

North Atlantic forcing of climate and its uncertainty from a multi-model experiment

By M. J. RODWELL^{1*}, M. DRÉVILLON², C. FRANKIGNOUL³, J. W. HURRELL⁴,
H. POHLMANN⁵, M. STENDEL⁶ and R. T. SUTTON⁷

¹*Met Office, Exeter, UK*

²*Centre Europeen de Recherche et de Formation Avancée en Calcul Scientifique, Toulouse, France*

³*Université Pierre et Marie Curie, Paris, France*

⁴*National Center for Atmospheric Research, Boulder, USA*

⁵*Max-Planck-Institut für Meteorologie, Hamburg, Germany*

⁶*Danish Meteorological Institute, Copenhagen, Denmark*

⁷*Centre for Global Atmospheric Modelling, University of Reading, UK*

(Received 17 November 2003; revised 11 February 2004)

SUMMARY

To understand recent climate change in the North Atlantic region and to produce better climate forecasts with uncertainty estimates it is important to determine the atmospheric ‘response’ to Atlantic sea-surface temperature (SST) forcing. There have been conflicting results regarding the strength, character and tropical-versus-extratropical origin of this response. For model-based studies, this may indicate differing sensitivities to Atlantic SST, but the comparison is complicated by changes in experimental design. Here, a highly controlled experiment with five atmospheric models is undertaken. The influence of realistic (if reasonably strong) and optimally chosen North Atlantic (equator to 70°N) SST anomalies is isolated. Unexpected global agreement between the models is found (e.g. the North Atlantic Oscillation (NAO), Eurasian temperatures, rainfall over the Americas and Africa, and the Asian monsoon). The extratropical North Atlantic region response appears to be associated with remote Caribbean and tropical Atlantic SST anomalies, and with local forcing. Some features such as the European winter-temperature response would be stronger than atmospheric ‘noise’ if the prescribed SST anomalies persisted for just two years. More generally, Atlantic air–sea interaction appears to be important for climate variability on the 30-year timescale and, thus, to be important in the climate-change context. The multi-model mean response patterns are in reasonable agreement with observational estimates, although the model response magnitudes may be too weak. The similarity between their responses helps to reconcile models. Inter-model differences do still exist and these are discussed and quantified.

KEYWORDS: Air–sea interaction Asian monsoon Intertropical convergence zone North Atlantic Oscillation

1. INTRODUCTION

At intraseasonal to interannual timescales it is well known that it is the transients in the North Atlantic atmosphere that drive much of the variability in sea-surface temperature (SST) (see, for example, Cayan (1992)). Nevertheless it is possible that two-way air–sea interaction could play a significant role in North Atlantic climate variability, particularly at longer timescales (Bjerknes 1964). Here, one aspect of this coupling is considered: that of the atmospheric response to North Atlantic SST forcing and the timescale dependence of the importance of this response for climate variability.

It is widely recognized that tropical North Atlantic SSTs can affect tropical deep convection and upper-tropospheric vorticity forcing which, through the action of barotropic Rossby waves can have a teleconnective impact on the climate of the North Atlantic (Hoskins and Karoly 1981). Rowntree (1976), Okumura *et al.* (2001), Terray and Cassou (2002) and many others have investigated the extratropical impact of such tropical Atlantic SST. Nevertheless, it is unclear how well current models represent the processes and interactions involved in this response: the generation of tropical SST

* Corresponding author, present affiliation: European Centre for Medium-Range Weather Forecasts, Reading, Berkshire RG2 9AX, UK.

anomalies (SSTAs), evaporation, convection etc. The role played by extratropical North Atlantic SST in forcing North Atlantic climate is probably less clear as there is little or no model consensus at present (see Kushnir *et al.* (2002) for a summary). The locally forced response can be thought of as the combination of a local baroclinic response to extratropical SST (e.g. Kushnir 1994) and the interaction between this response and the North Atlantic storm-track which produces a barotropic signal (e.g. Palmer and Sun 1985). Peng and Whitaker (1999) and Hall *et al.* (2001) suggested that this interaction is highly sensitive to the position of the storm-track and this may be one reason for the model discrepancies. The relatively low number of realizations and the high level of atmospheric internal variability may also help explain the differences. The combined response to tropical and extratropical SST is even more complicated. For example, Lau and Nath (1996) suggested that mid-latitude ocean–atmosphere coupling could enhance the amplitude and persistence of a tropically forced response.

The ‘analysis of variance’ (ANOVA, see, for example, Davis *et al.* (1997)) of ensembles of ‘AMIP*-style’ (Gates 1992) atmospheric general-circulation model (AGCM) simulations forced with observed SST has been used elsewhere to estimate the fraction of total atmospheric variance that can be attributed to SST forcing. ANOVA successfully highlights regions that are generally susceptible to forcing by SST in model simulations and is also useful for model intercomparison. The clear picture that emerges is that a high proportion of tropical atmospheric interannual variability in these models is forced by the prescribed SST. For December–February (DJF) mean-sea-level pressure (MSLP), over 70% of the tropical Atlantic interannual variability can be explained by SST forcing. In the extratropics, the value is much smaller (often less than 20% of total interannual variability). It could be argued that such a small percentage makes it fruitless to investigate the extratropical response to SST forcing. However, in view of the redness of the SSTA spectrum, the percentage should be larger at longer timescales. ANOVA can be extended to the frequency domain to assess the proportion of variance explained by SST forcing at longer timescales (Rowell and Zwiers 1999). The percentage of variance explained at decadal timescales is generally higher than at interannual timescales, but results are less robust across experiments. For example the French ARPEGE3 AGCM shows 40% of the decadal June–August (JJA) MSLP variance near Iceland to be explained by SST forcing (1947–1998) (Laurent Terray, personal communication) while the UK HadAM3 model shows less than 20%. The German ECHAM4 model gives an intermediate value. Using an earlier version of HadAM3, (HadAM1), Rowell and Zwiers (1999) showed that no significant decadal DJF variability was explained by SST forcing over the extratropical North Atlantic for the period 1949–1993 whereas, for the same period, HadAM3 suggests that over 40% of decadal variability as far north as Iceland can be explained by SST forcing. It is unclear how much of these differences can be attributed to model differences, as the relatively short length of the observed SST record puts a limit on achievable significance. Another limitation is that ANOVA does not indicate the relative importance of different oceanic regions (tropical Pacific, tropical Atlantic, extratropical Atlantic etc.) in the overall forcing. In addition, it is possible that particular patterns of SSTAs will have a disproportionately strong impact in certain regions or on certain ‘modes’ of atmospheric variability, but that this is obscured by the mean statistic given. For example, Rodwell *et al.* (1999) highlighted the response to a tripole pattern in North Atlantic SSTAs. This was seen in idealized experiments with fixed tripole SSTAs and in a regression analysis between the simulated North Atlantic Oscillation (NAO, Walker and Bliss 1932) and prescribed observed SST in

* Atmospheric Model Intercomparison Project.

a six-member ensemble of atmospheric model experiments. Their quoted correlation between observed and hindcast winter NAO index, 0.41, was confirmed by Mehta *et al.* (2000), but Mehta *et al.*'s results suggested that a six-member ensemble may be too small to get a robust result.

Barsugli and Battisti (1998) developed a one-dimensional energy-balance model of atmosphere–ocean coupling. The model, which is widely credited with capturing some of the essential features of air–sea interaction, appears to throw doubt on the meaningfulness of using AMIP-style AGCM simulations to study atmospheric predictability. For example, Bretherton and Battisti (2000) noted that correlations between observed variations and those of AMIP-style simulations may be exaggerated. This is because the AGCM will adjust to the observed SSTAs that were, in reality, partly forced by unpredictable atmospheric anomalies, and ensemble averaging can inflate the correlations by reducing the simulated atmospheric ‘noise’. These results suggest some inaccuracy in the ‘percentage of decadal variance explained by SST forcing’, if such a quantity is actually meaningful. However, the model of Barsugli and Battisti (1998) is intentionally simple and it does omit potentially important feedback mechanisms such as temperature advection by Ekman currents (e.g. Bjerknes 1964; Palmer and Sun 1985; Rodwell *et al.* 1999) and the effects of lower frequency variations in the ocean circulation. Adding quasi-geostrophic dynamics to the model, Ferreira *et al.* (2001) showed that the coupling to oceanic Rossby waves could lead to weakly unstable modes and a small climate predictability up to six years in advance. Using a more comprehensive coupled ocean–atmosphere general-circulation model (OAGCM), Collins (2002) did find decadal predictability of surface air-temperature anomalies over the North Atlantic. A working assumption in the study by Collins is that the model is perfect. Whether the model does capture well enough features such as the variability of the thermohaline circulation, its relationship with SST and the atmospheric response to SST forcing (represented by the parameter ‘b’ in the model of Barsugli and Battisti 1998) is not straightforward to validate against observations.

