in this article.

As a contribution to coupled model evaluation, we proposed an NAO subproject aimed at intercomparing the NAO simulated by the numerous global coupled ocean atmosphere models participating in the CMIP project. This report summarises the results obtained in this exploration of NAO in present-day coupled model simulations. Special care has been taken to use simple and reliable statistical techniques rather than more sophisticated techniques (e.g. canonical correlation analysis, multi-channel singular spectrum analysis, etc.) that can easily lead to erroneous interpretations. We have identified robust chararcteristics common to all the models, and also indicate some of the interesting exceptions. It is hoped that future studies will dig further into the model results to understand in more depth individual model behaviour. The results of this analysis do not necessarily reflect the "overall performance" of each of the coupled models.

The following section introduces the temperature data used in this study. Section 3 then discusses local variance in wintertime means followed by Sect. 4 which describes the behaviour of the leading large-scale NAO mode in surface temperature. Section 5 summarises the behaviour of the NAO indices using a simple stochastic model. The final section contains a summary and concluding remarks.

2

Surface temperature definition of NAO

2.1

Motivation

Simple (and rather crude) NAO indices are often obtained by taking the difference between standardised sea-level pressures from stations near the main "centres-of-action" of the NAO (the Azores high and the Icelandic low). Figure 1a shows the evolution of the NAO SLP index obtained by taking the difference between the standardised December-February mean sea-level pressures (SLP) at Gibraltar (UK) and Stykkisholmur (Iceland) (Jones et al. *1997*). In high index winters, the Icelandic low is deeper and the Azores high is stronger than normal, associated with a stronger mean westerly jet over the North Atlantic region.

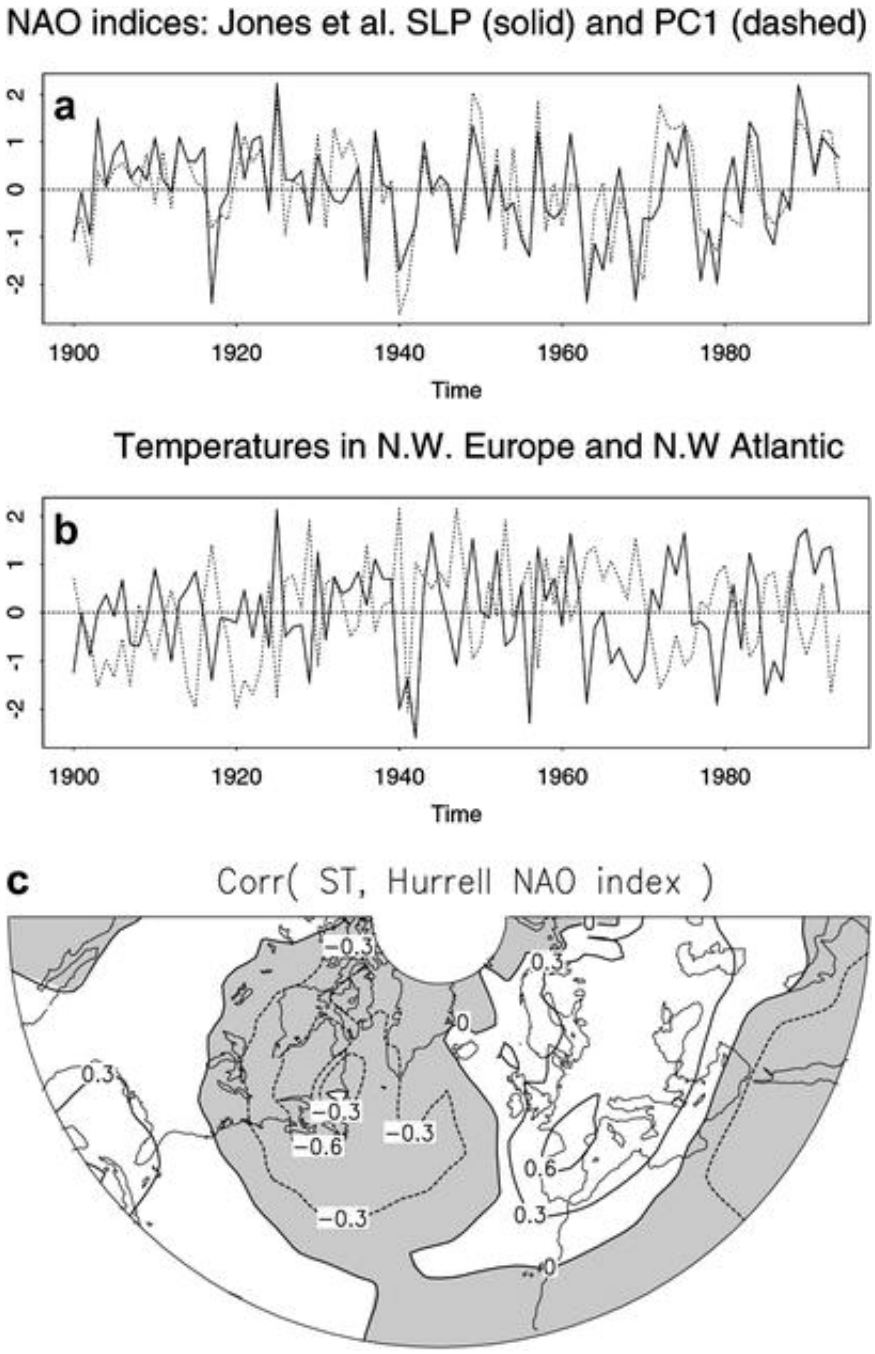

Fig. 1a-c. Some observed characteristics of the North Atlantic Oscillation. **a** Winter-time mean sea-level pressure index (Jones et al. *1997*); **b** winter-time mean surface temperatures averaged over large regions covering NW Europe (7.5°W-47°E, 52.5-72.5°N) and the NW Atlantic (22.5-77.5°W, 52.5-72.5°N); **c** correlation map of winter-time mean temperatures with the NAO SLP index

---

Unlike ENSO, the NAO is a broad-band phenomenon that has variations on both short and long time scales (Wunsch *1999*; Stephenson et al. *2000a*). Although more than 70% of the NAO variance is explained by "short-term" fluctuations having periods less than 10 years, marked decadal trends are also clearly present and have led to much speculation (Rogers *1984*; Wallace et al. *1996*; Trenberth *1990*; Hurrell *1996*; and references therein). Possible causes for decadal trends such as the increasing trend since the mid-1960s include: (a) external influences such as anthopogenically-induced global

warming or stratospheric ozone depletion, (b) long-term changes in the North Atlantic ocean, or (c) natural stochastic variations (Wunsch *1999*; Stephenson et al. *2000a*).

Although NAO in recent studies is most often defined in terms of sea-level pressure, many previous studies used the more robust signature in surface temperature to define the phenomenon. In fact, the NAO was originally discovered in the eighteenth century due to its 'seesaw' temperature signature between Greenland and northwestern Europe (van Loon and Rogers 1987). Furthermore, Walker and Bliss (*1932*) among others included temperature observations in their definition of NAO - for a comprehensive historical review see Stephenson et al. (*2002b*). The temperature seesaw is clearly apparent in Fig. 1c, which shows the correlation of surface temperatures with the Jones et al. (*1997*) NAO SLP index. Temperature has advantages in that its centres of action are more geographically fixed and are less noisy and trend-prone than those of sea-level pressure. These advantages plus the non-availability of sea-level pressure data in CMIP1 have led us to use surface (land and sea) temperatures to define NAO in this study.

2.2

Observed surface temperature data

For validation purposes, this study uses the Jones et al. (*1986*) 1851-1994 global land and sea monthly mean surface temperature data set (available from http://www.cru.uea.ac.uk/). This gridded data set was created by merging observed air temperatures at 2 m above the ground and temperatures at the sea surface into averages over 5° latitude-longitude boxes. It has been widely used for detection of global warming (IPCC *2001*). Unlike model data, the observed temperature data contains many "*missing values*" especially over grid boxes in remote high-latitude locations. Missing values can be problematic if improperly treated, and can give rise to misleading estimates of variability. For example, there are very few measurements near Greenland at the beginning of the record yet more in recent years. When all the available data are used to estimate the area-weighted average over the northwest Atlantic region (80°W-20°W, 50-75°N), the increasing number of cold observations close to Greenland gives the misleading impression that there has been a cooling trend of 5 °C since 1900. In an attempt to alleviate such problems, we have adopted a simple quality control approach that only uses data after 1900, and does not include ANY values from grid points that have less than 80 winter means defined over the period 1900-94. Winter time means were made at each grid point whenever one or more of the DJF months were available. Winter time anomalies were then calculated by subtracting the climatological means from the wintertime means.

Figure 1b shows the evolution of simple regional indices constructed from area averages of winter mean surface temperatures over Northwest Europe (NWE) (10°W-50°E, 50-75°N) and the northwest Atlantic (NWA) (80°W-20°W, 50-75°N). A clear negative correlation (r = -0.49) can be seen in the temperatures of the two regions. The difference between the two standardised indices is strongly correlated with the NAO SLP index shown in Fig. 1a (r = 0.77). In addition to the northern temperature seesaw, there is also a less well-known "southern seesaw" with centres of action in the southeastern USA and the Middle East (Fig. 1c). For example, when winters in Europe are mild, winters in Florida are also generally warmer and North Africa is colder than normal (Lamb *1972*). Walker and Bliss (*1932*) briefly discuss the southern NAO connections with temperatures at Charleston (Virginia, USA) and Cairo (Egypt). The NAO temperature signature is a robust spatial quadrupole pattern consisting of the well-known northern seesaw and the less well-known southern subtropical seesaw. The temperature correlations are associated with the convergence and divergence of air at the entrance and exit regions of the North Atlantic jet. For example, when the NAO index is stronger than normal (i.e. strong North Atlantic westerly jet), more cold air is advected from the north into the jet entrance region causing Greenland to become colder than normal (van Loon and Rogers

1978). Over most of the northern part of the Atlantic ocean, the temperatures are negatively correlated with the NAO due primarily to wind-induced changes in latent cooling (Bjerknes *1962*, *1964*).

2.3

Coupled model surface temperature data

The Coupled Model Intercomparison Project (CMIP) was initiated in 1995 under the auspices of the Working Group on Coupled Models (WGCM) of the World Meteorological Organisation CLIVAR program (Meehl *1995*). The purpose of CMIP is to examine climate variability and predictability as simulated by global coupled ocean-atmosphere models, and to compare the model output with observations where available. The initial stage, CMIP1, is a "*Level 1*" intercomparison of "*control runs*" in which external forcing terms such as carbon dioxide concentration and solar luminosity are held constant at present-day values (Meehl *1995*). The participating modelling groups provided time series of gridded monthly mean surface temperatures produced by control runs of their coupled ocean-atmosphere models. The total data set amounts to 43,452 months (3621 years) of global surface temperature values (840 Mb of data) and was quality controlled, and reformatted into netCDF format by the PCMDI group at Lawrence Livermore National Laboratory before been redistributed to sub-project investigators. The models constitute a representive sample of global coupled ocean atmosphere models currently used in climate research. Many of the participating models have also been used to make *"projections*" of the possible climate in the next century. It is therefore of great interest to assess the extent to which such models can simulate present-day climate and its variability.

