Climate Dynamics (1997) 13: 459–487

Climate Dynamics

A predictability study of simulated North Atlantic multidecadal variability

S. M. Griffies[1], **K. Bryan**[2]

[1]Geophysical Fluid Dynamics Laboratory, Route 1, Forrestal Campus, Princeton NJ 08542, USA
[2]Princeton University, Atmospheric and Oceanic Sciences Program, Sayre Hall, Forrestal Campus, Princeton NJ 08544-0710, USA

Received: 28 October 1996 / Accepted: 21 March 1997

Abstract. The North Atlantic is one of the few places on the globe where the atmosphere is linked to the deep ocean through air–sea interaction. While the internal variability of the atmosphere by itself is only predictable over a period of one to two weeks, climate variations are potentially predictable for much longer periods of months or even years because of coupling with the ocean. This work presents details from the first study to quantify the predictability for simulated multidecadal climate variability over the North Atlantic. The model used for this purpose is the GFDL coupled ocean-atmosphere climate model used extensively for studies of global warming and natural climate variability. This model contains fluctuations of the North Atlantic and high-latitude oceanic circulation with variability concentrated in the 40–60 year range. Oceanic predictability is quantified through analysis of the time-dependent behavior of large-scale empirical orthogonal function (EOF) patterns for the meridional stream function, dynamic topography, 170 m temperature, surface temperature and surface salinity. The results indicate that predictability in the North Atlantic depends on three main physical mechanisms. The first involves the oceanic deep convection in the subpolar region which acts to integrate atmospheric fluctuations, thus providing for a red noise oceanic response as elaborated by Hasselmann. The second involves the large-scale dynamics of the thermohaline circulation, which can cause the oceanic variations to have an oscillatory character on the multidecadal time scale. The third involves nonlocal effects on the North Atlantic arising from periodic anomalous fresh water transport advecting southward from the polar regions in the East Greenland Current. When the multidecadal oscillatory variations of the thermohaline circulation are active, the first and second EOF patterns for the North Atlantic dynamic topography have predictability time scales on the order of 10–20 y, whereas EOF-1 of SST has predictability time scales of 5–7 y. When the thermohaline variability has weak multidecadal power, the Hasselmann mechanism is dominant and the predictability is reduced by at least a factor of two. When the third mechanism is in an extreme phase, the North Atlantic dynamic topography patterns realize a 10–20 year predictability time scale. Additional analysis of SST in the Greenland Sea, in a region associated with the southward propagating fresh water anomalies, indicates the potential for decadal scale predictability for this high latitude region as well. The model calculations also allow insight into regional variations of predictability, which might be useful information for the design of a monitoring system for the North Atlantic. Predictability appears to break down most rapidly in regions of active convection in the high-latitude regions of the North Atlantic.

Correspondence to: S. M. Griffies
E-mail: smg@gfdl.gov

1 Introduction

A primary objective of several climate research programs, such as the Atlantic Climate Change Program (ACCP), the Atlantic portion of the World Ocean Circulation Experiment (WOCE), and the second component of the Climate Variability and Prediction Research Programme (CLIVAR), is to design practical means for monitoring climate variability of decadal and multidecadal time scales, such as that observed in the North Atlantic. The monitoring would be focused on natural climate variations and the detection of possible long-term anthropogenic climate change. Presently, the TOGA/TAO (Tropical Ocean and Global Atmosphere/Tropical Atmosphere Ocean) array, in conjunction with other measurements made from satellites and ships of opportunity, provides a useful basis for model forecasts of El Niño – Southern Oscillation (ENSO) (e.g., Neelin et al. 1994). The question arises as to whether the very different air–sea interaction processes involved in North Atlantic climate variability can likewise be projected forward using models which allow for the added involvement of the thermohaline circulation.

Here we present details from the first study of the predictability of simulated Atlantic multidecadal climate variability. The coupled model used is that analyzed by Delworth et al. (1993 and 1997; hereafter DMS1 and DMS2). Through the analysis of ensemble experiments using the coupled model, the aim of this study is to make a quantitative assessment of the potential predictability of very long time scale climate variability in the North Atlantic. The study also is aimed at providing some insights into the actual requirements that a real-world monitoring system needs to maintain in order to produce forecasts whose skill approaches the idealized predictability time scales deduced here. Preliminary results from this study were reported in Bryan and Griffies (1996) and Griffies and Bryan (1997).

Before going into details of the predictability study, it is helpful to review some simple physical models of North Atlantic climate variability and the associated predictability. The simplest model is Hasselmann's (1976) random walk model of air–sea interaction. In this model, the ocean is taken to be a homogeneous reservoir of water, such as a mixed layer (which could extend to the abyss in the high latitudes of the North Atlantic), with random heating and cooling applied at the upper boundary. The long term effects of the surface heat balance are parametrized as a linear negative feedback which prevents the reservoir from moving too far away from a stable climate equilibrium. The combined effect of the random heating and cooling and the negative feedback is to allow random excursions of SST anomalies whose magnitude are determined by the strength of the feedback. The temporal variability arising from this mechanism displays a red noise power spectrum, which decays as ω^{-2} for high frequency ω and becomes a constant for low frequency (see Sect. A.1 in the Appendix for further discussion). As noted by Hasselmann (1976), and earlier by Lorenz (1973) (see also discussion in the Appendix Sect. A.1.2), useful forecasts in the form of *damped persistence*; i.e., damping an anomaly back to the climatological mean on a time scale of the auto-correlation time, can be extracted from this form of variability. Since this auto-correlation time can extend out well beyond that typically realized by the atmospheric forcing, the damped persistence forecast could be quite useful for climate predictions above that provided by pure persistence, where pure persistence simply forecasts no change.