One approach to investigate air–sea interaction in the observations and to validate these interactions in OAGCMs is to use a lagged maximal-covariance analysis technique. Using such techniques, Czaja and Frankignoul (1999, 2002), Rodwell and Folland (2002, 2003) and Rodwell (2003) have been able to identify statistically significant estimates of the observational responses to SST forcing, although the shortness of the instrumental record, atmospheric internal variability and autocorrelation, and external forcing factors such as the El-Niño–Southern Oscillation can complicate a simple interpretation. Rodwell and Folland (2003) and Rodwell (2003) applied the same technique to an ensemble of HadAM3 simulations and to a long ~ 1500 -year simulation of an OAGCM (HadCM3, which includes HadAM3 as its atmospheric component). Results suggest that the atmospheric model does respond to SST forcing with the correct (500 hPa geopotential height) patterns, but too weakly.

Here, the aim is to address some of the questions raised above concerning the response to SST forcing, while avoiding the issues concerning predictability. A multi-model experiment is conducted where prime consideration is given to producing robust, physically justifiable and statistically significant results. Specifically, the aim is to (1) produce a best multi-model estimate of the atmospheric response to ‘optimally chosen’ seasonally fixed SSTAs, (2) investigate the physics and origin of the response, (3) look for strong and coherent local anomalies (that may not have been highlighted by the more generalized ANOVA technique), (4) compare the response with observationally based estimates for validation purposes, (5) estimate the importance of the response in climate variability over a range of timescales and (6) examine inter-model differences.

TABLE 1. DETAILS OF ATMOSPHERIC MODELS USED IN THE STUDY

Model	Origin	Type	Resolution		Reference	
			Horizontal	Vertical		
<i>H3</i>	HadAM3	UK	Grid-point	2.5° lat × 3.75° long	19 levels	Pope <i>et al.</i> (2000)
<i>A3</i>	ARPEGE3	France	Spectral	T63 (2.8° × 2.8°)	31 levels	Déqué <i>et al.</i> (1994)
<i>E4</i>	ECHAM4	Germany	Spectral	T42 (2.8° × 2.8°)	19 levels	Roeckner <i>et al.</i> (1996)
<i>E5</i>	ECHAM5	Germany	Spectral	T42 (2.8° × 2.8°)	19 levels	Roeckner <i>et al.</i> (2003)
<i>C2</i>	CAM2	USA	Spectral	T42 (2.8° × 2.8°)	26 levels	

Physics terms are calculated on a T42 Gaussian grid in ARPEGE3. ECHAM5 simulations were run at the Danish Meteorological Institute.

In section 2, the models are summarized and experimental details, including the methodology used to optimize the SSTAs, are described. The individual model responses are compared in section 3. Section 4 gives the multi-model mean response, compares it with atmospheric internal variability and discusses the likely SSTs (tropical versus extratropical) responsible for the response. Section 5 compares the mean model response with observational estimates. Section 6 discusses inter-model differences. Section 7 summarizes the response, its uncertainty and significance. Section 8 demonstrates linearity in the multi-model response. Assuming linearity, the timescales over which North Atlantic air–sea interaction is important for natural climate variability are estimated in section 9. Further discussion and the conclusions are given in section 10.

2. MODELS AND METHODOLOGY

(a) *Models*

The five different atmospheric models investigated are detailed in Table 1. All the models have previously been well tested and show reasonable climates. There are mean biases from observational reanalysis estimates, although these are thought to have only a secondary effect on the sensitivities of interest here. Typical maximal values include a +4 hPa summer MSLP bias in model *A3* over the central North Atlantic and a similar bias in the winter Azores high in *E4*. *H3* shows a 30% deficit in North Atlantic winter blocking frequency (Pope *et al.* 2000). For each model, the NAO is well represented by the first empirical orthogonal eigenfunction (EOF1) of North Atlantic region winter MSLP. For example, the DJF EOF1 in an ensemble of six *H3* simulations forced with observed SST between 1946 and 1998 explains 42%, 31%, 36%, 45%, 39% and 42% of the total variance, respectively. The corresponding observational value is 43%. The observed strong anticorrelation between the Azores high and the Icelandic low in DJF ($r \approx -0.69$) is well captured in the models tested (*H3*, *E4*, *E5*). For example, four ensemble members of *E5* give values of -0.63 , -0.65 , -0.70 and -0.55 , respectively. The same three models also well capture the relationships between the NAO and European precipitation and surface temperatures (Walker and Bliss 1932; Hurrell 1995; Rodwell *et al.* 1999). There is no reason to believe that the other two models behave significantly differently in these respects.

(b) *Simulations of each model*

For each model, a 21-year control simulation *C* was forced with the 1948–1998 climatological mean annual cycle in SST and sea ice (and the same pre-industrial CO₂ and trace gases). Simulations *A*⁺ and *A*[−] were made with SSTAs (see below) added to, and subtracted from, the climatological SSTs in the North Atlantic. All integrations were started at the beginning of February. The first ten months were disregarded. Analysis was

made on the 20 years of seasonal-mean data for each standard season (DJF, MAM, JJA and SON where, as usual, letters correspond to the first letter of each month). Rodwell and Ingram (2000) showed, for an earlier version of the Met Office model (HadAM2b), that 20 years of data are adequate to identify field-significant responses in the North Atlantic region.

(c) Construction of SST anomalies

The aim is to define SSTA patterns that are realistic in structure and magnitude and, if possible, ‘optimal’ in the sense that they are thought to have the strongest impact on the North Atlantic climate. To do this, SSTAs are based on the observational lagged maximum covariance analysis (MCA) of Rodwell (2003) applied to monthly-mean SST and the subsequent seasonal-mean 500 hPa geopotential height (Z500). The procedure, which can be skipped by the reader if desired, is detailed in the next paragraph.

Monthly-mean SST data, $(SST_{m,y})$ where m is the month and y is the year, are taken from the HadISST1 dataset (Rayner *et al.* 2003) in the box (90°W – 15°E , 0 – 70°N) for 1948 to 1998. Three-month-mean Z500 data, $(Z500_{m,y})$ where m is the central month of the ‘season’, are taken from the National Centers for Environmental Prediction (NCEP) reanalyses (Kalnay *et al.* 1996) in the box (90°W – 45°E , 10°N – 80°N) for the same period. For each month m the lagged MCA is applied to $(SST_{m-2,y}, Z500_{m,y})_{y \in \{1948, \dots, 1998\}}$ to obtain the first pair of anomaly covariance patterns $(SSTA_{m-2}, Z500A_m)$. As with Czaja and Frankignoul (1999), statistical significance is based on a Monte Carlo test of the squared covariance. Eight out of the 12 three-month running seasons are found to be significant at the 10% level. For the significant patterns, the use of a lag suggests that an ocean-forced response is obtained and additional tests (Rodwell and Folland 2002) tend to confirm this interpretation. The SST anomaly $SSTA_{m-2}$ is thought to affect the atmosphere by persisting through the months $m - 1, m, m + 1$ and forcing the Z500 field quasi-instantaneously (at a monthly temporal resolution). Rodwell and Folland (2002) show SSTA-pattern autocorrelations of 0.8–0.9 between November and the months of DJF. $SSTA_{m-2}$ could have been used as the optimal forcing SSTA for month m but, since not all $SSTA_{m-2}$ patterns are statistically significant, a simplification using just the significant MCA patterns was sought. It was found from pattern correlations that there is a reasonably clear separation between the SST patterns corresponding to $m \in \{S, O, N, D, J, F\}$ (with three patterns statistically significant) and $m \in \{M, A, M, J, J, A\}$ (with five patterns significant). Hence an Autumn/Winter pattern, $SSTA_{\text{SONDJF}}$, and a Spring/Summer pattern, $SSTA_{\text{MAMJJA}}$, were defined as the means of the three and five statistically-significant MCA patterns, respectively. For each month m the original SST data, $(SST_{m,y})_{y \in \{1948, \dots, 1998\}}$, is projected onto either $SSTA_{\text{SONDJF}}$ if $m \in \{S, O, N, D, J, F\}$, or $SSTA_{\text{MAMJJA}}$ if $m \in \{M, A, M, J, J, A\}$, to obtain a time coefficient. A scaling factor $SCALE_m$ is defined as 2.5 times the standard deviation of the time coefficient. For each $m \in \{S, O, N, D, J, F\}$, the ‘optimal’ SST forcing anomaly is defined as $A_m = SCALE_m \times SSTA_{\text{SONDJF}}$. Similarly, for each $m \in \{M, A, M, J, J, A\}$, the optimal SSTA forcing anomaly is defined as $A_m = SCALE_m \times SSTA_{\text{MAMJJA}}$. A_m is assumed to occur at the central date of the month. For the model forcing, it is added to (for A^+), or subtracted from (for A^-), the climatological monthly-mean SST, and linear time-interpolation is made between the centres of consecutive months. The transition between the February and March patterns (and the August and September patterns) is not thought to have a detrimental impact on the total atmospheric circulation, as the total circulation at any instant is likely to be dominated by atmospheric internal variability.