Table 1 summarises the different CMIP1 coupled model runs that have been analysed as part of this study, more details can be found at http://www-pcmdi.llnl.gov/modeldoc/cmip/. Our study is based on the CMIP1 model simulations that were made before January 1997, and so are suggestive but not necessarily indicative of the most recent model versions. Most of the model runs are labelled by the name of the institution that performed the run and supplied the data. However, data from two quite different runs was supplied by MPI/DKRZ: ECHAM3 and ECHAM4. ECHAM3 denotes a run made with the ECHAM3 atmospheric model coupled to the LSG ocean model, whereas ECHAM4 denotes a run made with the ECHAM4 atmospheric model coupled to the quite different ocean model OPYC3. For brevity, these shortened denominations will be used throughout the rest of this work. The UKMO data was generated using the HadCM2 version of the coupled model at the UKMO, and CERFACS denotes data generated from the coupled ARPEGE-OPA model, with ARPEGE being the atmospheric model from Météo-France and OPA being an ocean model from the LODYC group at the University of Paris VI. There is a wide diversity in the various model experimental designs. For example, two of the coupled runs are longer than 1000 years (GFDL and UKMO), whereas five of the coupled runs are shorter than 50 years (CCSR, CERFACS, COLA, IAP/LASG, LMD/IPSL). Care should be exercised not to over-interpret results from these shorter runs. The equivalent spatial resolution is the square root of the average area on the sphere covered by a grid box ($2\pi a^2/p^2$) and provides a useful *single* measure of the model horizontal resolution. It ranges from distances larger than 500 km for the lowest resolution models such as CERFACS, COLA, ECHAM3, GFDL, IAP/LASG, NCAR (W&M) to values less than 300 km for the highest resolution models such as ECHAM4, LMD/IPSL, NCAR(CSM). The majority of the models represent the atmospheric fields using spherical harmonics, except for five models which use finite differences [UKMO, MRI, GISS(Russell), GISS(Miller), LMD/IPSL]. Flux adjustment is a pragmatic way of correcting for systematic errors in coupled model air-sea fluxes, but it is only applied in less than half (eight out of 17) of the models participating in CMIP1. It should be noted, however, that four out of the six runs longer than 100 years correct the heat and freshwater fluxes in order to prevent excessive model drift. Flux adjustment may become less necessary in the future for models having smaller systematic errors, for example, more recent UKMO runs now no longer use any kind of flux adjustment. CMIP1 is a level-1 intercomparison not only

because the models have different initial conditions etc. but also because the surface temperature is not defined the same way for all the models. The majority of models define the temperature at either 1.5 m or 2 m above the ground, yet some models even define it on model levels or at 10 m above the ground. Fortunately, temperatures are generally well-mixed vertically in the tropospheric boundary layer over the North Atlantic and so the different level definitions are not expected to cause major differences in temperature behaviour. However, for stable regions over cold land or sea-ice, the choice of level could cause noticeable differences. All but two of the models calculate sea-ice thickness and extent prognostically, and seven out of these also try to account for the advection of sea-ice. We will avoid discussing the sea-ice behaviour in detail since this is the focus for another CMIP sub-project.

Table 1. Description of the CMIP1 model runs used in this study

Model	IPCC name	YRS	RESOLUTION		FADJ	TLEV	ICE	Reference
UKMO	HadCM2	1085	286 km	2.5° × 3.8°	*HW*	1.5 m	*F*	Johns et al. (*1997*)
GFDL	GFDL_R15_a	1000	543 km	R15	*HW*	0.990	*F*	Manabe and Stouffer (*1996*)
NCAR(CSM)	CSM1.0	300	262 km	T42	-	2 m	*R*	Boville and Gent (*1998*)
ECHAM4	ECHAM4/OPYC	240	262 km	T42	*HW*	2 m	*R*	Roeckner et al. (*1996*)
CCCma	CGCM1	150	355 km	T32	*HW*	2 m	*T*	Flato et al. (*2000*)
BMRC	BMRCa	105	396 km	R21	-	1.5 m	*T*	Power et al. (*1993*)
CSIRO	CSIRO Mk2	100	396 km	R21	*HWM*	2 m	*R*	Gordon and O'Farrell (*1997*)
MRI	MRI1	100	418 km	4° × 5°	*HW*	0 m	*F*	Tokioka et al. (*1996*)
NCAR (W&M)	NCAR1	100	543 km	R15	-	0.991	*R*	Washington and Meehl (*1996*)
GISS(Russell)	GISS2	98	418 km	4° × 5°	-	≈ 0.985	*T*	Russell et al. (*1995*)
GISS(Miller)	GISS1	89	418 km	4° × 5°	-	10 m	*T*	Miller and Jiang (*1996*)
COLA	COLA1	50	543 km	R15	-	Skin	*T*	Schneider et al.(*1997*)
ECHAM3	ECHAM3/LSG	50	524 km	T21	*HWM*	2 m	*T*	Voss et al.(*1998*)
IAP/LASG	GOALS	50	543 km	R15	*H*	0.991	*T*	Zhang et al. (*2000*)
CCSR	CCSR/NIES	40	524 km	T21	*HW*	0.995	*T*	Abe-Ouchi et al. (*1996*)
CERFACS	ARPEGE/OPA1	40	524 km	T21	-	2 m	-	Guilyardi and Madec (*1997*)
LMD/IPSL	IPSL/CM1	24	231 km	1.6 × 3.8	-	0.979	-	Branconnot et al. (*1997*)

YRS is the total integration length in years. RESOLUTION gives the equivalent spatial resolution (see text for details), and either the spectral truncation for spectral models or the latitude-longitude grid spacing for non-spectral models. FADJ indicates what type of flux adjustments were applied to the

heat (*H*), water (*W*), or horizontal momentum (*M*) fluxes at the air-sea interface (no flux correction is denoted by -). TLEV gives the altitude or model σ -level at which the surface temperature was defined. ICE denotes the type of sea-ice parameterization used: thermodynamic sea-ice model (*T*), thermodynamic model allowing free-drift of sea-ice (*F*), and thermodynamic model including sea ice rheology (*R*) (- signifies that sea-ice was obtained in a simple diagnostic way from SSTs). The reference cites the most recent comprehensive publication documenting the coupled model

3

Local variance in winter time mean temperatures

3.1

Total variation

Figure 2 shows maps of the standard deviation of the winter time mean temperatures. The largest temperature variations occur over the interior of the North American and Eurasian continents with typical amplitudes of 2-5 °C. Several of the models overestimate the standard deviation over the continental regions: BMRC, CCSR, NCAR (CSM), and NCAR (W&M). BMRC and NCAR (W&M) also exhibit excessive summer minus winter temperature differences over the continents (not shown) which suggests that these models may be underestimating surface heat capacity over land (perhaps due to deficiencies in soil moisture). Several of the models have local maxima in variance near the North Atlantic sea-ice edge, for example, BMRC, ECHAM4+OPYC, GISS(Miller), GISS(Russell), NCAR(CSM), UKMO. All these models determine sea-ice extent and thickness prognostically and exhibit particularly strong gradients in the summer-winter temperature difference at the sea-ice edge (Covey et al. *2000*). It is possible that these models are overestimating the thickness of the sea ice, which commonly occurs when using prognostic sea-ice models (Garric et al. *1997*). It is also of interest to note that the CCSR model has an unrealistic maximum in variance over northeast Canada and Greenland, and the COLA model has an unrealistic maximum in variance over the Arctic region.

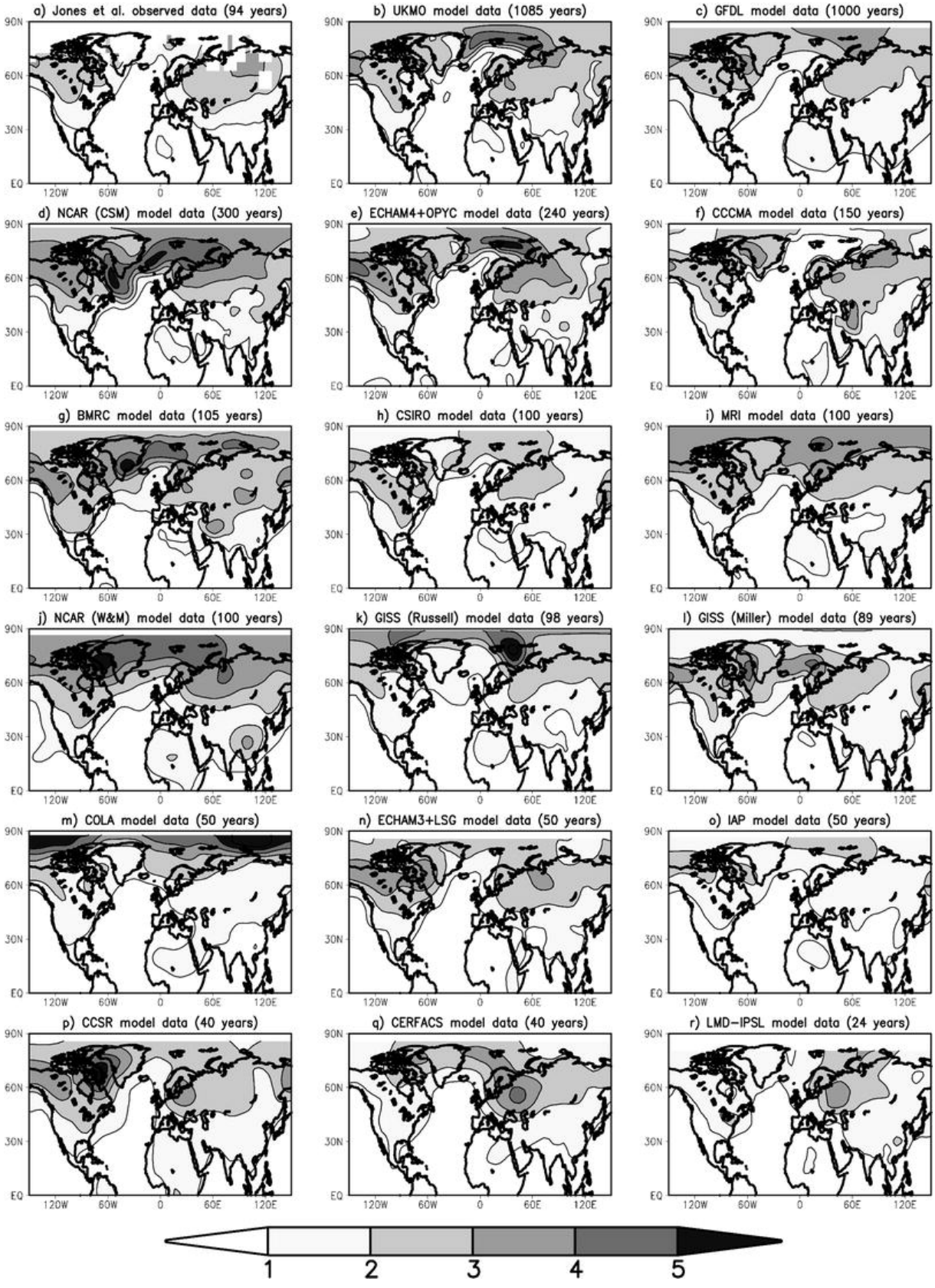
a) Jones et al. observed data (94 years)
b) UKMO model data (1085 years)
c) GFDL model data (1000 years)
d) NCAR (CSM) model data (300 years)
e) ECHAM4+OPYC model data (240 years)
f) CCCMA model data (150 years)
g) BMRC model data (105 years)
h) CSIRO model data (100 years)
i) MRI model data (100 years)
j) NCAR (W&M) model data (100 years)
k) GISS (Russell) model data (98 years)
l) GISS (Miller) model data (89 years)
m) COLA model data (50 years)
n) ECHAM3+LSG model data (50 years)
o) IAP model data (50 years)
p) CCSR model data (40 years)
q) CERFACS model data (40 years)
r) LMD−IPSL model data (24 years)
90N
60N
30N
EQ
120W
60W
0
60E
120E
1
2
3
4
5