The relevance of the Hasselmann idea to the problem of North Atlantic variability was suggested by Bryan and Hansen (1995), who considered a stochastically forced Stommel (1961) two-box model of the thermohaline circulation. Building on the ideas of Bryan and Hansen (1995), Griffies and Tziperman (1995) introduced another conceptual model based on the stochastic forcing of a four-box model originally considered by Huang et al. (1992). In these box models, thermohaline overturning is driven by latitudinal buoyancy gradients set up by surface fluxes of heat and salinity. In addition to a time-mean tendency to increase the density of northern waters in the North Atlantic, these fluxes have a large stochastic component associated with atmospheric disturbances similar to the Hasselmann (1976) model. When introduced into the box model of Griffies and Tziperman (1995), the effect provides a driving for the thermohaline circulation (THC) on all time scales. The model's thermohaline dynamics provides a response which has power concentrated in the multidecadal time scale. In particular, the feedbacks from temperature anomalies (acting as a negative feedback) and salinity anomalies (acting as a positive feedback) contribute to the decay and growth, respectively, of the THC circulation. A time delay in the THC response to north–south pressure gradients allows for damped oscillatory changes in the ocean circulation. The resulting mechanical analog is a noise-driven, damped linear oscillator. Depending on the strength of the damping, the damped oscillator model can be much more predictable than the simpler red noise model of Hasselmann. In analyzing the behavior of the North Atlantic variability in the global coupled model considered here, an attempt will be made to determine whether the Hasselmann or oscillator mechanism, a suitable combination, or an alternative nonlinear mechanism, is an appropriate simplified analog.

The main body of this study can be divided into two halves. The first half discusses the model simulation of Atlantic variability through analysis of the dominant EOF patterns for certain oceanic fields. An underlying theme here is that the dominant patterns associated with the multidecadal time scales are more predictable than the patterns associated with smaller space and shorter time scale variability. It is for this reason that the analysis of EOF patterns, which effectively sorts the variability according to space and time scales, also provides a sorting of the predictability scales. The second half of the work discusses the predictability of these dominant patterns as determined by ensemble experiments with the coupled model. More specifically, Sect. 2 briefly describes the GFDL coupled climate model, the preparation of the model data used for the subsequent analysis, and the design of the ensemble experiments. Section 3 considers the model's variability as seen in the zonally averaged meridional-depth plane as well as horizontal patterns localized to the North Atlantic. Section 4 discusses issues related to certain regimes of multidecadal variability realized during a particular 1000 year model integration as well as the climatologies used for determining the predictability of these regimes. Sections 5, 6, and 7 discuss the predictability times deduced from four ensembles of coupled model experiments. Section 8 provides discussion and conclusions. Appendix A presents a mathematical discussion of the predictability found in the red noise and noise driven damped harmonic oscillator processes. Appendix B discusses a method for bounding the ensemble predictability with that based on a red noise process.

2 The model and the experimental design

2.1 The model

The coupled ocean-atmosphere-ice-land surface climate model employed here was developed at the Geophysical Fluid Dynamics Laboratory (GFDL) to study climate and the possible consequences of enhanced greenhouse warming. It has been well documented through studies of various climate change scenarios (Manabe and Stouffer 1988,

1995; Manabe et al. 1991, 1992; Stouffer et al. 1989, 1994; Knutson et al. 1997), as well as studies of the model's natural variability (DMS1, DMS2, and Manabe and Stouffer 1996). Relevant to the current work in the North Atlantic variability seen in a multi-millenial control integration which results in a fairly realistic representation of the present-day stable climate system. The variability whose predictability time scales are of interest here is that seen in the model's middle and high-latitude regions in the Atlantic. The spectral power for the model variability of interest is concentrated in the 40–60 y range (DMS1, DMS2). The strongest multidecadal power arises from mid-latitude to sub-polar variations in the THC (DMS1) and variability in the high latitude Greenland Sea region (DMS2), and will be discussed further in subsequent sections.

It should be recognized that this model's simulation of multidecadal variability in the North Atlantic and Greenland Sea may not be adequate to describe all observed multi-decadal climate variability for this region. Nevertheless, the successes that it has had in representing certain phenomena lend support to our using it for the current predictability study. The experiments to be presented comprise a total of four ensembles amounting to a total of 1200 y of model integration. They provide a means for quantifying, at least to leading order, predictability time scales associated with the model's multidecadal North Atlantic variability. The model approach follows that pioneered by studies of atmospheric predictability (Lorenz 1969) and is in use today (see, e.g., Palmer 1996 for a review). The purpose of using a coupled ocean-atmosphere model is to reduce the uncertainities related to boundary conditions introduced when using ocean models forced with highly simplified atmospheric parametrizations (see, e.g., the review by Weaver and Hughes 1992). Also, a realistic atmospheric component is important since we believe that climate predictability cannot be properly estimated without taking into account the effectively random fluctuations of synoptic atmospheric forcing.

All model fields analyzed in this work are yearly mean fields. For developing the various model climatologies, a linear detrending was used to help remove certain long-term model adjustment trends. These adjustments were seen to affect mostly the deep water properties, with only mild effects seen in the upper ocean (i.e., the upper 1000–2000 m). Analysis of the multidecadal variability without the linear detrending does not alter the results in any significant manner. EOF patterns were computed as eigenvectors of the zero lag covariance matrix. A geometrical weighting was applied to the time series in order to account for the spherical grid in the horizontal and stretched grid in the vertical. Spectra for the EOF's principle component time series are computed with a maximum of a 50 y lag using a Tukey window (Chatfield 1989, chapter 7).

2.2 Ensemble design and method for measuring predictibility

A perturbation of the synoptic-scale atmosphere typically results within a few days or weeks in a randomly different atmosphere within the climatology (Palmer 1996 and references). This behavior suggests an ensemble design on which the initial conditions for the atmospheric model are chosen randomly from model climatology with the constraint that the new initial conditions correspond to exactly the same time of year (1 January) as the original atmospheric initial conditions. Furthermore, the ocean's original initial state will remain unchanged for each member of a particular ensemble. Both of these initial states are unrealistic. For example, realistic observations can never specify the oceanic state perfectly, and atmospheric analysis can provide more information than climatology. Additionally, predictability times are necessarily dependent on the details of the perturbaton applied to the initial states. Nevertheless, the design appears to us to provide a simple and useful starting point for this initial investigation of middle and high latitude predictability of multidecadal variability.