There are several advantages for using the above procedure: the SSTAs have realistic patterns, they may lead to a stronger atmospheric response than randomly chosen patterns, they do not favour any particular model (as they are based on observations) and, importantly, by using them it may be possible to compare the model responses with the best estimates of the real response (assumed to be the Z500 patterns arising from the same lagged MCA). The difference in seasonal-mean SST ($A^+ - A^-$) is shown in the surface temperature (TEMP) fields in Fig. 2. Although reasonably large, such SSTAs do exist in the raw seasonal-mean data.

(d) *Extra simulations of the HadAM3 model*

The HadAM3 simulations, A^+ and A^- , were repeated with different initial conditions to ensure better statistical significance when this model is considered alone. A parallel pair of HadAM3 simulations (termed H_{CK}) was made with a change to a single model parameter (the Charnock parameter) to check for sensitivity in the results. A further additional pair of simulations of HadAM3 (termed H_{TR}) was made with just the tropical SSTAs applied (the SSTAs south of the green line in the top panels of Fig. 2 (TEMP)). H_{CK} and H_{TR} were also repeated with different initial conditions to improve significance.

3. INDIVIDUAL MODEL RESPONSES

Figure 1 shows the mean seasonal-mean MSLP difference ($A^+ - A^-$) for each model. This difference is an estimate of each model's response to North Atlantic SSTs. Statistical significance at the 10% level using a two-sided t -test is indicated by shading. The percentage of the total area shown that is significant is quoted in the bottom right corner of each panel. Using the binomial distribution test of Livezey and Chen (1983) all panels are found to be field significant at the 10% level, assuming 55 spatial degrees of freedom for the global MSLP (Livezey and Chen suggested 30–60 degrees of freedom for the northern hemispheric Z500). All but two panels are significant at the 5% level assuming just 14 spatial degrees of freedom. A major result, which was unexpected based on a knowledge of the discrepancies highlighted in the introduction, is that the difference fields of the models appear quite similar, both locally in the North Atlantic region and also globally. JJA shows the best field significance, with negative anomalies over the tropical and subtropical Atlantic and over North and South America, and generally positive anomalies elsewhere. A striking feature for MAM, Fig. 1 (third column), is that all models (with the exception of $E4$) show a negative NAO dipole response in the North Atlantic region, with high pressures to the north and low pressures to the south. The Azores-minus-Iceland MSLP (i.e. the NAO-index anomaly) is negative for all models including $E4$. Values are -0.22 , -2.39 , -0.88 , -0.30 , -2.65 , and -3.37 hPa for $H3$, H_{CK} , $A3$, $E4$, $E5$, and $C2$, respectively (mean = -1.64 hPa, standard deviation = 1.34 hPa). In DJF, all models show consistent positive anomalies in the Aleutian low and, although not individually significant, all models show pressure anomalies of around -1 hPa over the whole of Eurasia. There is also a consistent pattern of MSLP anomalies over the North Atlantic, and even Greenland, during SON. Although the striking feature is the similarity in global and North Atlantic responses, there are also differences between models. The extent to which these differences reflect true model differences rather than being associated with chaotic atmospheric internal variability is discussed later.

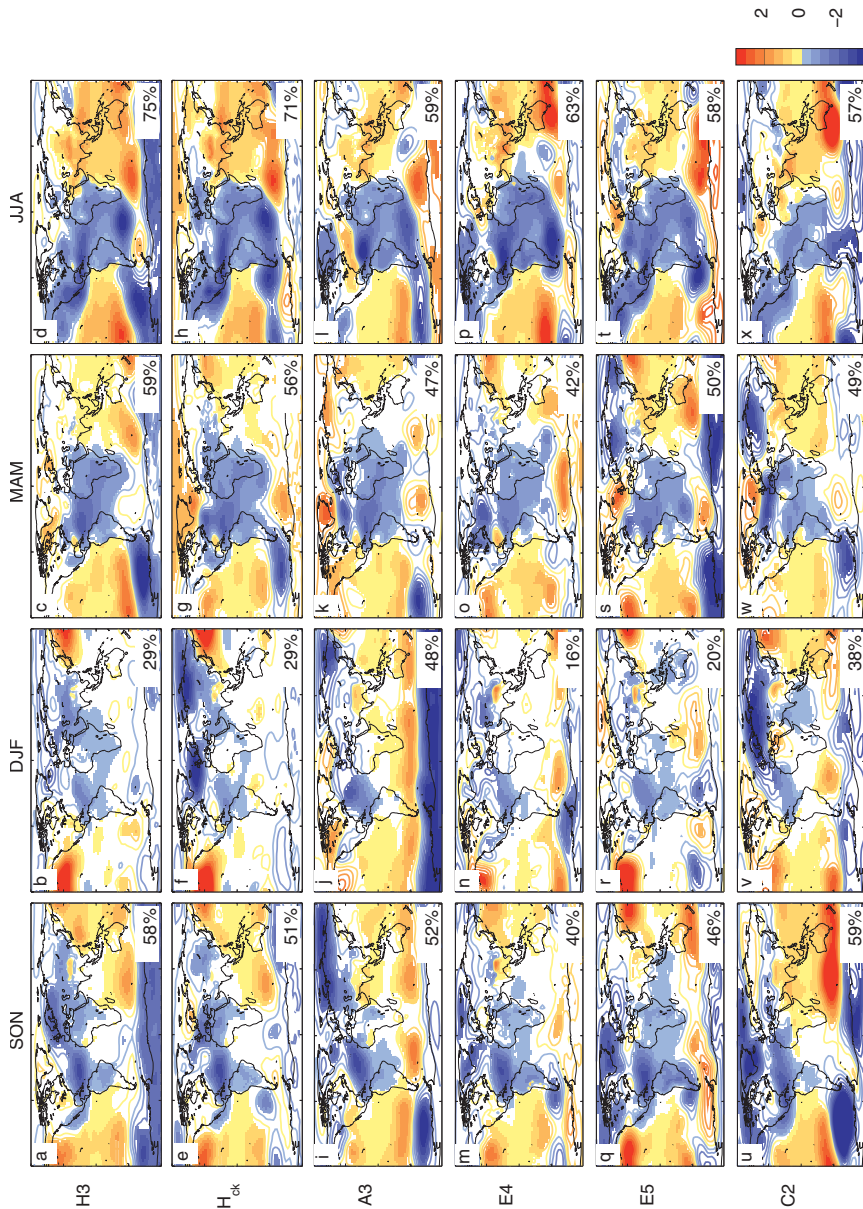


Figure 1. The 20-year-mean (40-year for H3 and H_{ck}) mean-sea-level pressure (MSLP) responses (hPa) of six atmospheric models in each season to imposed sea-surface temperature anomalies (SSTAs) in the North Atlantic. Note that SSTAs, ($A^+ - A^-$), are shown in the top panels of Fig. 2. Signals that are significant at the 10% level using a two-sided t -test are filled in colour, other values are contoured. The percentage given in the bottom right-hand corner of each panel is the percentage of the area shown that is significant at the 10% level (and, therefore, indicates the degree of field significance). See the text for a definition of the models.

4. MULTI-MODEL MEAN RESPONSE, ITS ORIGIN AND SIGNIFICANCE

(a) *Methodologies*

The similarity between models makes it legitimate to construct a multi-model mean and to refer to the mean anomalies as the ‘response’ to SST forcing. An issue highlighted in the introduction concerns the timescale-dependence of the importance for climate variability of forcing by SST. To determine what timescales, if there are any, at which the response is important for natural climate variability, the response is compared with the magnitude of temporally filtered atmospheric internal variability. Before introducing the model results, the simple method of determining these timescales is discussed.

For each model, i (1 to 6), and year, j (1 to 20), the difference between the two simulations, $d_{ij} = A_{ij}^+ - A_{ij}^-$, can be partitioned into an SST-forced response and random chaotic variability which is generated internally by the atmosphere. For a particular season, the response for model i can be estimated by the mean, $d_i = \frac{1}{20} \sum_{j=1}^{20} d_{ij}$, and the internal variability by the residual time-series $(d_{ij} - d_i)_{j=1}^{20}$. The response is likely to be smaller than the standard deviation of the residual time-series at most grid points. However, if timescales shorter than, say, n years are filtered from the residual time-series, then the signal to noise ratio will increase. The relevant question is at what filtering timescale, n , does the response to the fixed SSTAs used become comparable with the internal variability? Here the timescale is estimated by looking for equality between the magnitudes of the response and the filtered internal variability:

$$\frac{\mathbb{V}(d_{ij} - d_i)}{2n} = d^2, \quad (1)$$

where the left-hand side estimates the variance of the filtered internal variability and the right-hand side is the square of the estimated (multi-model) response, $d = \frac{1}{6} \sum_{i=1}^6 d_i$. Note that $\mathbb{V}(d_{ij} - d_i) = \frac{1}{120} \sum_{i,j} (d_{ij} - d_i)^2$ is the sample variance of the concatenation of the residual time-series of all models. The factor $\frac{1}{n}$ in (1) reflects the reduction in variance associated with taking the mean of n -years, and the factor $\frac{1}{2}$ arises from the fact that the difference of two simulations, A^+ and A^- has been taken. The assumption has been made that the concatenation of residual time-series at each grid point represents white noise. This latter assumption could be compromised by (for example) possible reddening effects due to interactions with the land surface. However, the variance of internal variability, as a function of filtering timescale, has also been estimated in another way (not assuming white noise) by using non-overlapping consecutive n -year time intervals (where n is a factor of 20). Similar timescales are found. Note also that internal variability is found to have a negligible impact on the multi-model mean.