Fig. 2a-c. Standard deviation of the winter-time means of surface temperature (in °C). Note the large values over the interior of the continents and near the sea-ice edge

---

3.2

Short-term interannual variations

Bjerknes (*1962*) made the important distinction between the passive North Atlantic upper-ocean response resulting in "short-term 2-5 year trends" (his words) and the longer term coupled variations. To isolate shortterm interannual variations, he and his student assistants at UCLA, made and examined many maps of year-to-year differences ($\Delta z = z_y - z_{y-1}$) in North Atlantic SSTs and SLPs for the period 1880 to 1915 (Bjerknes *1964*). First-order differencing is widely used by statisticians and economists to reduce long term trends in non-stationary time series (Box and Jenkins *1976*). Differencing reduces piecewise linear trends to small constant terms, and has the advantage over removing a linear-in-time fit in that it is local in time. This is desirable since climatic time series often contain randomly varying trends (stochastic trends), for example, the NAO SLP index had a decreasing decadal trend from 1940-60, but then an increasing one from 1960-1990. Because (Δz)/2 is simply the result of subtracting a 2-year moving average ($z(y) + z(y - 1)$)/2 from the original series, Δ/2 can be considered to be a simple high-pass filter. It attenuates the amplitude of low-frequency signals by a factor of $\sin(\pi f \tau)$ where f is the frequency and τ is the time between samples (e.g. 1 year). Therefore, biennial signals with periods of two years suffer no attenuation whereas signals with periods of four years have amplitudes attenuated by a factor of $1/\sqrt{2}(\sin(\pi/4))$. Decadal and even lower frequency variability are attenuated to less than 10% of their original amplitudes. These features and the fact that year-to-year differences in bulk quantities such as upper ocean heat content can be simply interpreted in terms of integrated fluxes justify the use of this method.

Figure 3 shows the standard deviation of the short-term $z'_y = \Delta z/2$ contribution to winter time mean temperatures. The spatial distribution generally resembles that of the total standard deviation shown in Fig. 2, yet with smaller amplitudes of 1-3 °C over the continents. An important feature that emerges in the observations and some of the models (UKMO, ECHAM4) is a local maximum in the eastern equatorial Pacific related to ENSO. However, many of the coupled models underestimate the ENSO variance possibly due to insufficient latitudinal resolution at the equator in the ocean models. This is supported by noting that only the higher resolution ECHAM4 and UKMO runs capture the maximum in the Pacific with a realistic amplitude. It is worth noting that despite underestimating the variance of ENSO, most of the CMIP1 runs still manage to simulate realistic amounts of variance over the Northern Hemisphere extra-tropics.

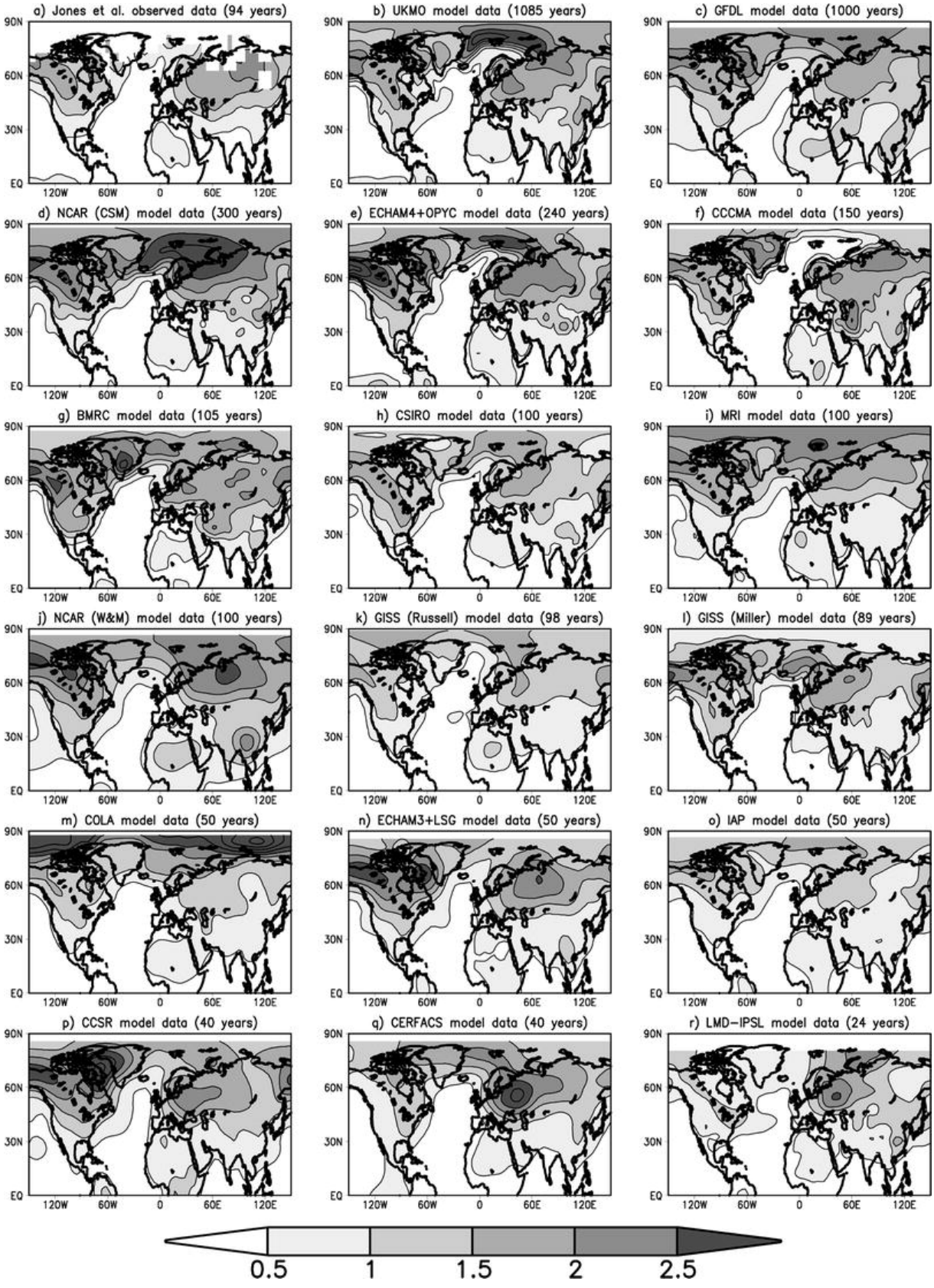
a) Jones et al. observed data (94 years)
b) UKMO model data (1085 years)
c) GFDL model data (1000 years)
d) NCAR (CSM) model data (300 years)
e) ECHAM4+OPYC model data (240 years)
f) CCCMA model data (150 years)
g) BMRC model data (105 years)
h) CSIRO model data (100 years)
i) MRI model data (100 years)
j) NCAR (W&M) model data (100 years)
k) GISS (Russell) model data (98 years)
l) GISS (Miller) model data (89 years)
m) COLA model data (50 years)
n) ECHAM3+LSG model data (50 years)
o) IAP model data (50 years)
p) CCSR model data (40 years)
q) CERFACS model data (40 years)
r) LMD−IPSL model data (24 years)
90N
60N
30N
EQ
120W
60W
0
60E
120E
0.5
1
1.5
2
2.5

Fig. 3a-c. Standard deviation of short-term variations in winter-time means of surface temperature (in °C). Note the resemblance to the geographical distribution of the total standard deviations in Figure 2

3.3

Serial correlation

In order to quantify the fraction of variance in short-term variations, Fig. 4 shows the ratio (in %) of the short-term standard deviation (Fig. 3) to the total standard deviation (Fig. 2). The ratio of standard deviations is related to the 1-year lag autocorrelation coefficient, r_1, by the expression $\sqrt{(1-r_1)/2} \times 100$ %, and therefore provides a simple measure of the amount of "serial correlation" or "redness" in the signal. Very "red" signals have a large positive autocorrelation and small standard deviation ratios, whereas strong biennial "blue" signals have a large negative autocorrelation and standard deviation ratios close to 100%. "White" noise is intermediate between these two extremes with an autocorrelation of zero and a ratio of standard deviations equal to 70.7%. Ratios in Fig. 4 are typically 60-80% over the continents and polar sea-ice (white), and 10-60% (red) over the oceans. One reason for the increased serial correlation (redness) over the ocean is that the heat capacity of the upper ocean mixed-layer is much larger than that of the drier continental surfaces (Hasselmann *1976*). Particularly red regions occur in the North Atlantic ocean in several models: BMRC, GISS(Miller), GISS(Russell), NCAR(CSM) and NCAR(W&M). The North Atlantic redness in BMRC, NCAR(W&M), and GISS(Miller) is associated with significant monotonic warming trends (climate drift) that will be discussed in the next section. The redness over the Atlantic ocean is underestimated in the two longest runs, UKMO and GFDL, which perhaps indicates that flux adjustment may be over damping long-term Atlantic fluctuations in these two models.

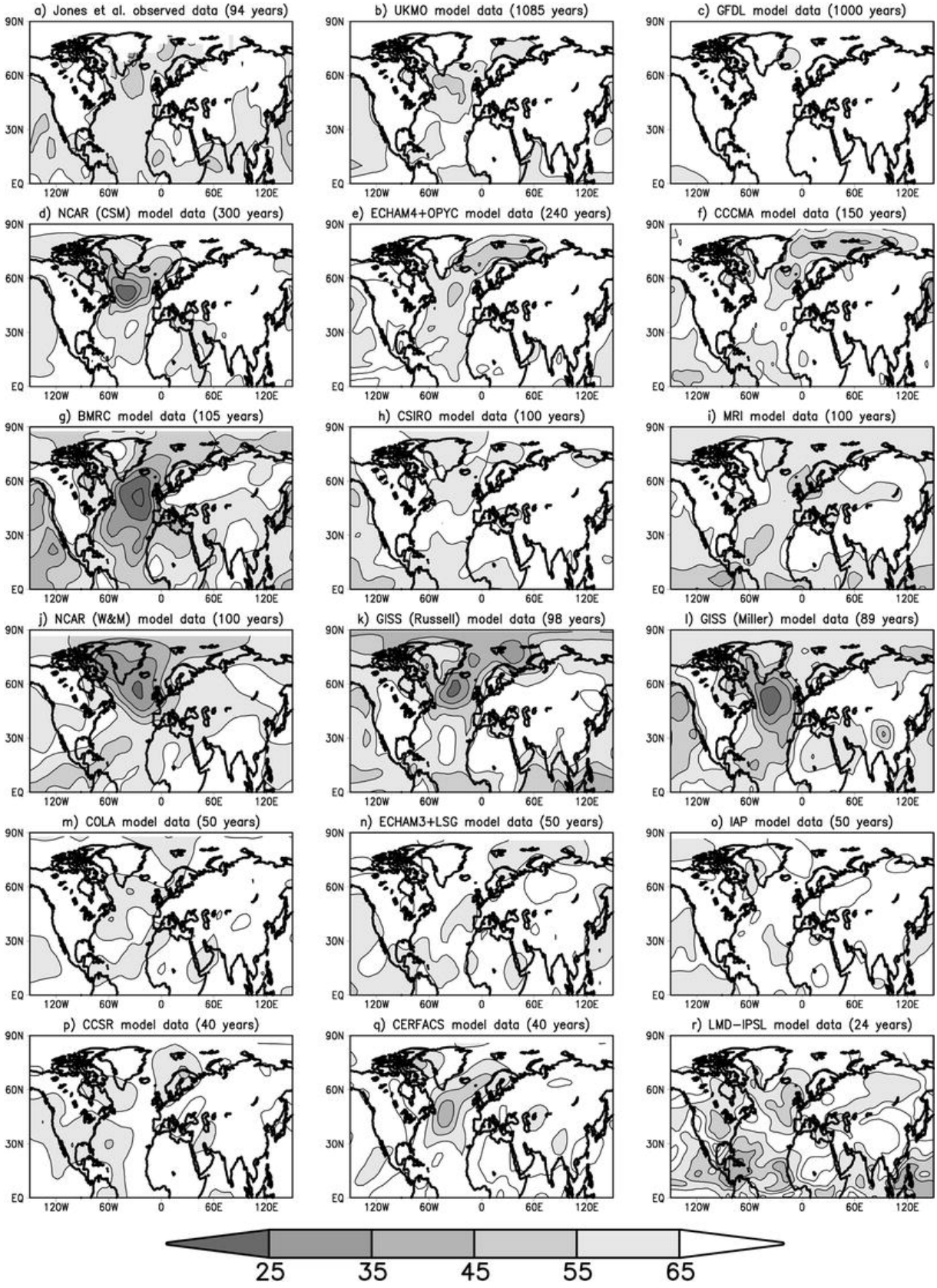
a) Jones et al. observed data (94 years)
b) UKMO model data (1085 years)
c) GFDL model data (1000 years)
d) NCAR (CSM) model data (300 years)
e) ECHAM4+OPYC model data (240 years)
f) CCCMA model data (150 years)
g) BMRC model data (105 years)
h) CSIRO model data (100 years)
i) MRI model data (100 years)
j) NCAR (W&M) model data (100 years)
k) GISS (Russell) model data (98 years)
l) GISS (Miller) model data (89 years)
m) COLA model data (50 years)
n) ECHAM3+LSG model data (50 years)
o) IAP model data (50 years)
p) CCSR model data (40 years)
q) CERFACS model data (40 years)
r) LMD−IPSL model data (24 years)
90N
60N
30N
EQ
120W
60W
0
60E
120E
25
35
45
55
65