Because members in the ensemble will exhibit sensitivity to the perturbed initial conditions, after some time the information in the ensemble will be of no more use than information already contained in the climatology (assuming that the system is ergodic, which means that infinite time and infinite ensemble averages are equivalent). It is at this time scale that the system has reached its predictability limits. There are various methods used to measure this information and hence to determine the predictability times. For this study, the ensemble variance (i.e., mean of the squared distance of the ensemble members from the ensemble mean; see Appendix A for more details) normalized by that of the climatology is used. This statistic measures the spread amongst the ensemble member and has been used in other predictability studies (e.g., Shukla 1981). If the statistical behavior of the system is well described by a linear stochastic process, and the initial conditions of the ensemble members are distributed in a Gaussian manner, the ensemble variance completely describes the amount of useful predictive information contained in the ensemble. Although some of the variability we will consider exhibits a distinctly nonlinear behavior, for the present purposes it is sufficient to employ the variance in order to provide a general understanding of the predictability time scales active in the model. Therefore, the predictability time scales will be defined by those times during the integration of the ensemble when the ensemble variance shows no significant difference from that of the climatology. Anderson and Stern (1996) provide discussion and references for more general approaches to measuring predictability.

It is important to place some confidence limits on the time scales deduced from the ensemble variance. Namely, how different do the variances from the ensemble and climate need to be in order to be sure of the significance of these differences? A standard procedure for answering this question is the Fisher F rest (see, e.g., Jenkins and Watts 1968; Shukla 1981; Anderson and Stern 1996). Applying the procedures discussed in Jenkins and Watts (1968), to the various fields used in this study, we arrive at the result that whenever the ensemble variance is less than roughly 50% of the climatological variance, we can be at least 95% confident that this difference is significant. Different fields show slightly different cutoffs, but for purposes of

discussion, 50% is sufficient and will be employed in this study. If the system is behaving in a linear noise-driven sense, as discussed in Appendix A, once the ensemble variance reaches 50% of the climatological variance it will continue to rise towards that of the climatology. It is conceivable, however, that some physical mechanism will allow for a reduction of ensemble variance, which indicates the potential for recovery of predictability and is indicative of nonlinear dynamics. The F-test provides a means for placing significance on any such fluctuations.

In addition to placing confidence bounds on the ratio of variances, there is a question regarding how different the ensemble variance is from that expected from a red noise ensemble. The distinction is important since it indicates the presence or absence of different physical mechanisms associated with the variability as well as to provide for different predictability times for this variability. In the spirit of considering the Hasselmann (1976) red noise process as a null hypothesis (Wunsch 1992; Sarachik et al. 1996), we will consider the significance of differences in the coupled model ensemble variance from that expected from a red noise ensemble. As shown in Sect. A of the Appendix, differences from red noise may manifest most distinctly and significantly through a slower than red noise initial growth of the ensemble variance. As long as the ensemble variance is underneath the 50% climatological value discussed, we can consider any differences from red noise to be of significance. Differences of the ensemble variance from red noise may arise from an oscillatory component, higher order linear processes, or nonlinear components in the variability. Therefore, it is interesting to consider 95% red noise confidence bounds on the ensemble variance. Further details and discussion related to this red noise bound are given in Appendix B.

3 Dominant patterns of North Atlantic variability

For the multidecadal variability in the coupled model's North Atlantic, there is a strong statistical association between fluctuations in the oceanic overturning and changes in the North Atlantic thermocline density distribution, SST, and sea surface salinity (SSS). In particular, SST variations in the model's northwest Atlantic have a dominant pattern of variability (i.e., its first empirical orthogonal function or EOF) with two main centers of action: one near Labrador and the other south of Greenland (DMS1 and Fig. 8 to be discussed later). This simulated pattern compares well with the multidecadal changes found in historical data as analyzed by Kushnir (1994). Since the mechanism for the low frequency SST variability is associated with fluctuations in the ocean model's thermohaline circulation, the model successfully embodies some essential features of Bjerknes' hypothesis for North Atlantic variability (Bjerknes 1964; Bryan and Stouffer 1991). The high-latitude Greenland Sea variability has recently been analyzed by DMS2 and is associated with the propagation of fresh water anomalies southward from the Arctic within the East Greenland Current into the North Atlantic, much as has been suggested by studies of the "Great Salinity Anomaly" by Dickson et al. (1988). The model signals from these events appear strongly in the SST averaged over a particular region just east of central Greenland, whose spectrum for a 2000 y integration of the model shows a clear peak near 50 y (DMS2).

The purpose of this section is to summarize the dominant patterns associated with the mutlidecadal variability in the model's North Atlantic. Model years 1–200, which are the same years analyzed by DMS1, are used for defining these patterns. We start with the patterns seen in the zonally averaged meridional-depth plane, then take the complementary perspective of the vertically averaged horizontal plane and finish with the dominant surface patterns in the temperature and salinity fields.

3.1 Meridional stream function

Figure 1 shows the first, second and third EOFs for the Atlantic meridional stream function. When used as basis