Where short timescales (e.g. 1–2 years) are highlighted by this method, these indicate regions where the response would play a major role in observed climate anomalies (assuming the multi-model is perfect and the observed SSTAs agree in pattern and magnitude with those applied). Of particular interest here are regions that are not already highlighted by the analysis of variance applied to AMIP-style simulations. These are regions where the general influence of SSTAs may be small, but where a strong impact is achieved from the particular SSTAs used here. The location of these regions will clearly be sensitive to the choice of SSTAs used (this ‘caveat’ is the ‘trade-off’ in these experiments for greater statistical significance).

Where timescales longer than about two years are indicated, these timescales are likely to be *underestimated*. This is because the SSTA magnitudes used are probably too large to be realistic beyond one- to two-year averages and, therefore, the response is too strong for a simple comparison with filtered internal variability on these timescales.

To include a timescale-dependence of the forcing strength, a linear scaling parameter can be applied to the response. Justification for using a linear scaling and how it is defined is discussed later.

Figure 2 shows, coloured, the multi-model mean difference $d = (A^+ - A^-)$ for surface temperature, precipitation, Z500 and MSLP. All responses are clearly field significant (5% field significance is attained when 44% of the area is point-wise significant, assuming nine degrees of freedom for the regional Z500). Below, we only highlight features in the multi-model mean that are *stronger than decadal filtered internal variability* (inside the thick black contour) and *common to all models*.

(b) Tropical response

In all seasons there is a tropical-Atlantic baroclinic response (positive Z500 anomalies overlying negative MSLP; Fig. 2, rows 3 and 4) and this is clearly related to locally forced changes in precipitation, a northward shift of the intertropical convergence zone (ITCZ) (Nobre and Shukla 1996) and an associated strengthening of the subtropical jet. The weaker DJF precipitation response may be because the tropical SSTAs are weakest in DJF and do not engage strongly with the ITCZ, which is near its March southern extreme (Nobre and Shukla 1996). The positive ITCZ precipitation anomalies over the Sahel in JJA (Fig. 2(h)) (Folland *et al.* 1986) and the Amazon in MAM (Fig. 2(g)) (Nobre and Shukla 1996) are stronger than biennially filtered internal variability (inside the thin black contour) and lead to evaporative cooling and cold surface temperatures (Figs. 2(d) and (c), respectively). These regions are also generally highlighted in an analysis of variance as having high ‘potential predictability’. Throughout the rest of the tropics, throughout the year, the response is more equivalent-barotropic (Z500 and MSLP anomalies with the same sign—see, for example, Figs. 2(l) and (p)), with reduced ITCZ precipitation apparently subordinate to the large-scale forcing. The striking one-standard-deviation reduction in the Indian summer monsoon rainfall (Fig. 2(h)) is the strongest manifestation of this, suggesting causality for the ‘observed’ palæo relationship with North Atlantic SST (Gupta *et al.* 2003).

(c) Extratropical response

For MAM and JJA, wave-like responses are seen in the Z500 field over the extratropical North Atlantic region (Figs. 2(k) and (l)) and these are equivalent-barotropic in structure (compare Figs. 2(k) and (l) with Figs. 2(o) and (p)). These responses are reminiscent of ray-tracing results based on the linear barotropic response to subtropical upper-tropospheric vorticity forcing (Hoskins and Karoly 1981). For MAM, the (zonal wave-number 1) response (Fig. 2(k)) is stronger than decadal filtered internal variability, even over Greenland and, as seen previously, it projects onto the NAO at the surface (Fig. 2(o)). For JJA (Fig. 2(l)), an anticyclonic centre to the west of the UK appears to mark the southward turning point for this (higher zonal wave-number) response. A further cyclonic centre, possibly associated with the same wave-like response, is seen over southern Europe and the Mediterranean in JJA. Row 1 of Fig. 3 shows the 40-year Z500 results for $H3$. Although the first 20 years of $H3$ are included in the multi-model mean, the similarity between $H3$ (Fig. 3, row 1) and the multi-model mean (Fig. 2, row 3) suggests that further experiments with HadAM3 can help clarify the roles of the tropical and extratropical SSTAs in the multi-model mean. (Clearly, one doesn’t need to appeal to this similarity to assess the origin of the response in HadAM3 alone.) Row 2 of Fig. 3 shows parallel results for the tropical SSTA experiment, H_{TR} . For MAM, and to

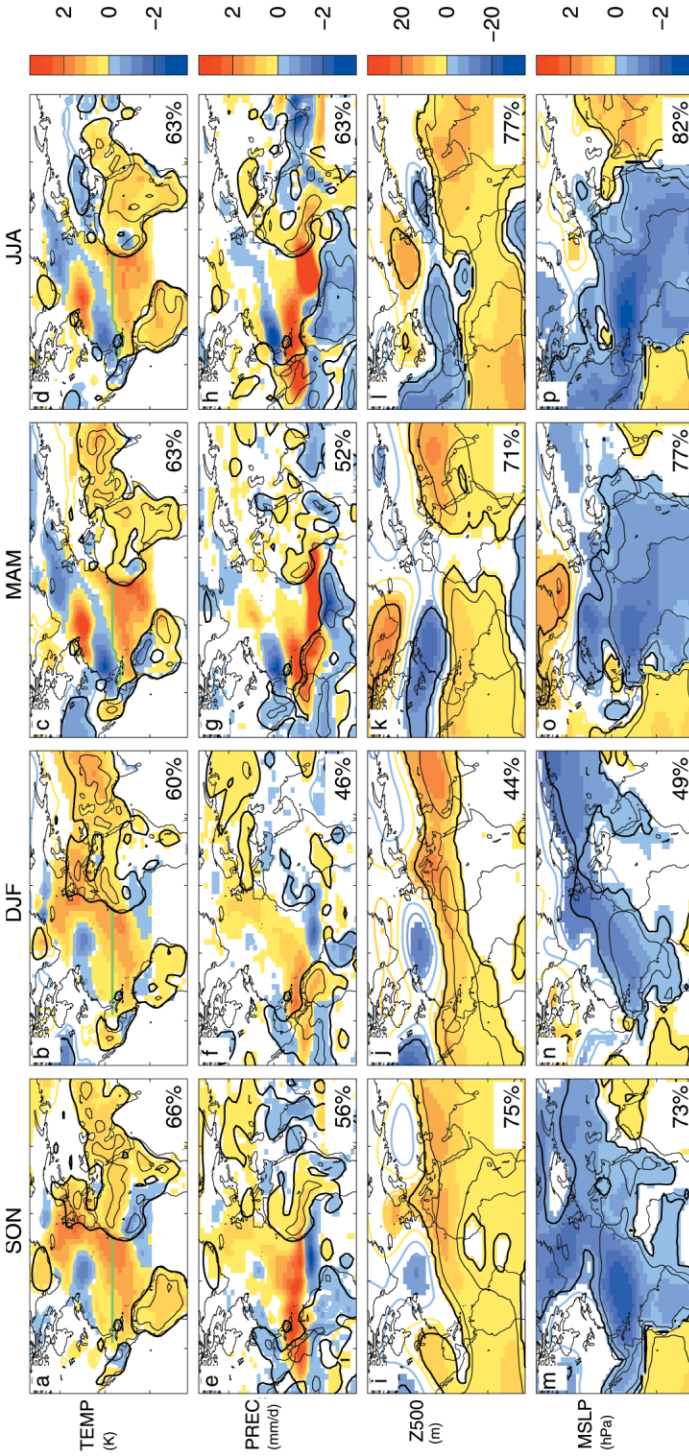


Figure 2. Seasonal-mean results from the multi-model mean of six models forced with the same North Atlantic sea-surface temperature (SST) anomalies. The rows correspond to surface temperature (TEMP), total precipitation (PREC), 500 hPa geopotential height (Z500) and mean-sea-level pressure (MSLP), respectively. The SST anomalies, ($A^+ - A^-$), used in the North Atlantic region can be seen in the top panels. The duplicate H3 and HcK simulations are not used. Signals that are significant at the 10% level using a two-sided t -test are filled in colour, other values are contoured (contours indicate values at the centre of each colour range). The quoted percentages in the bottom right-hand corner of each panel refer to the percentage of the area shown (land area for TEMP) for which the anomaly is statistically significant at the 10% level (and, therefore, indicate the degree of field significance). Black contours indicate the timescale $n = \sqrt{(d_{ij} - d_i) / 2d^2}$ (thick: ten years, thin: two years) at which the multi-model mean response is an important component of climate variability. The timescale is not applicable for TEMP over the sea. For clarity, the timescale for PREC is not plotted over the SST forcing region where it is nearly always less than two years, and it is smoothed for TEMP and PREC using a triangular truncation at T31. When the timescale dependence of the forcing strength is considered, the thick black contour is associated with a 30-year timescale. See the text for more details.