Fig. 4a-c. Ratio in % of the standard deviation of the short-term variations to the total standard deviation of the winter-time means. Small ratios indicate the presence of serial correlation caused either by persistence or decadal/secular trends

---

4

Large-scale NAO temperature patterns

4.1

The leading NAO temperature mode of variability

In this study, the leading large-scale patterns and indices associated with NAO have been extracted by performing a principal component analysis (PCA) on each of the gridded temperature datasets. To optimally extract the NAO, the spatial domain was limited to the North Atlantic ocean and surrounding land regions (120°W-60°E, 20-80°N). The Pacific sector and tropical Atlantic were excluded in order to minimise any contamination from strong tropical signals such as ENSO. Stephenson et al. (*2000a*) showed that if these regions are included the NAO pattern still emerges but only as the second leading EOF. The Arctic region north of 80°N was also excluded in order to reduce contamination that may be caused by anomalous Arctic sea-ice behaviour. Because surface temperature anomalies over land are typically between 3-10 times larger than those over the oceans, it is necessary to do PCA on the correlation matrix (standardised anomalies) rather than on the covariance matrix (non-standardised anomalies). The PCA loading weights (EOFs) can therefore be considered as "correlation maps" obtained by correlating the temperatures with the respective PCs. Large magnitudes in the EOF patterns indicate strong correlations with other regions.

Figure 5 shows the leading spatial correlation pattern, EOF1, for the observations and for all the models. Because of the strong correlation between PC1 and the NAO SLP index (r = 0.74), the leading EOF spatial pattern of the observed temperatures shown in Fig. 5a closely resembles the correlation map of temperatures with the NAO SLP index (Fig. 1c). Principal component analysis over the chosen extratropical domain is capable of extracting a robust NAO index and associated correlation pattern. The spatial pattern is dominated by a large-scale quadrupole pattern with nodes centered over Northwest Europe and the northwest Atlantic (the northern seesaw), and the southeastern United States and the Middle East (the southern seesaw). It can be noted in Fig. 5 that many of the models are able to capture the salient features of the quadrupole pattern. The northern seesaw is simulated to varying extents by all of the models except for BMRC, MRI, NCAR(W& M), and GISS(Miller). The southern seesaw is captured less well by many of the models but appears quite realistic in the long runs made with UKMO, GFDL, NCAR(CSM), and ECHAM4. In some models, the western node of the southern seesaw is displaced too far east across the North Atlantic ocean e.g. CCCma, CSIRO, and CCSR.

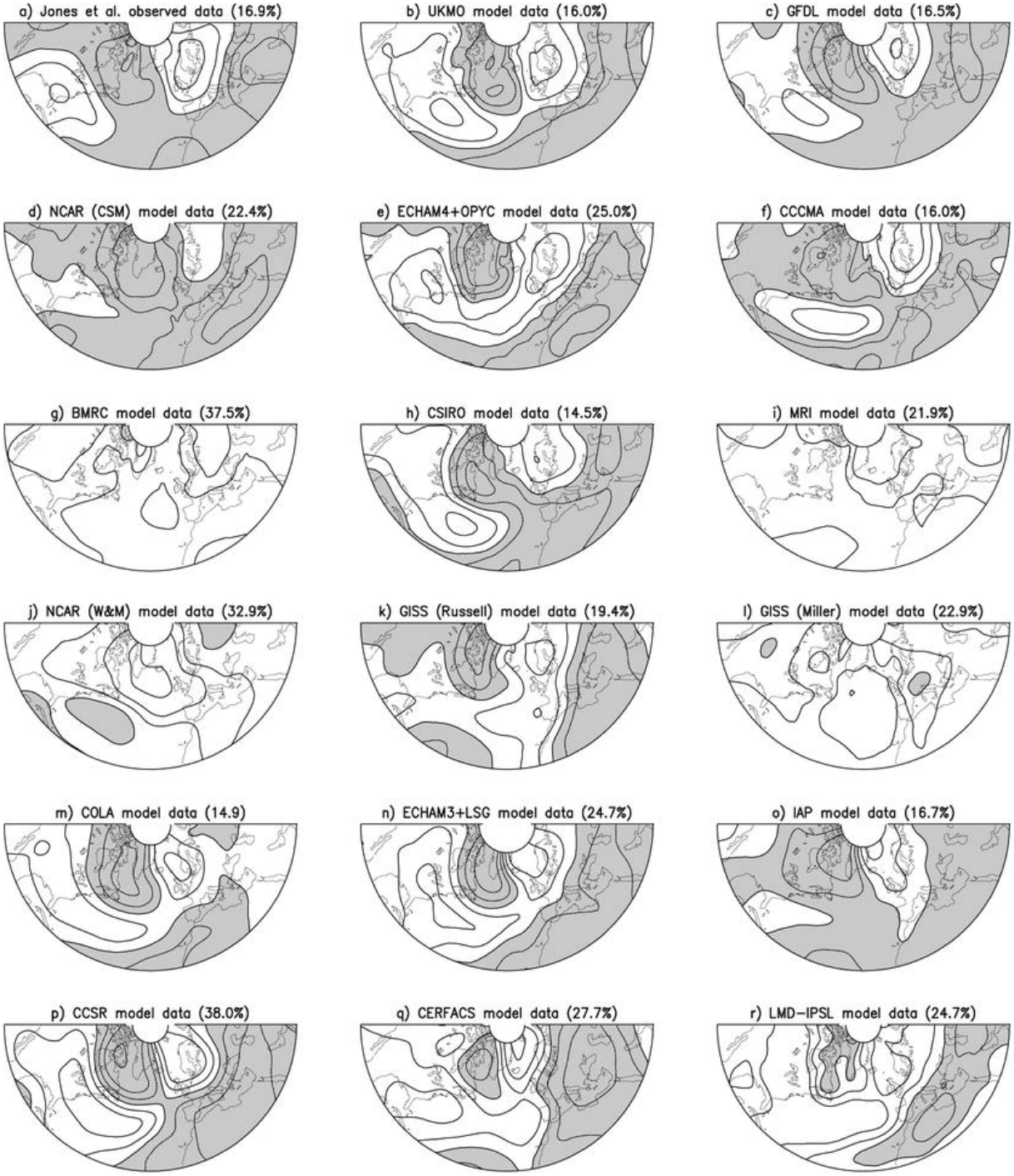

Fig. 5a-c. Leading eigenvector of the correlation matrix of the winter-time mean surface temperature (EOF1). Note the large-scale quadrupole pattern with nodes centered over NW Europe, NW Atlantic, the southeastern states of North America, and the Middle East. Values have been normalised so as to represent correlations with PC1 (in %)

Figure 6 shows the leading EOF recalculated for "detrended" year-to-year differences in the winter time mean temperatures. For the observations and many of the models, the leading EOFs of the detrended short-term variations closely resemble the EOF1 of the non-detrended variations shown in Fig. 5. This is due to the dominance of short-term variations in the NAO signal, more than 70% of the

observed NAO SLP index variance is explained by signals having periods less than ten years (Stephenson et al. *2000a*). It is encouraging to note that an NAO-like pattern now also emerges in the EOFs of the detrended data from the BMRC, NCAR(W&M), and GISS(Miller) models. The leading EOF of the non-detrended data from these models was capturing the coupled model drift instead of the desired NAO natural variability.