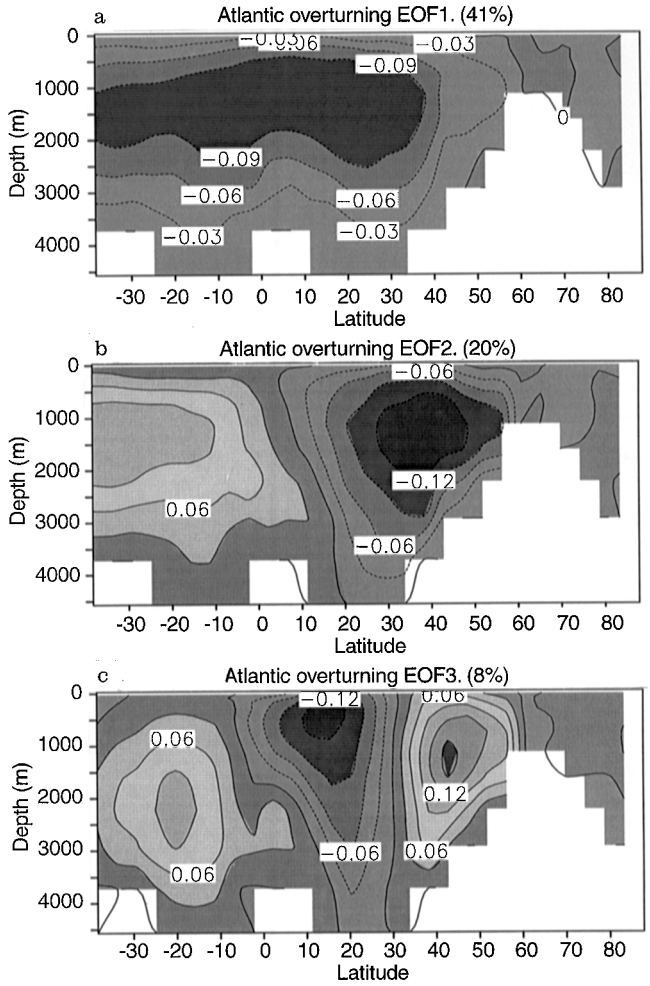


Fig. 1. **a** Atlantic meridional stream function EOF-1 (representing 41% of the variance) from model years 1–200. This pattern is interpreted as drift in the deep waters. **b** EOF-2 (representing 20% of the variance), which represents variability in the North Atlantic portion of the model. **c** EOF-3 (representing 8% of the variance), which represents a three-cell variability in the Atlantic basin. Units are dimensionless

functions for representing the stream function's variability, amplitude of the EOF patterns define the principal component (PC) time series. The auto-correlation functions and spectra for these time series are given in Fig. 2.

EOF pattern 1 (accounting for 41% of the total variance) represents a basin-wide adjustment in the Atlantic circulation which remains as a residual after the data were subjected to a linear detrending. The time scale for this pattern is quite long, hence the drift shown in its statistics in Fig. 2 (upper panels); i.e., this pattern is not stationary over the 200 y sample. The pattern has most of its variance in the southern deep waters. Such variability is not of direct concern in this study.

EOF pattern 2 (accounting for 20% of the total variance) represents a multidecadal fluctuation of the stream function with dominant variability in the northern North Atlantic and less variability in the sourthern half of the basin. The magnitude of the largest anomalies encompassed by this pattern are roughly 2 Sv ($1Sv = 10^6$ m^3/s). The statistics for PC-2 shown in Fig. 2 indicate very little trend and show a damped oscillatory character. The spectral power is concentrated in the 40–60 y band. It is this variability which is central to the present study. DMS1 define a THC index based on the maximum value of the meridional transport stream function. The multidecadal variability of this index is also coherent, much as the stream function's EOF-2. The THC index and the EOF-2 are of opposite phase but with good correlation (around -0.7; see Table 1 to be discussed in Sect. 3.5).

EOF-3 (representing 8% of the total variance) has a three-cell pattern in the Atlantic basin. The space–time scales for this and higher EOFs represent variability of the circulation on smaller temporal and spatial scales than that of interest in this study. As there is generally a non-trivial projection of these higher EOFs onto the North Atlantic region over which the THC index is defined, variability of the THC index will incorporate some of the smaller time scale variability of these patterns.

3.2 Zonally averaged density

The zonally averaged density characterizes the zonally averaged pressure fluctuations associated with variability of the thermohaline circulation. The analysis in this section provides some support for directly associating the large-scale THC fluctuations with the zonally averaged pressure gradients. For this purpose, Fig. 3 shows EOFs 1, 2, and 3 for the Atlantic's zonally averaged *in situ* density and Fig. 4 represents the statistics of their principle components. The contours in Fig. 4 are chosen to highlight the mid-depth large-scale density gradients, which are more relevant for understanding the structure of the large-scale density fluctuations associated with THC fluctuations.

EOF-1 (representing 17% of the variance) has a roughly uniform negative density anomaly spreading from the northern part of the basin into the deep waters to the south. Down to around 3000 m depth, there is a mean north–south density gradient, with a large density in the south during this phase of the pattern. Although not significant at the 95% red noise level, the statistics for PC-1 show signs of a damped oscillation with power concentrated around 50 y, as well as an indication of the drift coming from deeper waters which was also seen in EOF-1 of the meridional stream function (Fig. 2a). This principle component shows a relatively close temporal correlation with PC-2 of the meridional stream function, with a maximum correlation near 2–5 h prior to the peak in the stream function PC-2. Hence, when stream function pattern EOF-2 is strong and positive (i.e., a weak phase of the North Atlantic overturning circulation), the zonal density pattern EOF-1 tends to be strong and positive. Density EOF-2 (representing 14% of the variance) has a more damped temporal behavior than EOF-1, yet there is again a dominance of multidecadal variability in its spectrum. This pattern contains generally uniform density across the upper 500–1000 m, with this water also penetrating into the northern sinking region of the model. The deep waters to the south are generally uniform and of opposite anomaly sign. This pattern is roughly in quadrature (i.e., 90° out of phase) with EOF-1. Note that by computing the EOFs for the North Atlantic (not shown) rather than the full Atlantic, and statistics for both of these EOFs can be made somewhat more stationary and oscillatory, suggesting that the multidecadal power seen in the patterns for the full Atlantic are manifestations of the same oscillatory variability seen in the overturning stream function EOF-2. EOF-3 (representing 12% of the variance) contains more spatial structure than EOFs 1 and 2 and it also has power concentrated in the 40–60 y time scale. The higher EOF patterns likewise contain more small-scale spatial structure but with successively less coherence in time and less power at the multidecadal time scale. An interesting picture of the fluctuations in the density field can be seen through animations of EOFs 1, 2, and 3 (not show). What is apparent in these animations, and to a large extent also in animations of the full anomalous density field, is that the multidecadal variability manifests in this field as fluctuations in the large scale zonally averaged north–south density gradient, which is predominantly represented by EOFs 1 and 2. Correspondingly, the variability of the meridional stream function is closely associated with these fluctuations, with a roughly 2–5 y lag. It should be noted that such a picture was also indicated by the regression analysis of DMS1 (see their Fig. 8 and 10).