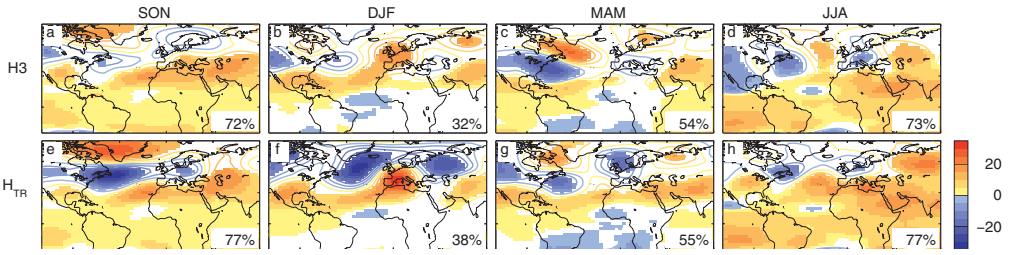


Figure 3. The 40-year mean 500 hPa geopotential-height (Z_{500}) responses (m) in HadAM3 when forced with the full North Atlantic sea-surface-temperature (SST) anomaly patterns, H_3 (top row) and when forced with just the tropical part of the SSTA patterns, H_{TR} (bottom row). Signals that are significant at the 10% level using a two-sided t -test are filled in colour, other values are contoured. The percentage given in each panel is the percentage of the area shown which is significant at the 10% level (and, therefore, indicates the degree of field significance).

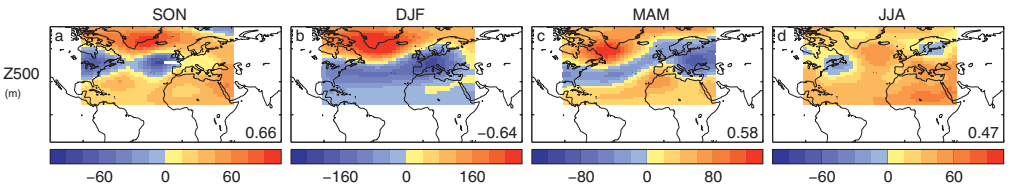


Figure 4. The 500 hPa geopotential-height (Z_{500}) anomaly patterns that arise in the maximal covariance analysis used to produce the forcing sea-surface-temperature anomalies. The patterns are scaled by five (2×2.5) times the standard deviation of their principal components so that they can be roughly compared with the multi-model mean. Area-weighted pattern correlations with the statistically significant parts of the multi-model mean are quoted.

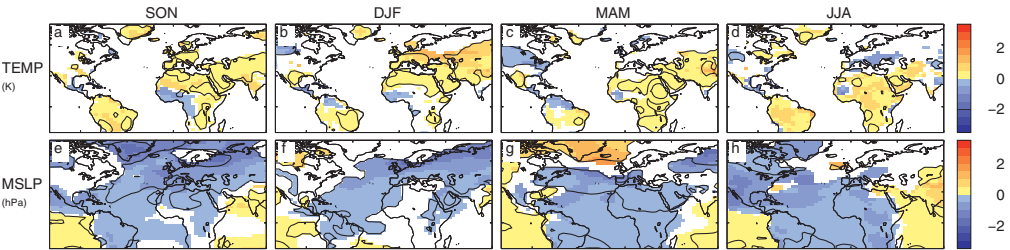


Figure 5. The standard deviation of inter-model variability, coloured and plotted with the same sign as the multi-model mean, and only where the multi-model mean is statistically significant. The rows correspond to surface temperature (TEMP) and mean-sea-level pressure (MSLP). The black contour indicates where residual atmospheric internal variability is estimated to account for 50% of inter-model variance. See the text for more details.

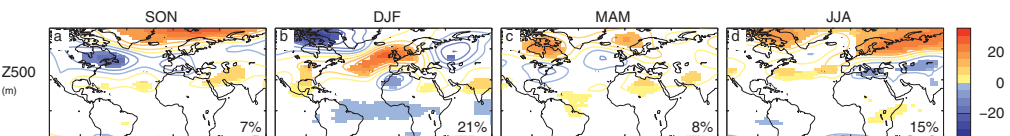


Figure 6. The multi-model mean nonlinear 500 hPa geopotential-height (Z_{500}) anomalies, $(A^+ - C) + (A^- - C)$. Signals that are significant at the 10% level using a two-sided t -test are filled in colour, other values are contoured. The percentage given in each panel is the percentage of the area shown which is significant at the 10% level (and, therefore, indicates the degree of field significance).

a lesser extent JJA, the similarity between $H3$ and H_{TR} (and the multi-model mean) does indeed indicate a tropical origin for these waves (the off-equatorial Caribbean may be particularly important—Hoskins and Sardeshmukh (1987)). Where significant, $(H3 - H_{TR})$ would highlight the influence of extratropical SSTAs alone, or features that require both tropical and extratropical SSTAs to be present. Statistical significance is likely to be more difficult to obtain for $(H3 - H_{TR})$, if only because a difference between two such fields contains twice the internal variability. Nevertheless, there are statistically significant differences. For example, surface latent-heat fluxes and precipitation differences (not shown) clearly highlight the role of extratropical SSTAs to the east of Florida and in the Gulf of Mexico. For MAM, these appear to force a similar wave-like pattern (perhaps by inducing a similar Rossby-wave source (Sardeshmukh and Hoskins 1988) over the Caribbean) although there are some phase differences. For JJA, the extratropical SSTAs may actually set the position of the high to the west of the UK (compare Figs. 3(d) and (h)) and be necessary, if not necessarily sufficient, for the remote coolness of the southern European low.

The extratropical multi-model Z500 responses for SON and DJF, Figs. 2(i) and (j), (which are forced by similar SSTAs) are similar to each other, suggesting that the response is particularly robust. H_{TR} (Figs. 3(e) and (f)) suggests that much of the HadAM3 Z500 response is forced by tropical and Caribbean SSTAs. To achieve a similar magnitude of response in DJF and SON, one can speculate that compensation may occur in the DJF Rossby-wave source with stronger mean absolute-vorticity gradients in the winter North American jet stream compensating for weaker tropical precipitation (compare Figs. 2(e) and (f)) and upper-tropospheric divergent-flow anomalies. In addition to the likely tropical forcing of a barotropic component of the extratropical response, Figs. 2(i) and (m) and Figs. 2(j) and (n) suggest that the multi-model extratropical response is quite baroclinic in nature (Kushnir 1994; Kushnir and Held 1996). Local SSTAs are seen in Figs. 2(e) and (f) to force important changes in precipitation (Kushnir and Held 1996; Rodwell *et al.* 1999; Frankignoul and Kestenare 2002) and atmospheric latent heating (Peterson *et al.* 2002) as far north as 60°N . The SON and DJF European temperature anomalies of $+1\text{ K}$ (Figs. 2(a) and (b)) are stronger than biennially filtered internal variability (indicated by the thin black contour) and would, therefore, be discernible if similar SSTAs were sustained for just two years. This response is a clear example of a strong forced signal that is not highlighted by the analysis of variance of AMIP-style simulations. HadAM3 suggests that extratropical SSTAs are important for this response. The increased precipitation seen downstream over central and northern Asia in these seasons (Figs. 2(e) and (f)) may be a combined advective effect of tropically enhanced westerlies (Figs. 3(e) and (f)) and extratropically increased moisture supply. An interesting feature of the $H3$ and H_{TR} results is that the statistically significant response to extratropical SSTAs in HadAM3 appears to damp the tropically forced response in SON and DJF (compare Figs. 3(a) and (e) with Figs. 3(b) and (f)). The extratropically forced Z500 response does appear to have a physical basis although the apparent damping may be coincidental and model dependent. For this reason it is discussed in the inter-model differences section 6.