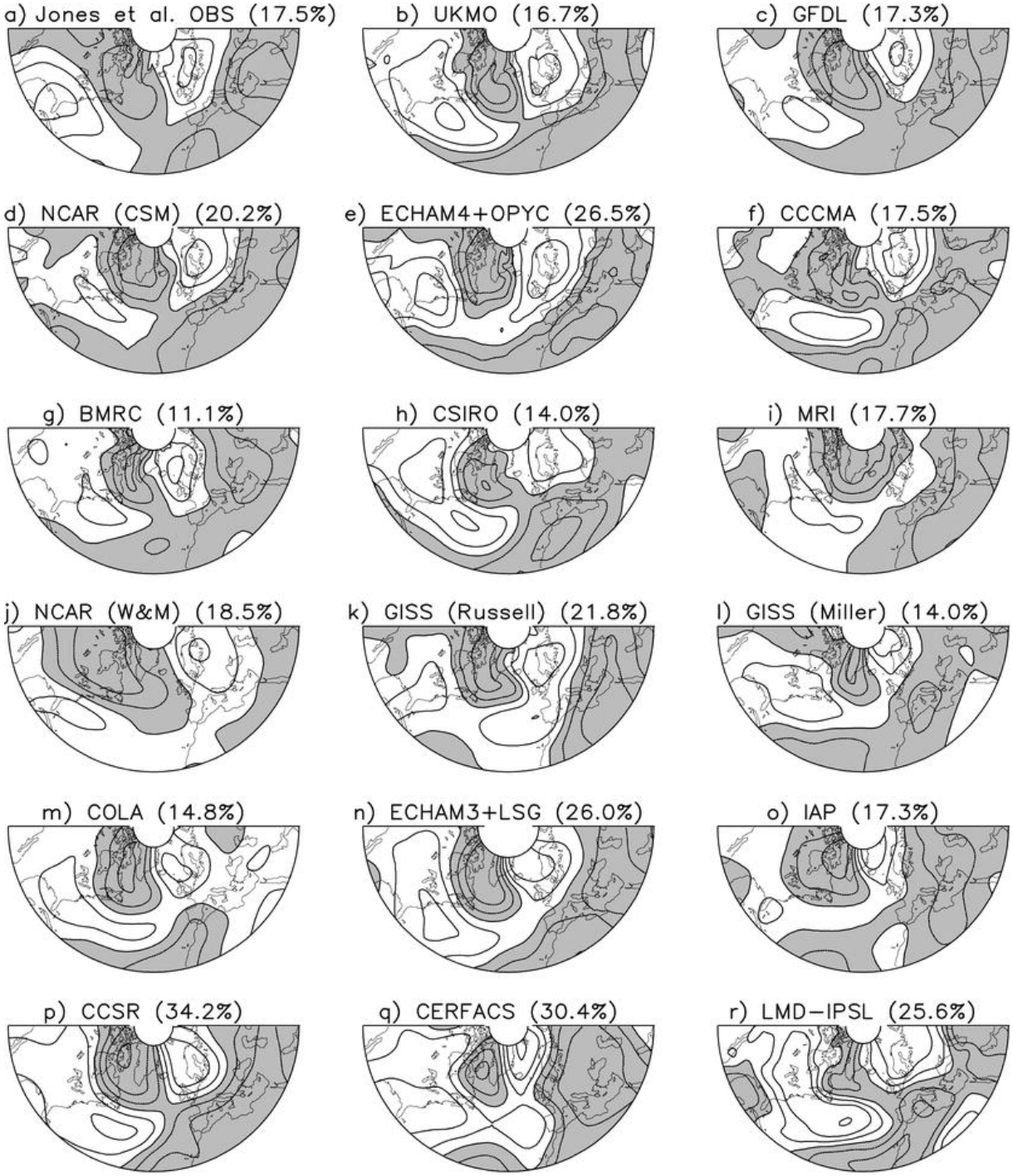

Fig. 6a-c. Same as Fig. 5 except for short-term interannual variations in wintertime mean temperatures (obtained from year-to-year differences). Note the similar patterns to those in Fig. 5. for many of the models

---

Figure 7 shows the leading principal component time series. There is a rich diversity of behaviour with some of the models containing obvious monotonic trends whereas other models appear to give more stationary and stochastic variations. Primarily because of the large inertia of the deep ocean, coupled ocean atmosphere models require a long time to reach climatic equilibrium. During the approach to equilibrium (the "spin-up" phase), results from such models often contain strong monotonic trends

("climate drift"). Such non-stationarity represents an unnatural form of climate variability that can easily contaminate natural variations simulated by such models. Monotonic trends are clearly apparent in Fig. 7 in the NAO PC1 time series from BMRC, MRI, NCAR(W&M), GISS(Miller) and LMD/IPSL. Furthermore, surprisingly all of these trends are towards increasing NAO in time. Monotonic trends due to model drift can be robustly quantified using the Spearman rank correlation coefficient of PC1 with time. The Spearman rank correlation coefficient is invariant for any monotonic function of the time series and so can be used to test for all types of monotonic trend (not necessarily linear). Column 2 of Table 2 gives the Spearman correlation coefficients for all the PC1 time series, statistically significant trends different from zero at the 1% level of significance are indicated by asterisks (two-tailed Student *t*-test assuming no serial correlation). Five of the models have significant increasing trends in PC1: BMRC, MRI, NCAR(W&M), GISS(Miller), and LMD/IPSL. In all the models except GISS(Miller), the trends are towards warming in surface temperature over northwest Europe and most of the rest of the Northern Hemisphere. BMRC, NCAR (W&M), GISS(Miller) also exhibited a very marked redness in the North Atlantic ocean in Fig. 4. This suggests that the drift may be related to dynamical coupled errors in the North Atlantic ocean induced by coupling (Stephenson *2002*). The remaining 12 models do not have significant monotonic trends and resemble more closely stationary noise. It is encouraging to note that almost half of these stationary models do not use flux adjustment.

Table 2. Correlations between various indicies calculated from observations and model simulations

Model	(PC1, year)	(NWA, NWE)	(NH, PC1)	(NWA, PC1)	(NWE, PC1)	(NINO3, PC1)
Observations 1901-91	-0.02	-0.49 (*)	+0.03	-0.60 (*)	+0.72 (*)	-0.24
UKMO	+0.03	-0.38 (*)	+0.08	-0.72 (*)	+0.76	-0.20 (*)
GFDL	-0.01	-0.37 (*)	-0.18 (*)	-0.85 (*)	+0.61 (*)	+0.00
NCAR(CSM)	-0.01	-0.01	-0.71 (*)	-0.91 (*)	-0.03	-0.33 (*)
ECHAM4	-0.01	+0.08	+0.07	-0.71 (*)	+0.45 (*)	-0.36 (*)
CCCma	-0.01	-0.57 (*)	-0.10	-0.79 (*)	+0.84 (*)	-0.15
BMRC	-0.62 (*)	+0.46 (*)	+0.97 (*)	+0.90 (*)	+0.70 (*)	+0.57 (*)
CSIRO	+0.04	-0.50 (*)	+0.17	-0.70 (*)	+0.70 (*)	-0.01
MRI	-0.59 (*)	+0.05	+0.80 (*)	+0.49 (*)	+0.81 (*)	+0.41 (*)
NCAR (W&M)	-0.89 (*)	+0.38 (*)	+0.77 (*)	+0.96 (*)	+0.50 (*)	-0.01
GISS(Russel)	-0.09	-0.19	-0.14	-0.29 (*)	+0.66 (*)	-0.01
GISS(Miller)	-0.92 (*)	+0.33 (*)	+0.96 (*)	+0.91 (*)	+0.54 (*)	-0.04
COLA	-0.02	-0.20	+0.06	-0.77 (*)	+0.56 (*)	-0.21
ECHAM3	-0.10	-0.47 (*)	-0.11	-0.76 (*)	+0.64 (*)	-0.05
IAP/LASG	-0.05	-0.36 (*)	-0.28	-0.77 (*)	+0.40 (*)	-0.00
CCSR	-0.01	-0.82 (*)	-0.63 (*)	-0.93 (*)	+0.85 (*)	-0.43 (*)
CERFACS	+0.16	-0.26	-0.25	-0.19	-0.01	+0.27
LMD/IPSL	+0.72 (*)	-0.21	+0.68 (*)	+0.17	+0.67 (*)	+0.44

Spearman rank correlation has been used for the 2nd coloumn to robustly test the existence of monotonic time trends in the NAO indices. Asterisks indicate strong correlations significantly different from zero at 95% confidence (two-sided t-test)

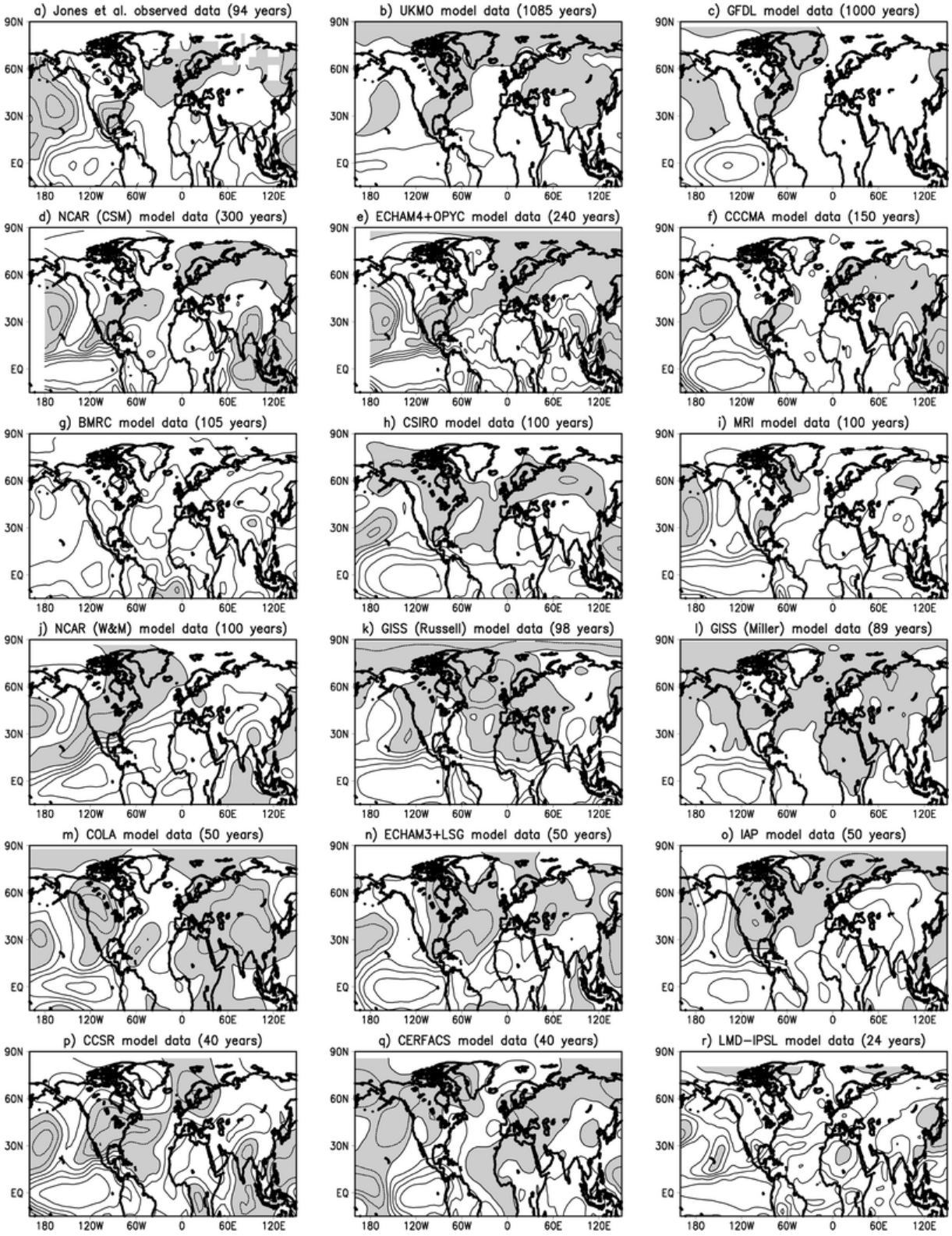

Fig. 7a-c. Time series of the leading principal component (PC1) of the winter-time mean surface temperatures. All series have been standardised to have unit variance

4.2

Northern Hemisphere NAO teleconnections

The North Atlantic Oscillation is a large-scale phenomenon that has "teleconnections" (correlations) with climate in many widely-separated regions. To quantify such teleconnections, we have constructed several simple climatic indices by averaging surface temperatures over the following regions: Northern Hemisphere (NH), northwest Atlantic (NWA), Northwest Europe (NWE), eastern equatorial Pacific (NINO3). These simpler indices are often easier to interpret than multivariate indices such as principal components and can provide useful complimentary information. Table 3 gives the cross-correlations between the various indices based on observations from 1901-91 both for total non-detrended variations (above the leading diagonal) and for shortterm year-to-year differences (below the leading diagonal). With the exception of correlations with Northern Hemisphere temperatures NH, many of the correlations are similar for both the total and shortterm variations. PC1 has a large and significant correlation (r = 0.74) with the NAO SLP index justifying its use as a temperature based NAO index.