Since the work of Stommel (1961), many have pointed to the association between the zonally average north–south density gradient and the long-term (order of centuries) stability and variability properties of the Atlantic circulation. Tziperman et al. (1994), Hughes and Weaver (1994), and Rahmstorf (1995), show recent results to this effect. The assumption that the meridional circulation is proportiona to the north–south density gradient was also the basis for the meridional-depth plane thermohaline box model of Griffies and Tziperman (1995) previously discussed in Sect. 1. This box model study indicated that the temperature and salinity phase relations exhibited by the coupled model's circulation oscillations can be captured quite well from such a simplified meridional-depth perspective. However, as Greatbatch and Peterson (1996) indicate, one may question whether the assumption of the near direct relation between meridional circulation and

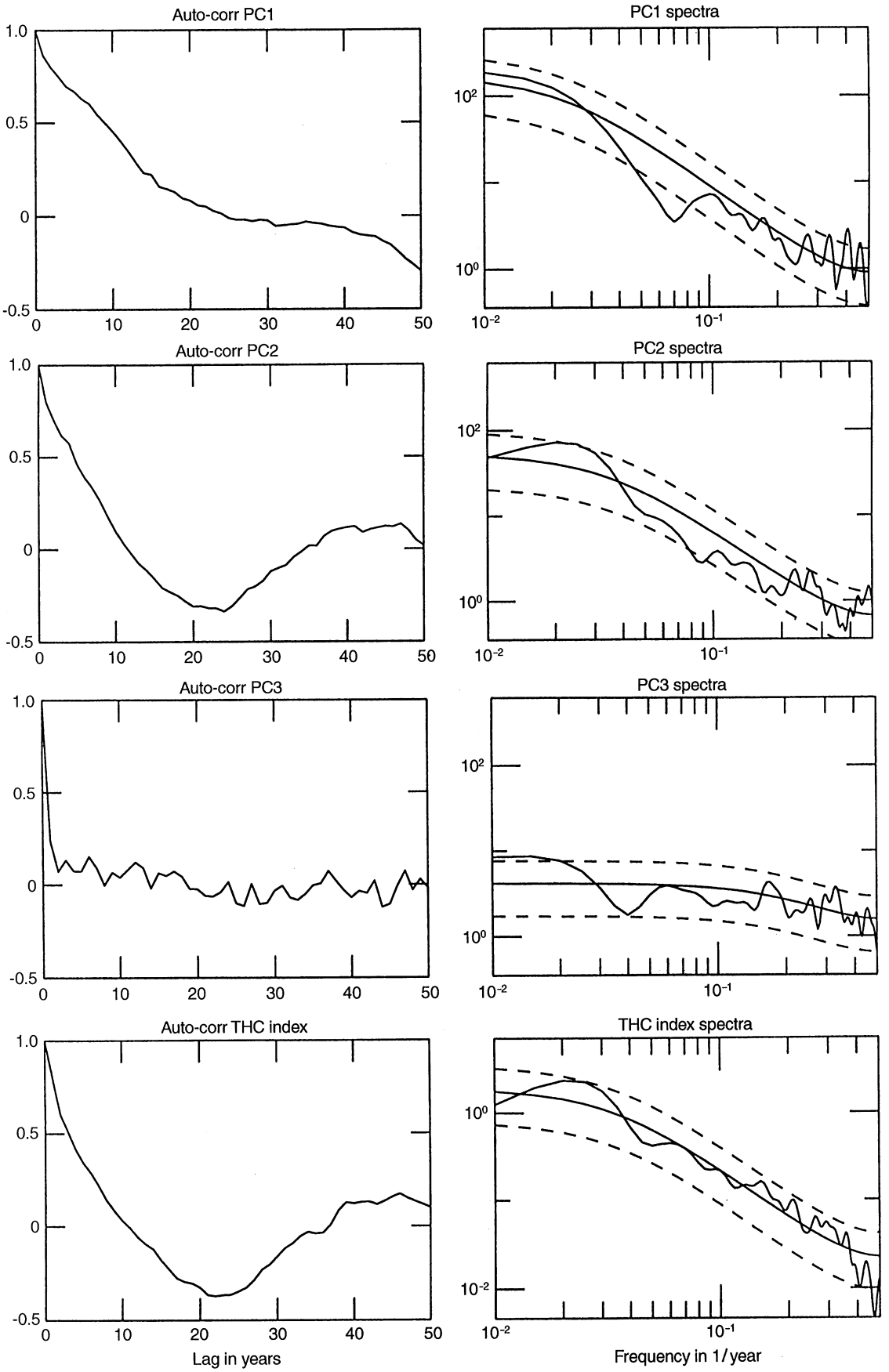


Fig. 2. Auto-correlation functions (lag units in years) and spectra (frequency units in years^{-1}) for meridional stream function principle components 1, 2, and 3 from model years 1–200. Also shown are the corresponding statistics for the THC index