Over the southern USA and Mexico, precipitation is reduced in all seasons but JJA (Fig. 2, row 2). The experiments with HadAM3 offer an interesting explanation of the origin of this response. H_{TR} suggests that the tropical SSTAs force a reduction in precipitation and an increase in land-surface temperature throughout the year (not shown), but that the cold extratropical SSTAs in $(H3 - H_{TR})$ off Florida and over the Gulf of Mexico tend to counteract this effect in JJA, possibly by enhancing the land–sea temperature contrast during the North American monsoon. The multi-model precipitation

signal over the southern USA and Mexico is stronger in places than biennially filtered internal variability for SON, MAM and JJA, but the region is not highlighted by a general analysis of variance. The MSLP anomalies seen over much of the USA in JJA (Fig. 2(p)) are also stronger than biennially filtered internal variability. Analysis of variance of the ECHAM3.5 AGCM does highlight this region as more strongly influenced by the SST (see, for example, Kushnir *et al.* (2002)), and HadAM3 shows particularly strong decadal potential predictability in this region (based on an analysis of variance, not shown). The interesting result here is that the Atlantic is implicated in the forcing (in addition to the generally accepted view that the tropical Pacific plays an important role (Kushnir *et al.* 2002)).

5. COMPARISON WITH OBSERVATIONAL RESPONSE

Figure 4 shows the Z500 patterns that arise in the MCA (Rodwell 2003) which was used to produce the SSTAs. (The shaded area is limited to that used in the MCA). While keeping in mind the uncertainties associated with these patterns (Rodwell and Folland 2002), it is worth comparing them with the model results. Area-weighted pattern correlations are made between the MCA Z500 pattern and the statistically significant parts of the multi-model mean response. These pattern correlations are quoted in Fig. 4 and indicate that the multi-model mean, where point-wise significant, generally agrees quite well (but not perfectly) with these MCA patterns in SON, MAM and JJA. Scaling Livezey and Chen's (1983) mid-estimate of 45 degrees of freedom for hemispheric Z500 to the area of significant multi-model mean response in the MCA region would suggest eight, eight and nine degrees of freedom for SON, MAM and JJA, respectively. Assuming these degrees of freedom, the pattern correlations are significantly different from zero at the 4%, 7% and 10% levels, respectively based on a one-sided *t*-test. The *significant* extratropical wave-like pattern in MAM in the multi-model mean (Fig. 2(k)), and also in *H3* and *H_{CK}* alone (Figs. 3(c) and (g)), agrees well with that in the MCA (Fig. 4(c)). There is also extratropical wave-like agreement for JJA (compare Fig. 2(l) and Fig. 4(d)) albeit with some phase or positional differences (*H3* alone, Fig. 3(d), shows the best extratropical agreement with Fig. 4(d)). For SON, the best agreement between the multi-model mean (Fig. 2(i)) and MCA (Fig. 4(a)) appears to occur in the subtropics, with poorer agreement in the extratropics (although *H3* and *H_{CK}*, Figs. 3(a) and (e), show reasonable extratropical agreement with Fig. 4(a)). The correspondence between model responses and MCA clearly breaks down for DJF (compare Fig. 2(j) and Fig. 4(b)), with little agreement, even in the subtropics. This may possibly be associated with increased storminess in the extratropical winter observations that could 'confuse' the MCA. There is better subtropical agreement between the multi-model DJF response and that detected in the observations for NDJ using a slightly different MCA technique (Czaja and Frankignoul 2002). Further discussion of storm-track-related differences is given in section 6.

Overall, considering the uncertainties and possible sources of discrepancy, these results do tend to validate the MCA technique *and* suggest that the model response patterns, at least in terms of Z500, are approximately correct. The increased magnitude in the MCA results may partly reflect the optimization inherent in the MCA technique although it may also suggest that the multi-model response to SST forcing is too weak. The latter explanation is consistent with the conclusions of Rodwell and Folland (2003).

6. INTER-MODEL RESPONSE DIFFERENCES

(a) *Quantification of differences*

Until now, the main concentration has been on the similarity between the model responses. Clearly, however, there are differences. Here, these differences are quantified and an attempt is made to partition the differences into a true difference between the statistically expected responses of the models and a part associated with sampling uncertainty. Figure 5 shows (coloured) inter-model standard deviations for TEMP and MSLP which are plotted, for ease of comparison, with the same sign as the multi-model mean. These standard deviations are generally weaker than the multi-model response (compare with Fig. 2). Hence, although there are differences between the individual model MSLP responses in Fig. 1, these differences do not obscure the mean response.

Each model mean anomaly can be considered to be the sum of the model's response to SSTA forcing and the average of 20 years of internal variability. It is found that internal variability *does* contribute significantly to the inter-model standard deviation. The black contour in Fig. 5 shows where (20-year filtered) atmospheric internal variability would be expected to account for 50% of the inter-model variance, i.e. where:

$$\frac{1}{20}\nabla(d_{ij} - d_i) = \frac{1}{2}\nabla(d_i - d), \quad (2)$$

with definitions as before. The left-hand side represents the 20-year filtered internal variance and the right-hand side represents half the inter-model variance. By examining the area enclosed by the black contour in Fig. 5 it is clear that, with the exception of JJA, much of the inter-model variance is actually associated with residual internal variability. Hence the model responses obtained from longer simulations would be even more similar than they are here. The uncertainty in the MAM NAO response (Fig. 5(g)) is seen to be partly associated with residual internal variability.

(b) *Physical interpretation of inter-model differences*

For SON and DJF, ($H3 - H_{TR}$) (not shown) indicates a strong North Atlantic equivalent-barotropic high centred above, and downstream of, the cool SSTAs off Newfoundland. Anomalies in surface latent-heat fluxes, precipitation *and* MSLP are clearly associated with the imposed extratropical SSTAs (over the eastern as well as western North Atlantic). The anomalous circulation is similar to the observational results of Kushnir and Held (1996) and the idealized modelling result of Rodwell *et al.* (1999) and would act, as in the paper by Hoskins and Karoly (1981), to satisfy surface thermodynamic balance. The response leads to an apparent (possibly coincidental) damping of the tropically forced extratropical response in HadAM3 (as noted earlier). The response to such extratropical SSTAs is likely to be sensitive to model storm-track characteristics (Peng and Whitaker 1999; Peng *et al.* 2002; Kushnir *et al.* 2002). The more robust linear baroclinic response of the multi-model mean is, perhaps, indicative of inter-model storm-track differences, and this may be another factor in the poorer extratropical correspondence between multi-model mean and observations in SON and DJF.

A3 shows Caribbean precipitation differences much more strongly linked to local evaporation than the other models in JJA and this may help explain its somewhat different (not necessarily wrong) extratropical response (Fig. 1(i)). Such tropical differences highlight the fact that even the relatively well understood tropical forcing of the extratropical circulation is not well represented in current models. This clear model difference, together with the general reduction in internal variability in JJA, helps explain why inter-model differences are more strongly highlighted in Figs. 5(d) and (h) (i.e. not enclosed in a black contour).

H_{CK} had its Charnock parameter, which affects surface fluxes of heat and momentum (Janssen and Viterbo 1996), increased by a factor 2.5 compared with $H3$. The anticipation was that this might increase sensitivity to SSTAs, particularly in regions of high wind speed. Consistent with this anticipation, H_{CK} (Fig. 1, row 2) does appear to show a general strengthening of the extratropical response over $H3$ (Fig. 1, row 1), particularly in DJF and MAM. In addition, south of 30°N where surface wind speeds are generally smaller, $(H_{CK} - H3)$ Z500 and MSLP are almost invariably not statistically significant. For MAM there is a statistically significant (and physically consistent) enhancement of the SST-forced surface latent-heat-flux (LHF) pattern and the NAO MSLP dipole when the Charnock parameter is increased. The previously quoted NAO index anomalies for $H3$ and H_{CK} were -0.22 and -2.39 , respectively. However, field significance for $(H_{CK} - H3)$ in MAM, even over just the extratropical box (120°W – 90°E , 30°N – 80°N), is poor (11% of the area is statistically significant for MSLP), and so there must be some doubt at present over how much of this change in the NAO index can really be attributed to the change in the Charnock parameter. The season which generally shows the best $(H_{CK} - H3)$ field significance over all the parameters tested is DJF (the largest percentage area of point-wise significance is 25% for Z500). This is consistent with the fact that surface wind speeds are generally largest in winter. However, there is a (statistically significant) *reduction* in the SST-forced LHFs in DJF as the Charnock parameter is increased. This means that the apparent extratropical enhancement seen between Figs. 1(b) and (f) does not have a simple explanation. SON and JJA do not show statistically significant changes in SST-forced LHF associated with a change in the Charnock parameter.