Table 3. Cross-correlations between different indices baswed on observations for the common period 1900-91

Series	NAO SLP	PC1	NH	NWA	NWE	NINO3
NAO SLP	-	+0.74 (*)	+0.21	-0.63 (*)	+0.70 (*)	+0.02
PC1	+0.68 (*)	-	+0.03	-0.60 (*)	+0.72 (*)	-0.24
NH	+0.30 (*)	-0.05	-	+0.04	+0.32 (*)	+0.25
NWA	-0.60 (*)	-0.62 (*)	-0.22	-	-0.49 (*)	-0.06
NWE	+0.67 (*)	+0.66 (*)	+0.33 (*)	-0.61 (*)	-	-0.11
NINO3	-0.01	-0.27	+0.43 (*)	+0.06	-0.09	-

Values above the leading diagonal give correlations of total variations, whereas values below the diagonal give correlations of only short-term variations (year-to-year differences). Asterisks indicate strong correlations that are significantly different from zero at 99% confidence (two-tailed t-test)

Table 2 includes cross-correlations between non-detrended indices calculated for the observations and for all the models. Year-to-year differences give similar correlations for the stationary models (not shown). Many of the models capture the northern seesaw as can be noted by the significant non-zero (anti-)correlations between NWA and NWE temperatures (negative values marked with * in column 1 of Table 3). BMRC, NCAR(W&M) and GISS(Miller) have significant positive correlations between NWE and NWA caused by unrealistic coupled model drift towards warmer conditions. With the exception of NCAR(CSM) and CERFACS, all the stationary models have significant negative correlations between PC1 and NWA and significant positive correlations between PC1 and NWE as seen in the observations. NCAR(CSM) has an unrealistic poor correlation between PC1 and NWE because of a marked eastward shift in the position of the "northern seesaw" (Fig. 5d). CERFACS has an unrealistic poor correlation between PC1 and NWE caused by its ’northern seesaw’ being smaller longitudinally than the dipole in the observations (Fig. 5q).

Hurrell (*1996*) demonstrated that the NAO can account for 31% of the variance of the Northern Hemisphere mean temperature north of 20°N. However, from Table 2 it can be seen that the NAO SLP index accounts for less than 3% ($0.21^2 \times 100\%$) of the total variance and 9% ($0.30^2 \times 100\%$) of the

short-term variance in the observed Northern Hemisphere average temperature over the period 1901-91. The temperature based NAO index, PC1, is even less correlated with NH mean temperatures (r = 0.03). With the exception of the models that have trends in PC1, only three other models have significant correlations between PC1 and NH mean temperatures [GFDL, NCAR(CSM), CCSR] and all these correlations are negative. The negative correlation between the observed NAO index and surface warming between the equator and 20°N (especially over Africa) cancels much of the postive correlation with warming over northern Eurasia. The stationary coupled models confirm that NAO does not have a significant positive correlation with the mean surface temperatures averaged over the *whole* of the Northern Hemisphere.

4.3

Correlation of NAO and ENSO

Previous studies have shown that the association between ENSO and the NAO SLP index is weak and difficult to detect in short observational records (Rogers *1984*, Fraedrich *1994*). Table 2 shows that the NAO SLP index has almost no correlation (r = 0.02) with NINO3 temperatures in the tropical Pacific. However, the temperature-based NAO index, PC1, does have a small negative correlation with NINO3 (r = -0.24). Many of the coupled model simulations also exhibit an anticorrelation between PC1 and NINO3 (Table 2). Four of the stationary models have significant negative correlations with NINO3: UKMO, NCAR(CSM), ECHAM4, CCSR. The overly strong link between the NAO and the tropical Pacific in the UKMO model has also been briefly mentioned in Osborn et al. (*1999*).

The correlation maps in Fig. 8 show that ENSO is anticorrelated with an NAO-like quadrupole pattern in many of the models. This is particularly marked in the ECHAM4 and CCSR models suggesting that perhaps these models have an overly strong teleconnection between the tropics and the extra-tropics. The negative correlation between PC1 and NINO3 may also partly explain the negative correlation between PC1 and NH temperatures in several of the models [e.g. NCAR(CSM) and CCSR]. It is interesting that a similar quadrupole pattern in surface temperatures is excited by both NAO and ENSO despite ENSO having only weak correlations with mean sea-level pressure in the North Atlantic region. This teleconnection is perhaps mediated via the Pacific North American pattern, which is known to have an impact on the North Atlantic storm tracks (Ambaum et al. *2001*).

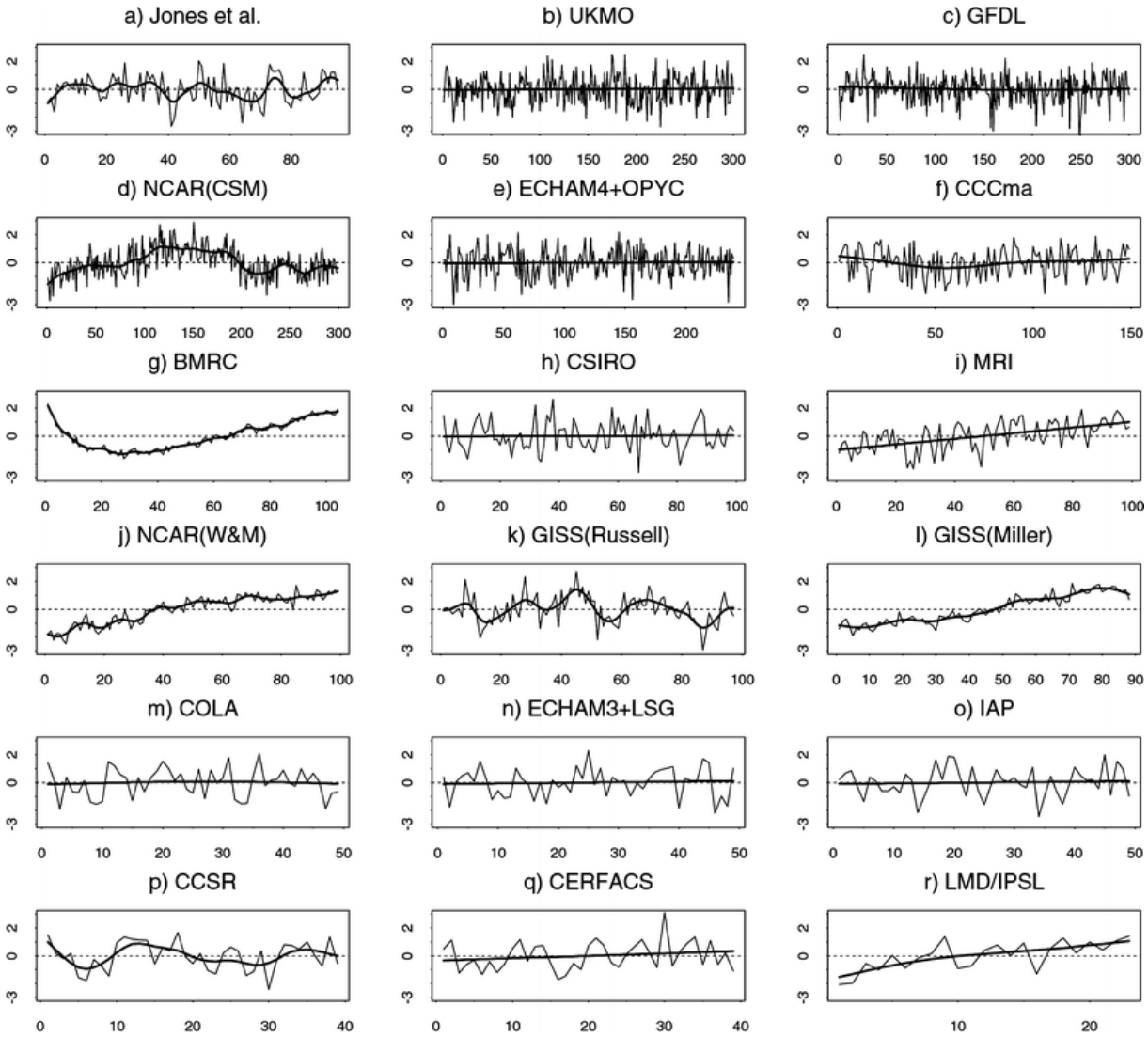

Fig. 8a-c. Correlation maps of surface temperature with the NINO3 index for each model showing the surface temperature teleconnections with ENSO. Contour interval 0.2 with negative correlations *shaded*

4.4

Brief summary

The results presented in this section show that 13 out of the 17 models have a leading temperature EOF that captures the main features of the NAO temperature pattern. Five of the models exhibit monotonic drift in the leading PC time series, but encouragingly four of these models are able to recapture the NAO spatial pattern once detrended. A weakness of many models is that the NAO index is too strongly negatively correlated with ENSO. There is little evidence from either the observations or the models for a correlation between the NAO indices and NH mean temperature.

5

Stochastic modelling of NAO variations

Stochastic modelling provides a useful way of characterising the time-evolution of noisy phenomenon such as the NAO (Wunsch *1999*; Stephenson et al. *2000a*). In this section, a two parameter (parsimonious) time series model will be used to provide a concise summary of the behaviour of the NAO time series from the models. The estimated model parameters are used to quantify both the short-memory and long-memory behaviour of NAO simulated by the different models.

5.1

Power spectra

Figure 9 shows the estimated power spectra of PC1 for the observations and all the models. The spectral power density of PC1 for the observed temperature data (Fig. 9a) is uniform with no significant spectral peaks. This agrees with studies of the NAO SLP index which also show a broad band spectrum close to that of white noise having no statistically significant peaks due to dominant periodicities (Wunsch *1999*; Stephenson et al. 2000). Interestingly, the NAO surface temperature index has less power at low frequencies than does the NAO SLP index there are fewer longterm trends in the temperature data than in the SLP data. With the exception of NCAR(CSM) and GISS(Russell) models, 10 of the 12 stationary models without monotonic trends also have almost flat power spectra with no significant peaks. GFDL is curious in that its power spectra slopes up slightly at high frequency (blue noise). The models with monotonic trends and NCAR(CSM) and GISS(Russell) show spectra with increasing power at low frequency caused by the presence of long-term trends. In the case of NCAR(CSM) and GISS(Russell) this is caused by the long-term trending clearly visible in the time series Fig. 7d and 7k.