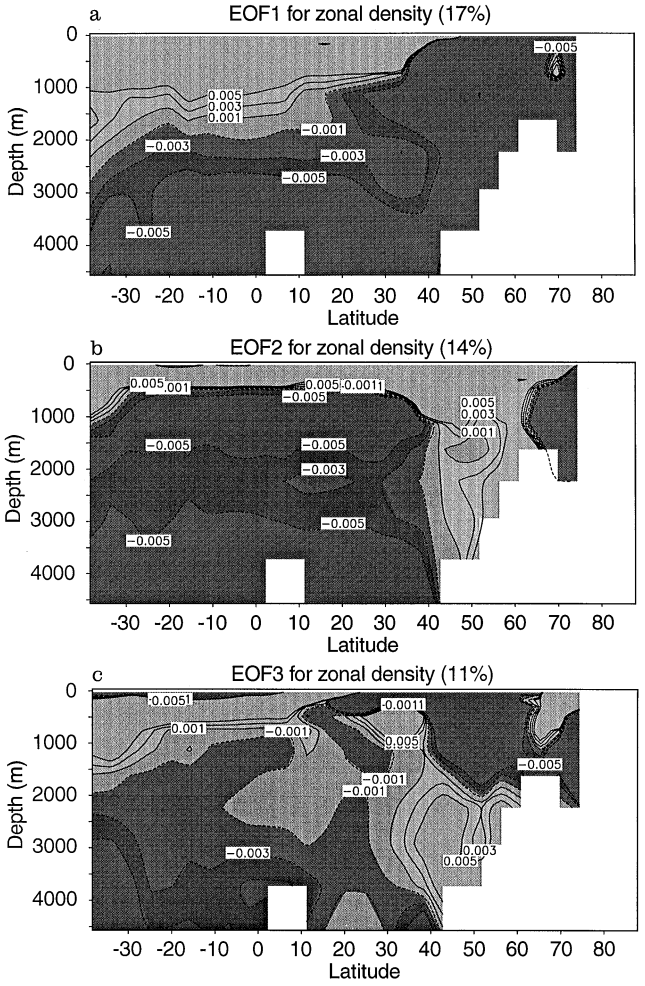


Fig. 3. a Atlantic zonally averaged density EOF-1 (representing 17% of the variance), **b** EOF-2 (representing 14% of the variance), and **c** EOF-3 (representing 11% of the variance) from annual mean fields over model years 1–200. The *light shading* indicates positive anomalies and the *dark shading* is for negative anomalies. Units are dimensionless

north–south density gradients will manifest itself on the multidecadal time scale. Nevertheless, the results shown here, as well as the regression analysis of DMS1, indicate that there is a sizeable correlation, with a relatively small time lag (on the order of 2–5 y), between multidecadal fluctuations in the North Atlantic portion of the overturning stream function and the north–south density gradient. The horizontal gyre dynamics discussed in DMS1, and in the subsequent section, is not captured by the zonally averaged box model. These dynamics are perhaps represented by the third EOF pattern of the zonally averaged density field. Nevertheless, the close correlation between the meridional density gradient and the meridional overturning described here, strongly suggests that the box model is a useful and relevant encapsulation of the leading order zonally averaged thermohaline dynamics active in the coupled model.

3.3 Dynamic topography

The dynamic topography EOF-1 and EOF-2 patterns (Fig. 5A, B) account for 21% and 17%, respectively, of the variability in the annual mean North Atlanic dynamic topography. These dipolar patterns bear a resemblance to those obtained by DMS1 using a regression analysis between dynamic topography and THC index (see Fig. 18 of DMS1). The statistics of the EOF's principle component time series (Fig. 6) indicate that these patterns undergo a significant multidecadal oscillation with enhanced power well above the 95% red noise null hypothesis in the 40–60 y time range. The statistics for higher EOFs contribute much less to this multidecadal power and so are not of relevance here. The maximum anomaly encompassed by EOF-1 and EOF-2 patterns are 5 cm in dynamic height. This amplitude is less than half that seen in the analysis of analogous observational data by Levitus (1990). We comment in Sect. 8 on the implications this discrepency may have on predictability.

In addition to the EOF patterns themselves, it is interesting to assess the amount of local variability which is captured by the patterns. This information provides more direct regional information necessary for assessing the relevance of predicting the EOF patterns. The temporal correlation between the EOF's principal component with the time series for each grid point provides a good measure of this relevance. Figures 5C, D shows these maps for the first and second dynamic topography EOF patterns. The maps indicate that in certain "hot" regions, which largely correspond to those regions in the EOF patterns with largest amplitude, as much as 80% of the local standard deviation, or greater than 60% of the variability, is captured by the first and second EOF patterns individually. Regions of significant overlap bring these numbers even higher. In other words, a useful forecast of these EOF patterns will be equivalent to a forecast of roughly 60–70% of the local dynamic topography variability. This result highlights the importance of quantifying predictability of the EOF patterns.

These EOFs occur temporally in quadrature. Figure 7 shows yearly maps of EOF-1 + EOF-2 over years 130–152, which is roughly one-half of a full oscillation during this part of the climatology. The initial year of this figure represents an anomalously weak overturning circulation. As seen in this figure, the dynamic topography anomaly is largest in the western portion of the basin, where there is a tendency for westward propagation from the center of the basin and then southward propagation along the east coast of North America. The dynamics seen in these figures are those whose predictability we are interested in quantifying.

3.4 Sea surface salinity and temperature

As described by DMS1 and Griffies and Tziperman (1995), both temperature and salinity effects are active components during the different phases of the model's North Atlantic oscillation. The first EOF patterns for SST and SSS defined over model years 1–200 are shown in Fig. 8A, B, and statistics for the first principle components