7. RESPONSE TO ATLANTIC FORCING, AND ITS UNCERTAINTY

We can, now, make a meaningful estimate of the response to the chosen SST forcing, of the uncertainty in the magnitude of the response and of the fraction of this uncertainty that is due to model error. For example, the mean northern European (0°E – 20°E , 45°N – 52°N) land surface-temperature anomaly in DJF is $+1.00$ K (Fig. 2(b)). The inter-model standard deviation is 0.58 K (Fig. 5(b)) with 69% of the inter-model variance explained by residual atmospheric internal variability. The remaining 31% can be attributed to model differences. For the UK, $\text{TEMP} \approx 0.90(\pm 0.37)$ K, which is comparable in magnitude (and sign) with that predicted to occur by 2050 under low CO_2 emission scenarios (Hulme *et al.* 2002). Note that the peak magnitudes of the SSTAs ($A^+ - A^-$) used here do occur in reality (e.g. as recently as JJA 2003) and are predicted to occur in a more sustained fashion under medium to high CO_2 emission scenarios (Cubasch *et al.* 2001).

8. LINEARITY OF THE RESPONSE

The multi-model nonlinear (anti-symmetric) component (Peng *et al.* 2002) to the Z500 anomalies, $(A^+ - C) + (A^- - C)$, Fig. 6, is smaller than the *collocated* significant linear component, $(A^+ - A^-)$, Fig. 2, row 3, and is not field significant. Pure statistics may partly explain this result, although it seems reasonable to infer that, to a first approximation, the multi-model response is primarily linear. Such linearity has been suggested previously based on observations for the tropics (Nobre and Shukla 1996) and extratropics (Kushnir 1994).

9. THE IMPORTANT TIMESCALES FOR THE RESPONSE

For natural variability, the typical magnitude of SSTAs is a function of the timescale over which they occur. SSTAs of the magnitude used here ($A^+ - A^-$) may well exist naturally for a single season or possibly even for the same season in two consecutive years, by chance or associated with the ‘re-emergence mechanism’, for example. The re-emergence mechanism involves the insulation of winter/spring-forced deep mixed-layer thermal anomalies by a summertime shallow stable layer and their re-emergence at the surface in the following late autumn and winter by mixing processes (see, for example, Alexander and Deser (1995)). Hence, the biennial timescales highlighted above indicate regions where the effect of SST forcing really could be important on biennial timescales (based on the SSTA patterns used here). This forcing is likely to be associated mainly with mixed-layer temperature anomalies. SSTAs of the magnitude of ($A^+ - A^-$) would *not* be expected to last for as long as a decade. Hence the response d , as defined in section 4(a), is also likely to be unrealistically large for long timescales. This implies, for timescales longer than about two years, an underestimation in (1) of the timescale for which the SST-forced response is an important component in climate variability. The aim here is to produce a better estimate for this timescale.

For a more realistic n -year forcing, for example $\alpha(n) \times (A^+ - A^-)$, where $|\alpha| < 1$, the approximate linearity established above suggests that the response would be $\approx \alpha d$. A better estimate of the timescale may, therefore, be made by solving (1) with d^2 replaced by $(\alpha d)^2$. Here, α is defined as:

$$\alpha(n) = \frac{\text{TRND}_{\text{TR}}}{\text{TEMP}_{\text{TR}}}, \quad (3)$$

where $\text{TRND}_{\text{TR}}(n)$ is the maximal n -year trend in the observed tropical SST, 1870–present, and TEMP_{TR} is the applied tropical SST difference, ($A^+ - A^-$).

There is no a priori guarantee that (1) can be solved for n when d^2 is replaced by $(\alpha d)^2$, but trial shows that it can be. The solution for each season gives a timescale for the thick contour in Fig. 2 of $n \approx 30$ years, at which the variance of filtered internal variability matches the squared scaled response. Hence, at a timescale of about 30 years, North Atlantic air–sea interaction may well play an appreciable role in natural climate variability. This timescale is considerably longer than the decorrelation time of the mixed layer and is likely, therefore, to be associated with lower-frequency variations, possibly involving variability of the thermohaline or ‘Gulf Stream’ circulations. Note that nothing is inferred here about predictability.

10. DISCUSSION AND CONCLUSIONS

This study has investigated the atmospheric response to North Atlantic (equator to 70°N) sea-surface temperature anomalies. This topic is meaningful if we restrict our interest to the mean quasi-instantaneous atmospheric response.

Because *interannual* atmospheric internal variability is strong compared with the responses of interest, it was considered essential to make the experiments under highly controlled conditions with as few degrees of freedom for the system as possible. This was done by using atmospheric models rather than coupled models, using the same SST and sea-ice climatologies for each model and restricting attention to fixed SSTA patterns. Other controlled conditions include using the same pre-industrial CO₂ and trace gases, although it is unlikely that this would have had a large impact on the results. In order to further improve the chances of achieving statistical significance,

'optimal' SSTAs were deduced from a lagged maximal-covariance analysis of observed monthly-mean SST and the following seasonal-mean 500 hPa geopotential height field.

Under these controlled and optimized conditions, the models show rather similar responses. It has been found that the mean response to SSTAs in the models used is generally larger in magnitude than the inter-model standard deviation. Further, the inter-model standard deviation has been shown to be often dominated by atmospheric internal variability so the true inter-model differences are likely to be considerably smaller than the magnitude of the multi-model mean response. The mean response for the particular SSTA patterns chosen includes, for example, a -1.6 hPa effect on the Azores-minus-Iceland North Atlantic Oscillation index, $+1$ K surface temperature anomalies over Europe and a remote one-standard-deviation decrease in Indian summer monsoon rainfall.

Much of the large-scale response appears to be forced by the tropical SSTAs and communicated via barotropic Rossby waves to the extratropics. Nevertheless, a clear role for extratropical SSTAs in forcing an extratropical response is evident throughout the year. Kushnir *et al.* (2002) highlighted the major discrepancies between (other) models in their *local* extratropical responses. For SON and DJF, a barotropic response associated with extratropical SSTAs is indicated for the one model (HadAM3) investigated here. However, the fact that the extratropical multi-model mean response is rather baroclinic suggests that such barotropic (storm-track related) discrepancies also exist between the current models.

There are reasonable similarities in SON, MAM and JJA between the statistically significant parts of the multi-model response patterns and the estimated observational responses. The magnitudes of the multi-model responses for each season are generally less than those estimated for the observations, and so it seems likely that the models do *not*, at the very least, overestimate the magnitude of the response.

The multi-model results suggest that, for strong SSTA magnitudes such as those that occurred in JJA 2003 for example, air-sea interaction may have a detectable influence on climate variations on timescales as short as two years. Mixed-layer processes are likely to be important at this timescale. Linear scaling of the response also suggests that air-sea interaction may play a significant role in natural climate variability on a longer timescale of about 30 years. For this timescale, variations in the thermohaline and 'Gulf Stream' circulations may be important. One implication of this result is that climate models need to represent accurately variability of these oceanic-circulation features if we are to have confidence in forecasts of climate change. The stronger estimates of the true observational response suggest that, if anything, the 30-year timescale identified here should be considered as an upper bound.

Attention here has focused on the mean response and its magnitude compared with atmospheric internal variability. Clearly SSTAs could also affect the relative importance of different 'modes' of variability about the mean climate (Peng and Robinson 2001). This possibility has not been addressed here. To help achieve statistical significance, fixed SSTA patterns have been used for each season. Other SSTA patterns would be likely to give different response patterns (for example the 'tripole' SSTA pattern could force an NAO response in DJF, as it does here for MAM). Hence the regions highlighted by the biennial and 10/30-year contours in Fig. 2 are also likely to be sensitive to the choice of SSTA patterns. The SSTA patterns most relevant for forcing the atmosphere may vary with timescale. This aspect has not been investigated here although the decadal and multi-decadal nature of the principal components arising from the MCA procedure of Rodwell (2003) suggest that low frequencies do play a major role in the definition of the SSTA patterns used here.

ACKNOWLEDGEMENTS

This work was completed while MJR was working at the Met Office and during a short sabbatical at the National Center for Atmospheric Research, USA. We thank Christophe Cassou and Adam Phillips for help with the CAM2 model and Chris Folland, Jeff Knight, Laurent Terray, the reviewers and associate editor for their help or useful comments. This work was funded by the European Union's PREDICATE project.