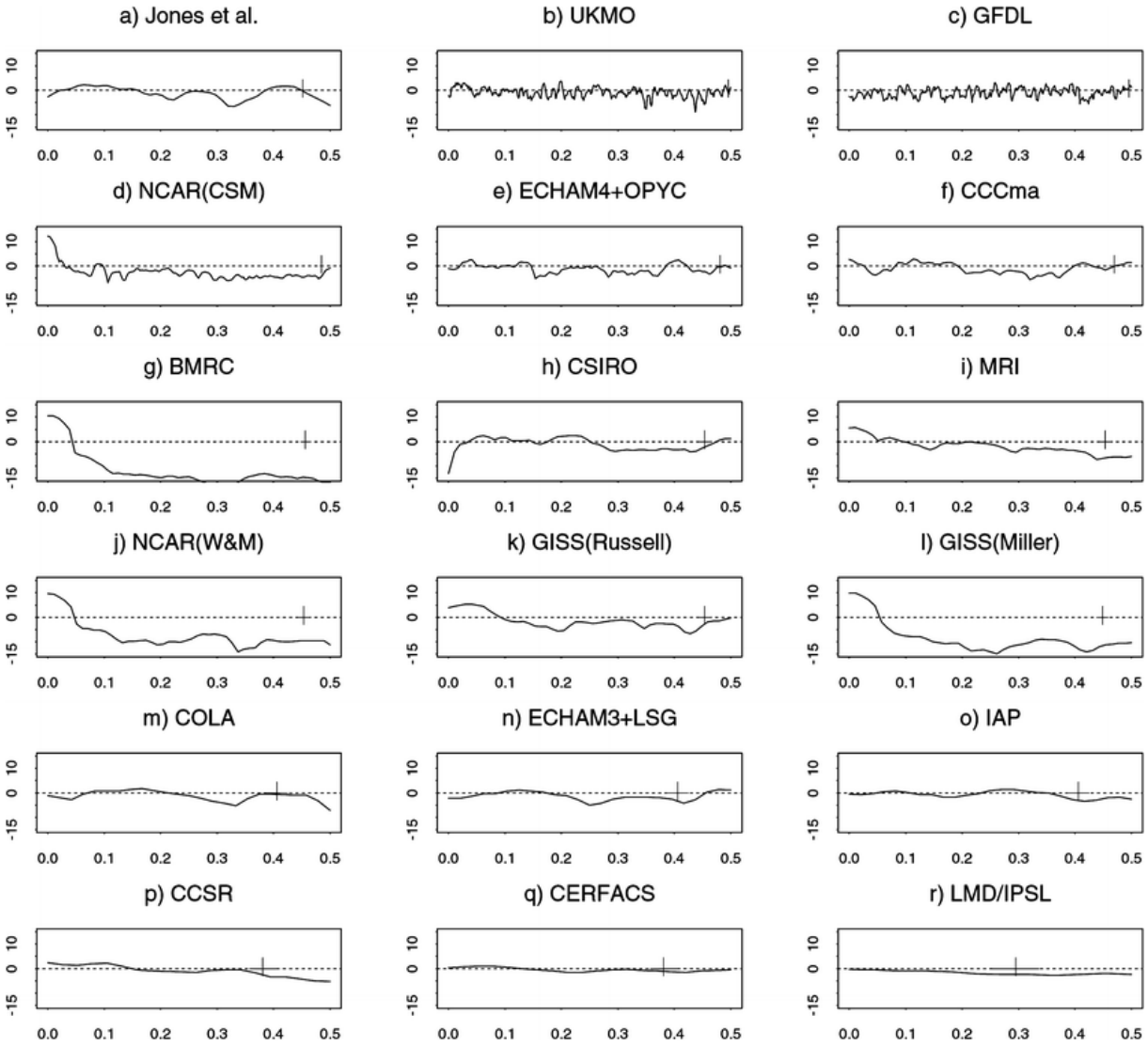

Fig. 9a-c. Estimated power spectra of PC1 for all the models on logarithmic scales. The 95% confidence interval is shown as a *small vertical line*

5.2

The simple AR(1) red noise model

Surface temperature anomalies in mid-latitudes may be considered to be serially correlated noise that results from low-pass filtering of heat fluxes associated with random weather events (Davis *1976*; Hasselmann *1976*). The finite thermal capacity of the surface layer generates serial correlations in the surface temperature anomalies. This continuous time stochastic model is not directly applicable to discrete time series such as winter time means, yet a suitable analogue having the same low frequency limit is provided by the first order AR(1) autoregressive model

$$z(y) = r_1 z(y-1) + a(y) \qquad (1)$$

where $z(y)$ is the winter time anomaly for year y, r_1 is the lag-1 year autocorrelation, and $a(y)$ is a Gaussian distributed shock/innovation (Box and Jenkins *1976*). This simple "red noise" time series model is widely used to model serially correlated discrete time series, and has proved to be useful in understanding and forecasting model generated SST variability (Saravanan and McWilliams *1997*; Griffies and Bryan *1997*).

5.3

Modelling short- and long-range dependence

Stephenson et al. (*2000a*) showed that the simple AR(1) model was too restrictive to describe all the features of the observed NAO SLP index. In particular, because of its exponentially decaying autocorrelation function, the AR(1) model predicts that the variance of means of length n should scale as n^{-1} for sufficiently long means (Beran *1994*). This inverse scaling relationship leads to the well-known expression $\sigma/\sqrt{n}$ for the standard error of the mean of n independent variables.

However, Stephenson et al. (*2000a*) showed that the variance of means of the NAO SLP index actually decays more slowly as $n^{-0.72}$. For instance, decadal means of the NAO SLP index from 1864-1998 have almost twice ($10^{-0.72}/10^{-1}$) the variance expected for an AR(1) process. The NAO SLP index has more trending on long time scales than is to be expected for a simple AR(1) process. This increased amount of natural variability needs to be considered when testing for climate change trends in the NAO.

Similar long-term persistence or "long memory" has been found to exist in many other naturally occurring time series, for example, river flow rates (Hurst *1951*), global mean temperatures (Bloomfield *1992*), and wind power variations (Haslett and Raftery *1989*). Time series with long memory are noisy yet appear to meander for long times away from the longterm mean as can be seen in PC1 of NCAR(CSM) in Fig. 7d. Unlike short memory processes such as AR(1), which flatten off at low frequency, long memory processes have increasing amounts of variance at low frequencies (spectral peak at zero frequency). Long memory processes forget only slowly their past behaviour, the autocorrelation function shows slow power law decay in lag rather than fast exponential decay which occurs for "short-memory" processes such as red noise and other ARMA processes (Beran *1994*). The amount of long memory is characterised by the dimensionless scaling exponent H first discovered by Hurst (*1951*) in his comprehensive study of river flows.

The simplest time series model that can capture both short-range and long-range dependence is a fractionally differenced autoregressive order 1, FAR(1) model:

$$(1-B)^d(1-r_1B)z(y) = a(y) \qquad (2)$$

where B is the backward shift operator $Bz(y) = z(y\text{-}1)$ and r_1 and $d = H - 0.5$ are model parameters summarising short and long memory, respectively. For the special case $H = 0.5$, one obtains the short-memory AR(1) model that gives variances of means that scale as n^{-1} where n is the length of the means. For larger values of the Hurst exponent, the variance of means falls off less quickly as n^{2H-2} due to the presence of more long-term stochastic trends. For Hurst exponents greater than unity the process becomes non-stationary with unbounded variance.

Table 4 gives maximum likelihood estimates of r_1 and *H* obtained by fitting the PC1 time series to the FAR(1) model. For the 12 models that do not have significant monotonic trends, the estimated lag-1 autocorrelation coefficient is generally small with values ranging from -0.05 to 0.39. This evidence provides a coupled model confirmation that the NAO spectrum is only "weakly red" (i.e. "pink noise"). With the exception of GFDL which is slightly blue, all the models have positive lag-1 autocorrelations, suggesting that the NAO temperature index may have some slight serial correlation between winters. This ("redness") can be exploited to give some skill in statistically forecasting the NAO one year ahead. However, such forecasts will at most be able to explain only a very small fraction (r_1^2) of the total NAO variance (Box and Jenkins *1976*; Wunsch *1999*).

Table 4. Some estimated characteristics of the observed and simulated PC1 time series of length NYRS

Model	NYRS	Lag-1 r_1	Hurst H
Obs. 1900-94	95	+0.17 ± 0.10	+0.52
UKMO	1085	+0.09 ± 0.03	+0.51
GFDL	1000	-0.05 ± 0.03	+0.45
NCAR(CSM)	300	+0.51 ± 0.05	+0.85
ECHAM4	240	+0.05 ± 0.06	+0.51
CCCma	150	+0.12 ± 0.08	+0.56
BMRC	105	+0.95 ± 0.03	+0.94
CSIRO	100	+0.17 ± 0.10	+0.44
MRI	100	+0.50 ± 0.09	+0.77
NCAR(W&M)	100	+0.87 ± 0.05	+0.92
GISS(Russell)	98	+0.39 ± 0.09	+0.72
GISS(Miller)	89	+0.90 ± 0.05	+0.95
COLA	50	+0.15 ± 0.14	+0.47
ECHAM3	50	+0.03 ± 0.14	+0.45
IAP/LASG	50	+0.07 ± 0.14	+0.51
CCSR	40	+0.25 ± 0.16	+0.64
CERFACS	40	+0.07 ± 0.16	+0.58
LMD/IPSL	24	+0.34 ± 0.20	+0.59

The lag-1 autocorrelation coefficient r_1 and its standard error σ_{r_1} are obtined by maximum likelihood fits of the time series to the AR(1) red noise model. The Hurst exponent, *H*, measures the amount of "long-memory" in the time series and has been estimated using log-log fits to aggregated variance (see text for details)

5.4

Brief summary

Figure 10 uses the estimated parameters of the FAR(1) model to summarise the short and long memory behaviour of the PC1 time series. With the exception of the models that drift, most of the models are clustered close to the point (0.0, 0.5) confirming that interannual variability in the NAO has a broad band spectrum which closely resembles white noise. However, it can be seen that all but one of the stationary models are to the right of the line $r_1 = 0$ implying some serial correlation between successive winters (shortterm memory). All of the models apart from NCAR(CSM), NCAR(W&M), GISS(Russell), and GISS(Miller) have Hurst exponents close to 0.5 signifying that they have no long-memory in temperature-based NAO indices. For example, the variance of means of these time series all scale as n^{-1} (not shown). The long-memory in GISS(Miller) and NCAR(W&M) is due to the presence of significant monotonic trends caused by coupled model drift. However, NCAR(CSM) and GISS(Russell) do not have significant monotonic trends yet still exhibit significant amounts of long-memory behaviour. This is caused by the presence of long-term stochastic trends that can be seen in the time series shown in Fig. 10. This intriguing non AR(1) behaviour merits further attention in these particular model simulations.

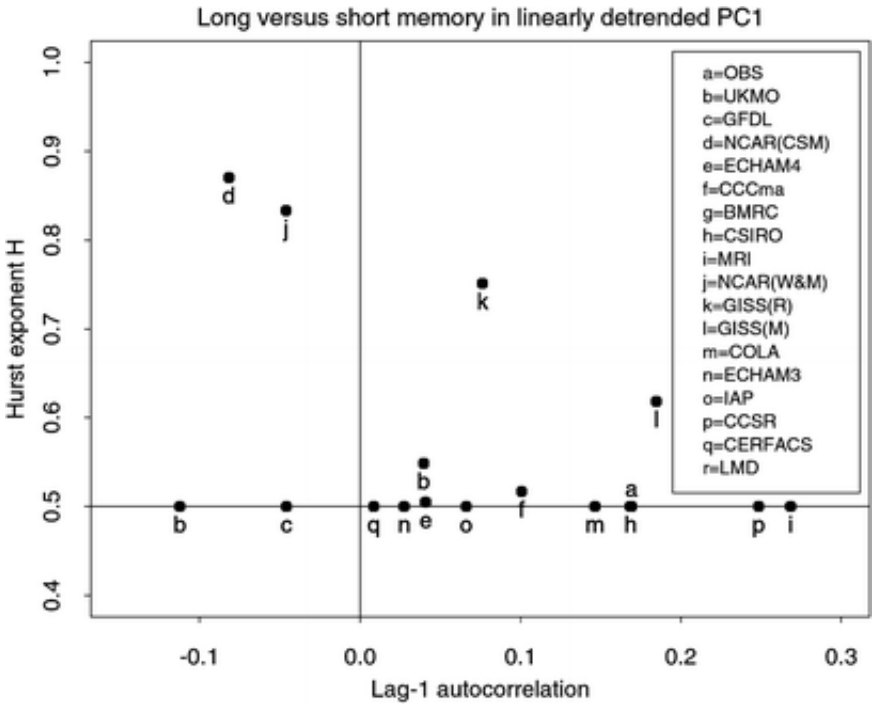

Fig. 10a-c. Scatter plot of long versus short memory summarising the behaviour of PC1 for all the models. The abscissa is the estimated lag-1 autocorrelation r_1 and the ordinate is the dimensionless Hurst exponent *H*. Short memory processes lie close to the line $H = 0.5$. Note that many of the models are close to the point $r_1 = 0$ and $H = 0.5$, which represents completely uncorrelated white noise

---

6

Concluding remarks

This preliminary coupled model intercomparison study has found that:

1. There is a wide diversity of model behaviour in simulating total and short-term interannual variance. Patterns resemble one another but a few models are much "redder" over the central North Atlantic ocean.