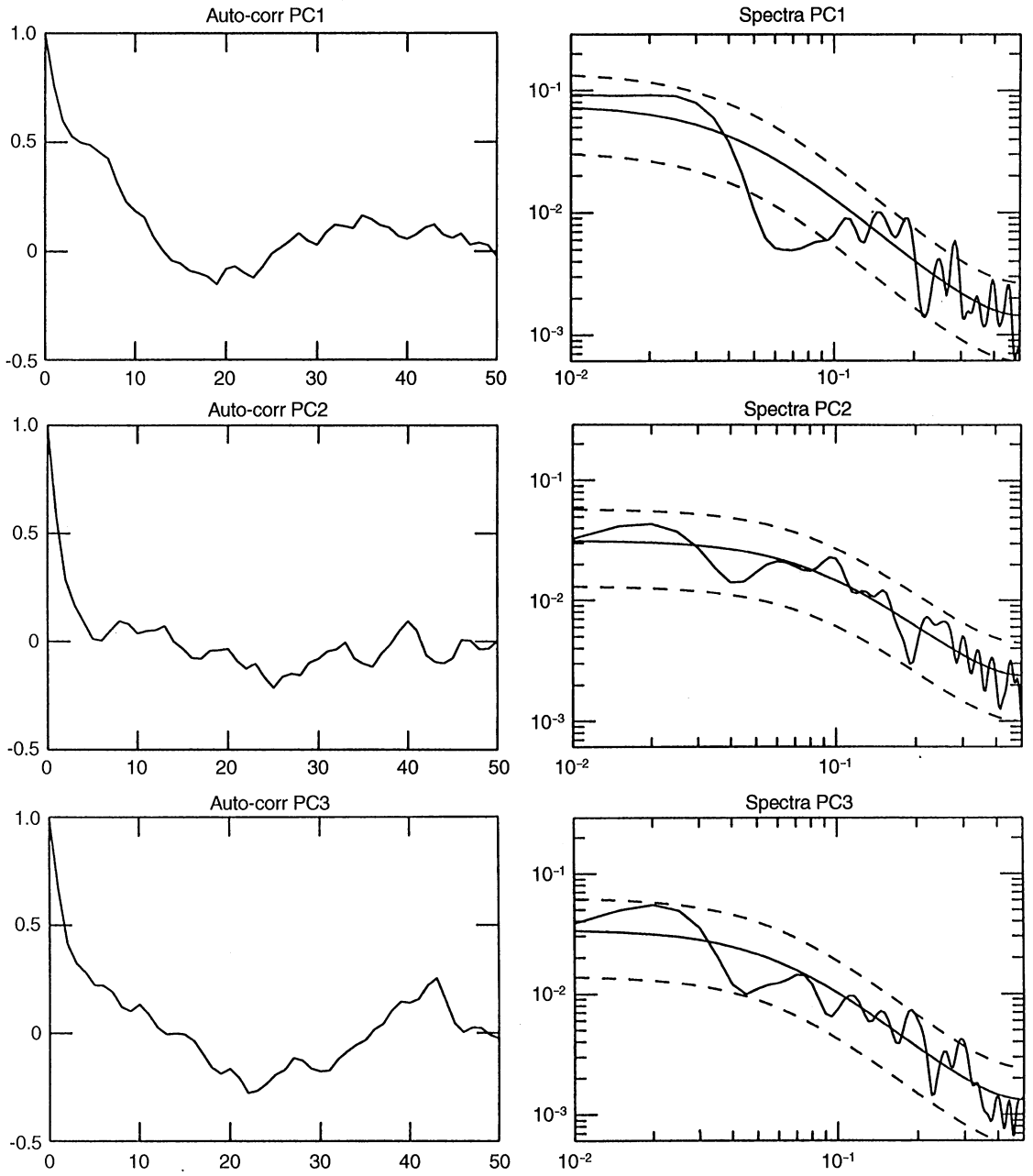


Fig. 4. Auto-correlation functions (lag units in years) and spectra (frequency units in years^{-1}) for the zonally averaged density principle components 1, 2, 3, from model years 1–200

are given in Fig. 9. EOF-1 for both fields shows a characteristic dipolar pattern with one center of action near the northeast United States extending towards Labrador, and the other center in the middle of the North Atlantic south of Greenland. Multiplied by the principle components, these patterns imply roughly 1.5 °C and 0.5 psu anomalies at their centers. Also shown in Fig. 8C, D are maps of the correlation coefficient between the principal component and the time series for the corresponding grid point. Both patterns are similar to the EOFs themselves, but with somewhat more broad spatial ranges. They indicate that the first EOF for each field accounts for up to 60% of the standard deviation in the raw fields.

The smoother oscillatory nature of SSS relative to SST shown in a comparison of the auto-correlation function in Fig. 9A and C, as well as the larger lag-one auto-correlation time (seen by looking at the fall-off of the auto-correlation function at small lags) can be attributed to the smaller amount of negative atmospheric feedbacks acting on SSS. This reduced negative feedback allows for relatively wider undamped swings for SSS with enhanced memory in its variability.