REFERENCES

- Alexander, M. A. and Deser, C. 1995 A mechanism for the recurrence of wintertime midlatitude SST anomalies. *J. Phys. Oceanogr.*, **25**, 122–137
- Barsugli, J. J. and Battisti, D. S. 1998 The basic effects of atmosphere–ocean thermal coupling on mid-latitude variability. *J. Atmos. Sci.*, **55**, 477–493
- Bjerknes, J. 1964 Atlantic air–sea interaction. *Adv. Geophys.*, **10**, 1–82
- Bretherton, C. S. and Battisti, D. S. 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
- Cayan, D. R. 1992 Latent and sensible heat-flux anomalies over the northern oceans—Driving the sea-surface temperature. *J. Phys. Oceanogr.*, **22**, 859–881
- Collins, M. 2002 Climate predictability on interannual to decadal time scales: The initial value problem. *Climate Dyn.*, **19**, 671–692
- Cubasch, U., Meehl, G. A., Boer, G. J., Stouffer, R. J., Dix, M., Noda, A., Senior, C. A., Raper, S. and Yap, K. S. 2001 ‘Projections of future climate change’. In *Climate change 2001: The scientific basis. Contribution of the working group I to the Third assessment report of the Intergovernmental Panel on Climate Change*. Eds. J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell and C. A. Johnson. Cambridge University Press, Cambridge, UK and New York, USA
- Czaja, A. and Frankignoul, C. 1999 Influence of North Atlantic SST on the atmospheric circulation. *Geophys. Res. Lett.*, **26**, 2969–2972
- 2002 Observed impact of Atlantic SST anomalies on the North Atlantic Oscillation. *J. Climate*, **15**, 606–623
- Davies, J. R., Rowell, D. P. and Folland, C. K. 1997 North Atlantic and European seasonal predictability using an ensemble of multidecadal atmospheric GCM simulations. *Int. J. Climatol.*, **17**, 1263–1284
- Déqué, M., Dreveton, C., Braun, A. and Cariolle, D. 1994 The ARPEGE/IFS atmospheric model: A contribution to the French community climate modelling. *Climate Dyn.*, **10**, 249–266
- Ferreira, D., Frankignoul, C. and Marshall, J. 2001 Coupled ocean–atmosphere dynamics in a simple midlatitude climate model. *J. Climate*, **14**, 3704–3723
- Folland, C. K., Palmer, T. N. and Parker, D. E. 1986 Sahel rainfall and worldwide sea temperatures, 1901–85. *Nature*, **320**, 602–607
- Frankignoul, C. and Kestenare, E. 2002 The surface heat flux feedback. Part I: Estimates from observations in the Atlantic and North Pacific. *Climate Dyn.*, **19**, 633–647
- Gates, W. L. 1992 AMIP: The Atmospheric Model Intercomparison Project. *Bull. Am. Meteorol. Soc.*, **73**, 1962–1970
- Gupta, A. K., Anderson, D. M. and Overpeck, J. T. 2003 Abrupt changes in the Asian southwest monsoon during the Holocene and their links to the North Atlantic Ocean. *Nature*, **421**, 354–357
- Hall, N. M. J., Derome, J. and Lin, H. 2001 The extratropical signal generated by midlatitude SST anomaly. Part I: Sensitivity at equilibrium. *J. Climate*, **14**, 2035–2053
- Hoskins, B. J. and Karoly, D. J. 1981 The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, **38**, 1179–1196
- Hoskins, B. J. and Sardeshmukh, P. D. 1987 A diagnostic study of the dynamics of the northern hemisphere winter of 1985–86. *Q. J. R. Meteorol. Soc.*, **113**, 759–778
- Hulme, M., Jenkins, G. J., Lu, X., Turnpenny, J. R., Mitchell, T. D., Jones, R. G., Lowe, J., Murphy, J. M., Hassell, D., Boorman, P., McDonald, R. and Hill, S. 2002 ‘Climate change scenarios for the United Kingdom’ UKCIP02 Scientific Report, Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK

- Hurrell, J. W. 1995 Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, **269**, 676–679
- Janssen, P. A. E. M. and Viterbo, P. 1996 Ocean waves and the atmospheric climate. *J. Climate*, **9**, 1269–1287
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woolen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. 1996 The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.*, **77** 437–471
- Kushnir, Y. 1994 Interdecadal variations in North Atlantic sea surface temperatures and associated atmospheric conditions. *J. Climate*, **7**, 141–157
- Kushnir, Y. and Held, I. M. 1996 Equilibrium atmospheric response to North Atlantic SST anomalies. *J. Climate*, **9**, 1208–1220
- Kushnir, Y., Robinson, W. A., Bladé, I., Hall, N. M. J., Peng, S. and Sutton, R. 2002 Atmosphere GCM response to extratropical SST anomalies: Synthesis and evaluation. *J. Climate*, **15**, 2233–2256
- Lau, N.-C. and Nath, M. J. 1996 The role of the ‘atmospheric bridge’ in linking tropical Pacific ENSO events to extratropical SST anomalies. *J. Climate*, **9**, 2036–2057
- Livezey, R. E. and Chen, W. Y. 1983 Statistical field significance and its determination by Monte Carlo techniques. *Mon. Weather Rev.*, **111**, 46–59
- Mehta, V., Suarez, M., Manganello, J. and Delworth, T. 2000 Oceanic influence on the North Atlantic Oscillation and associated northern hemisphere climate variations: 1959–1993. *Geophys. Res. Lett.*, **27**, 121–124
- Nobre, P. and Shukla, J. 1996 Variations of sea surface temperature, wind stress, and rainfall over the tropical Atlantic and South America. *J. Climate*, **9**, 2464–2479
- Okumura, Y., Xie, S.-P., Numaguti, A. and Tanimoto, Y. 2001 Tropical Atlantic air–sea interaction and its influence on the NAO. *Geophys. Res. Lett.*, **28**, 1507–1510
- Palmer, T. N. and Sun, Z. 1985 A modelling and observational study of the relationship between sea surface temperature in the north-west Atlantic and the atmospheric general circulation. *Q. J. R. Meteorol. Soc.*, **111**, 947–975
- Peng, S. and Whitaker, J. S. 1999 Mechanisms determining the atmospheric response to midlatitude SST anomalies. *J. Climate*, **12**, 1393–1408
- Peng, S. and Robinson, W. A. 2001 Relationships between atmospheric internal variability and the responses to an extratropical SST anomaly. *J. Climate*, **14**, 2943–2959
- Peng, S., Robinson, W. A. and Li, S. 2002 North Atlantic SST forcing of the NAO and relationships with intrinsic hemispheric variability. *Geophys. Res. Lett.*, **29**, doi: 10.1029/2001GL014043
- Peterson, K. A., Greatbatch, R. J., Lu, J., Lin, H. and Derome, J. 2002 Hindcasting the NAO using diabatic forcing of a simple AGCM. *Geophys. Res. Lett.*, **29**, doi: 10.1029/2001GL014502
- Pope, V. D., Gallani, M. L., Rowntree, P. R. and Stratton, R. A. 2000 The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Climate Dyn.*, **16**, 123–146
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C. and Kaplan, A. 2003 Global analyses of sea surface temperature, sea-ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, doi: 10.1029/2002JD002670
- Rodwell, M. J. 2003 ‘On the predictability of North Atlantic climate’. Pp. 173–192 in *The North Atlantic Oscillation*. Eds. J. W. Hurrell, Y. Kushnir, G. Ottersen and M. Visbeck. American Geophysical Union, Washington, DC, USA
- Rodwell, M. J. and Folland, C. K. 2002 Atlantic air–sea interaction and seasonal predictability. *Q. J. R. Meteorol. Soc.*, **128**, 1413–1443
- 2003 Atlantic air–sea interaction and model validation. *Ann. Geophys.*, **46**, 47–56

- Rodwell, M. J. and Ingram, W. J. 2000 'Notes on optimal run lengths for fixed SST anomaly experiments'. Hadley Centre Internal Note 96, (available from Hadley Centre for Climate Prediction and Research, Met Office, FitzRoy Road, Exeter, Devonshire EX1 3PB, UK)
- Rodwell, M. J., Rowell, D. P. and Folland, C. K. 1999 Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature*, **398**, 320–323
- Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U. and Schulzweida, U. 1996 'The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate'. Max-Planck-Institut für Meteorologie Report 218 (available from Max-Planck-Institut für Meteorologie, Bundesstr. 55, 20146 Hamburg, Germany)
- Roeckner, E., Buml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblüeh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U. and Tompkins, A. 2003 'The atmospheric general circulation model ECHAM 5: Part 1: Model description'. Max-Planck-Institut für Meteorologie Report 349 (available from Max-Planck-Institut für Meteorologie, Bundesstr. 55, 20146 Hamburg, Germany)
- Rowell, D. P. and Zwiers, F. W. 1999 The global distribution of sources of atmospheric decadal variability and mechanisms over the tropical Pacific and southern North America. *Climate Dyn.*, **15** 751–772
- Rowntree, P. R. 1976 Response of the atmosphere to a tropical Atlantic ocean temperature anomaly. *Q. J. R. Meteorol. Soc.*, **102**, 607–625
- Sardeshmukh, P. D. and Hoskins, B. J. 1988 The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.*, **45**, 1228–1251
- Terray, L. and Cassou, C. 2002 Tropical Atlantic sea surface temperature forcing of quasi-decadal climate variability over the North Atlantic–European region. *J. Climate*, **15**, 3170–3187
- Walker, G. T. and Bliss, E. W. 1932 World Weather V. *Mem. R. Meteorol. Soc.*, **4**, 53–84