2. Encouragingly 13 out of 17 models appear to have a leading EOF that captures the main features of the observed NAO quadrupole pattern such as the "northern dipole"

3. The southern dipole is less well simulated and is often displaced too far east over the Atlantic ocean.

4. Five of the 17 models have monotonic trends in surface temperature caused by coupled model drift towards warmer surface conditions.

5. The 12 other "stationary" models give NAO indices that are only slightly serially correlated ("weakly red") in agreement with what is found for the observations.

6. Two of the stationary models have significantly more longterm variability (trends) in the NAO index than can be expected from a short memory process.

7. There is little evidence of a significant correlation between the NAO index and Northern Hemisphere average temperatures in either the observations or in the results from the non-drifting models.

8. There is weak anti-correlation between the NAO temperature index and the NINO3 index in the observations and in several of the model results. Some models overestimate this weak teleconnection.

Paragraph 8.7.5 in IPCC (*2001*) claimed that "coupled climate models simulate the NAO quite well, although there are some differences in amplitude". However, our findings show that the NAO time series in many of the coupled models are unrealistic due mainly to monotonic trends and/or overly strong correlations with ENSO:

1. Category 1 models with monotonic trends, namely the five models BMRC, MRI, NCAR(W&M), GISS(Miller), and LMD/IPSL;

2. Category 2 models with overly strong correlations with ENSO, namely the six models BMRC, MRI, UKMO, NCAR(CSM), ECHAM4, and CCSR;

3. Category 3 models, namely the eight other models that do not have monotonic trends or overly strong ENSO teleconnections: GFDL, CCCma, CSIRO, GISS(Russell), ECHAM3, IAP/LASG, and CERFACS/IPSL.

Apart from the MRI model, all of the other category 1 models are not flux-adjusted. Categories 2 and 3 have slightly more flux-adjusted models than non-adjusted models, but the difference is not striking. The CCCma, GISS(Russell), and GFDL models are all in category 3, thereby supporting the IPCC (*2001*) claim that "a realistic AO is simulated in CCCma, GISS, and GFDL models".

As a continuation of this project, we are now assessing the behaviour of the NAO in both control coupled runs and also in climate change coupled runs having doubled carbon dioxide. Data for this is provided by the second phase, CMIP2, which includes control simulations and also transient simulations having 1% per year increase in atmospheric carbon dioxide concentration. CMIP2 provides monthly means of not only surface temperature but also sea-level pressure and precipitation, which allows us to investigate more thoroughly the NAO and its teleconnections simulated by the coupled models.

Acknowledgements. This work would not have been possible without the kind and able assistance of Curt Covey at PCMDI who made every effort to help us in this endeavour. We also would like to thank Brian Hoskins, Jim Hurrell, Larry Gates, Martina Junge, Tom Phillips, Uwe Ulbrich and two reviewers for their most useful remarks. The authors were supported during this study as part of the Storm Track Ocean interaction and European Climate (STOEC) project funded by European Commission grants ENV4-CT97-0499. Visits of DBS to CINECA were supported by the European Union ICARUS project. More information about the STOEC project can be found at http://www.met.rdg.ac.uk/cag/STOEC/.

References

Abe-Ouchi A et al. (1996) Outline of CCSR coupled atmosphere and ocean model and experiment. Internal Report, Center for Climate System Research, University of Tokyo, Japan

Ambaum MHP, Hoskins BJ, Stephenson DB (2001) Arctic Oscillation or North Atlantic Oscillation? J Clim 14: 3495-3507

Ambaum MHP, Hoskins BJ, Stephenson DB (2002) Arctic oscillation or North Atlantic oscillation? (vol 14, pg 3495, 2001). J Clim 15(5): 553-553

Beran J (1994) Statistics for long-memory processes. Chapman and Hall, pp 315

Bjerknes J (1962) Synoptic survey of the interaction of sea and atmosphere in the North Atlantic. Geophysica Norvegica 24(3): 115-145

Bjerknes J (1964) Atlantic air-sea interaction. Adv Geophys 10: 1-82

Bloomfield P (1992) Trends in global temperature. Clim Change 21: 1-16

Boville BA, Gent PR (1998) The NCAR Climate System Model, Version One. J Clim 11: 1115-1130

Box GEP, Jenkins GM (1976) Time Series Analysis - forecasting and control. Holden-Day, California, pp 575

Branconnot P, Marti O, Joussaume S (1997) Adjustment and feedbacks in a global coupled ocean-atmosphere model. Clim Dyn 13: 507-519

Covey CC, Abe-Ouchi AA, Boer GJ, Boville BA, Cubasch U, Fairhead L, Flato GM, Gordon H, Guilyardi E, Jiang X, Johns TC, LeTreut H, Madec G, Meehl GA, Miller R, Noda A, Power SB, Roeckner E, Russell G, Schneider EK, Stouffer RJ, Terray L, von Storch JS (2000) The seasonal cycle in coupled ocean-atmosphere general circulation models. Clim Dyn 16(10-11): 775-787

Davis RE (1976) Predictability of sea surface temperature and sea-level pressure anomalies over the North Pacific ocean. J Phys Oceanogr 6: 249-266

Flato GM, Boer GJ, Lee W, McFarlane N, Ramsden D, Weaver A (2000) The CCCma global coupled model and its climate. Clim Dyn 16: 451-467

Fraedrich K (1994) An ENSO impact on Europe -a review. Tellus 46(4): 541-552

Garric G, Stephenson DB, Terray L (1997) A global coupled atmosphere sea ice upper ocean simulation. Com Ren Acad Sci 324(7): 529-536

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

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

Griffies SM, Bryan K (1997) A predictability study of simulated North Atlantic multidecadal variability. Clim Dyn 13: 459-487

Guilyardi E, Madec G (1997) Performance of the OPA/ARPEGE-T21 global ocean-atmosphere coupled model. Clim Dyn 3: 149-165

Haslett J, Raftery AE (1989) Space-time modelling with long-memory dependence: assessing Ireland's wind power resource. Appl Stat 38(1): 1-50

Hasselmann K (1976) Stochastic climate models. Part I. Theory. Tellus 28: 473-484

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

Hurrell JW (1996) Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. Geophys Res Lett 6: 665-668

Hurst HE (1951) Long-term storage capacity of reservoirs. Trans Am Soc Civil Eng 116: 770-779

IPCC (2001) Climate change 2001: the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) Cambridge University Press, Cambridge, UK

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, Raper SCB, Bradley RS, Diaz HF, Kelly PM, Wigley TML (1986) Northern Hemisphere surface air temperature variations, 1851-1984. J Clim Appl Meteorol 25: 161-179

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

Lamb HH (1972) Climate: present, past and future, Vol I, fundamentals and climate now. Methuen, UK, pp 613

Manabe S, Stouffer RJ (1996) Low-frequency variability of surface air temperature in a 1000-year integration of a coupled atmosphere-ocean-land surface model. J Clim 9: 376-393

Meehl GA (1995) Global coupled general circulation models. Bull Am Meteorol Soc 76: 951-957

Miller RL, Jiang X (1996) Surface energy fluxes and coupled variability in the tropics of a coupled general circulation model. J Clim 9: 1599-1620

Osborn TJ, Briffa KR, Tett SFB, Jones PD, Trigo RM (1999) Evaluation of the Northern Atlantic Oscillation as simulated by a coupled climate model. Clim Dyn 15(9): 685-702

Paeth H, Hense A, Glowienka-Hense R, Voss R, Cubasch U (1999) The North Atlantic Oscillation as an indicator for greenhouse-gas induced regional climate change. Clim Dyn 15(12): 953-960

Pittalwala II, Hameed S (1991) Simulation of the North-Atlantic Oscillation in a general-circulation model. Tellus 18(5): 841-844

Power SB, Colman RA, McAvaney BJ, Dahni RR, Moore AM, Smith NR (1993) The BMRC Coupled atmosphere/ocean/sea-ice model. BMRC Res Rep 37. Bureau of Meteorology Research Centre, Melbourne, Australia, pp 58

Roeckner E, Oberhuber JM, Bacher A, Christoph M, Kirchner I (1996) ENSO variability and atmospheric response in a global coupled atmosphere-ocean GCM. Clim Dyn 12: 737-754

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

Russell GL, Miller JR, Rind D (1995) A coupled atmosphere-ocean model for transient climate change studies. Atmos Ocean 33: 683-730

Saravanan R, McWilliams JC (1997) Stochasticity and spatial resonance in interdecadal climate fluctuations. J Clim 10: 2299-2320

Schneider EK, Zhu Z, Giese BS, Huang B, Kirtman BP, Shukla J, Carton JA (1997) Annual cycle and ENSO in a coupled ocean-atmosphere general circulation model. Mon Weather Rev 125: 680-702

Stephenson DB (2002) The impact of ocean-atmosphere coupling on mid-latitude storm tracks and blocking. UGAMP Techn Rep, Reading, UK

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

Stephenson DB, Wanner H, Brönnimann S, Luterbacher J (2002b) The History of Scientific Research on the North Atlantic Oscillation. In: Hurrell J, Visbeck M, Kushnir Y, Ottersen G (eds) Chapter 2 in the AGU monograph on the North Atlantic Oscillation. American Geophysical Union, USA

Tokioka T, Noda A, Kitoh A, Nikaidou Y, Nakagawa S, Motoi T, Yukimoto S, Takata K (1996) A transient CO_2 experiment with the MRI CGCM: annual mean response, volume 2. CGER's Supercomputer Monograph Report, CGER-IO22-96, ISSN 1341-4356, Center for Global Environmental Research, National Institute for Environmental Studies, Environment Agency of Japan, Ibaraki, Japan, pp 86

Trenberth KE (1990) Recent observed interdecadal climate changes in the Northern Hemisphere. Bull Am Meteorol Soc 71: 988-993

Ulbrich U, Christoph M (1999) A shift of the NAO and increasing storm track activity over Europe due to anthropogenic greenhouse gas forcing. J Clim 15(7): 551-559

van Loon H, Rogers JC (1978) The seesaw in winter temperatures between Greenland and Northern Europe. Part I: general description. Mon Weather Rev 106: 296-310

Voss R, Sausen R, Cubasch U (1998) Periodically synchronously coupled integrations with the atmosphere-ocean general circulation model ECHAM3/LSG. Clim Dyn 14(4): 249-266

Walker GT, Bliss EW (1932) World weather V. Mem Roy Meteorol Soc 4: 53-84

Wallace JM, Zhang Y, Bajuk L (1996) Interpretation of interdecadal trends in Northern Hemisphere surface air temperature. J Clim 9: 249-259

Wanner H, Bronnimann S, Casty C, Gyalistras D, Luterbacher J, Schmutz C, Stephenson DB, Xoplaki E (2001) North Atlantic Oscillation - concepts and studies. Surv Geophys 22(4): 321-382

Washington WM, Meehl GA (1996) High-latitude climate change in a global coupled ocean-atmosphere-sea ice model with increased atmospheric CO_2. J Geophys Res 101(D8): 12,795-12,801

Wunsch C (1999) The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillation. Bull Am Meteorol Soc 80: 245-255

Zhang X-H, Shi G, Liu H, Yu Y-Q (2000) IAP global ocean-atmosphere-land system model. Science Press, New York, pp 3-100

Zhou T, Zhang X-H, Yu Y-Q, Yu R, Wang S (2000) The North Atlantic Oscillation simulated by versions 2 and 4 of IAP/LASG GOALS Model. Adv Atmos Sci 17(4): 601-616