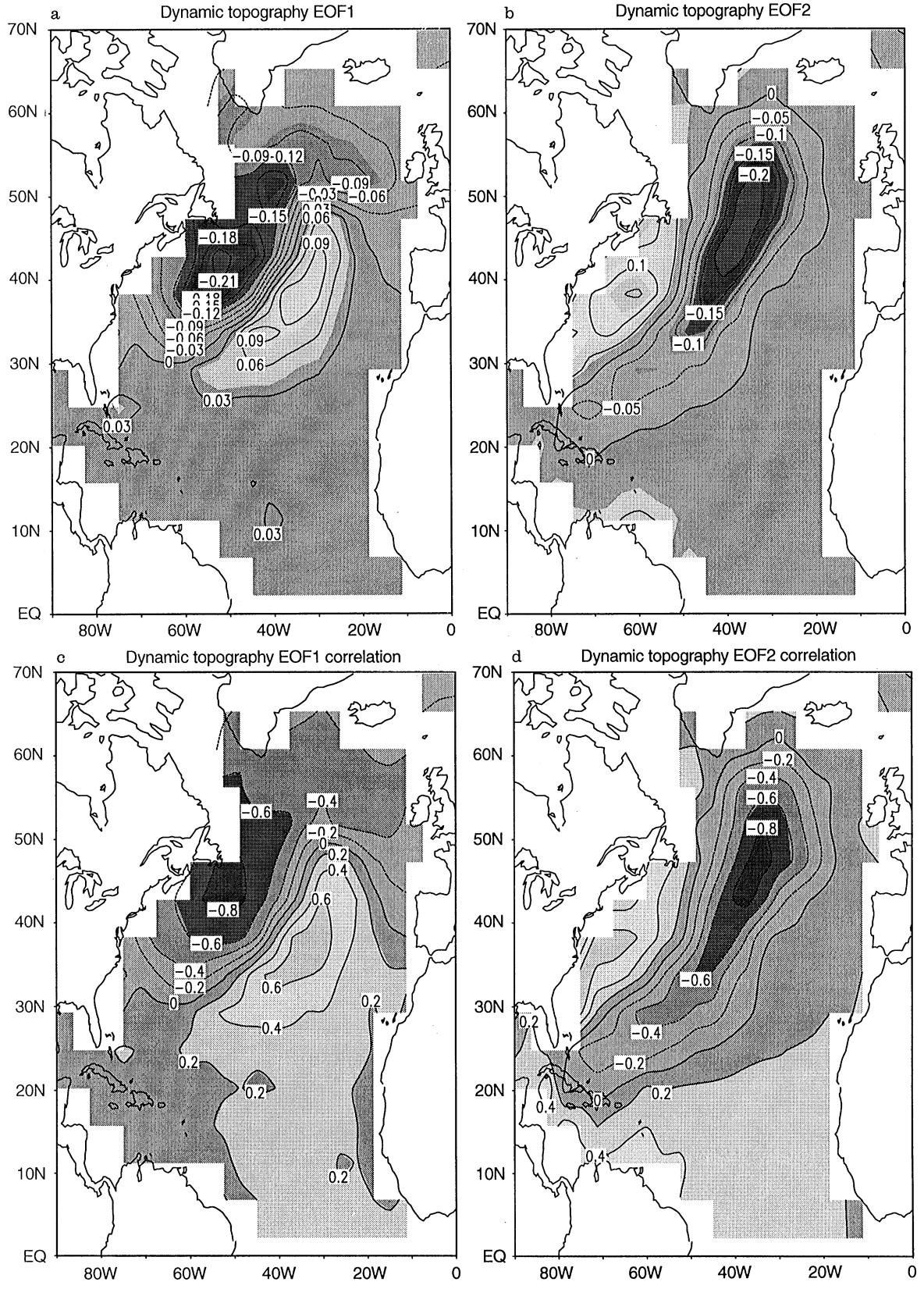


Fig. 5. **a** EOF-1 (representing 21% of the variance) and **b** EOF-2 (representing 17% of the variance) patterns for the North Atlantic dynamic topography defined from annual mean data over years 1–200. The patterns were defined over the spatial domain of 0–90 °*W*, 0–70 °*N*, for those portions of the model ocean extending to at least 1100 meters depth. **c** and **d** are the corresponding correlation maps where the time series at each grid point was correlated to that from the EOF's principal components. Units are dimensionless

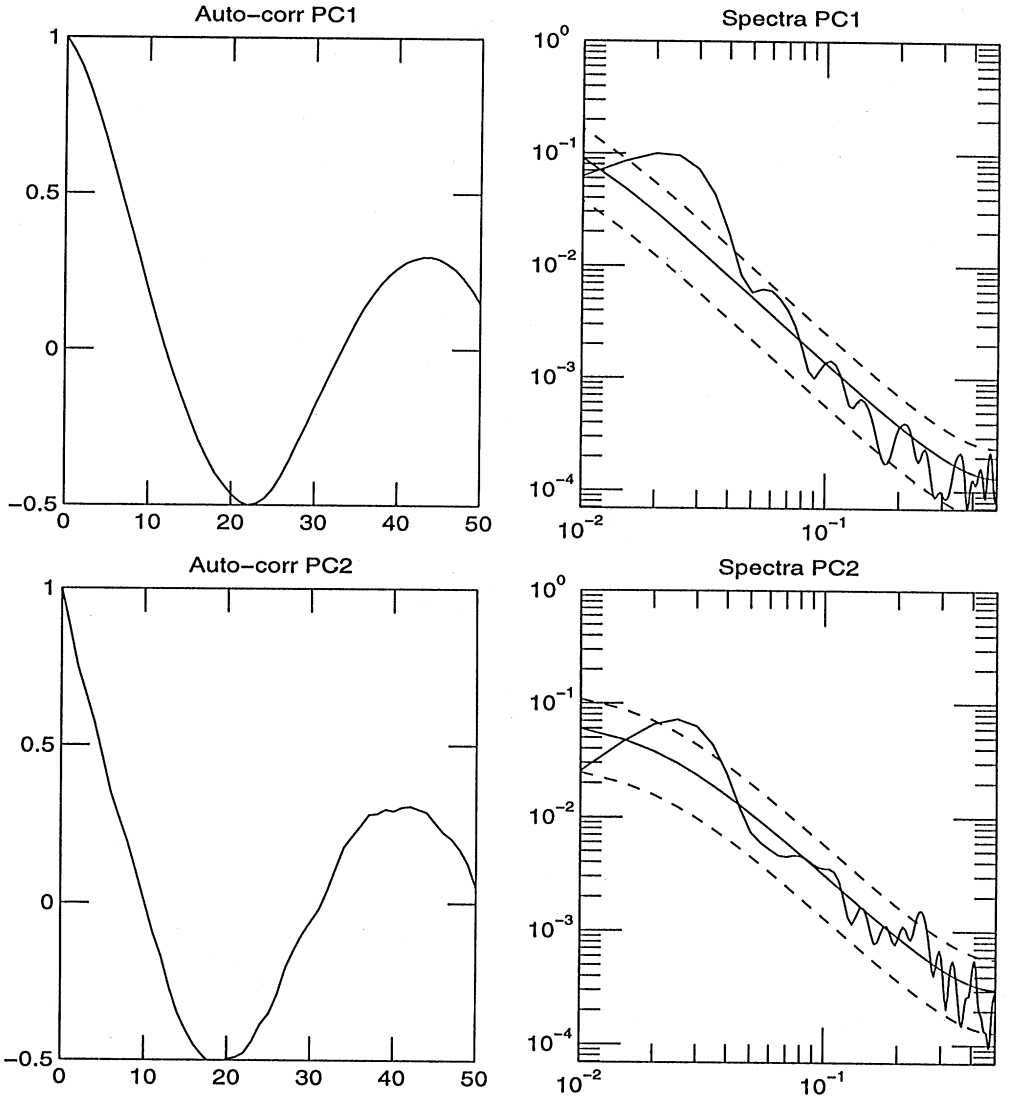


Fig. 6. Auto-correlation functions (lag units in years) and spectra (frequency units in years^{-1}) for the North Atlantic dynamic topography principle components one and two using model years 1–200

EOF-1 for both SST and SSS embody the low frequency oscillatory nature of the North Atlantic multidecadal variability. Higher EOFs for these surface fields are much more damped and have a neglible contribution to the multidecadal power. Note that the SST EOF-1 pattern bears a resemblence to that pattern seen in the analysis of observational data by Kushnir (1994). DMS1 noted this similarity through taking difference maps between model SST for a period of anomalously weak circulation and anomalously strong circulation.

3.5 Summary of the dominant patterns

In general, the signature of the multidecadal variability is much cleaner for subsurface ocean properties, which might be expected since the mechanism for the variability involves fluctuations in the thermohaline circulation. As revealed through the dynamic topography, variability is concentrated in the western half of the basin, with a tendency for westward propagation from the center of the basin and then southward propagation along the North American coast. Higher resolution models, such as that studied by Döscher et al. (1994) and Gerdes and Köberle (1995) produce patterns of response to sudden increases in the thermohaline circulation which are not unlike these EOF patterns, but are somewhat more concentrated near the western boundary. This correspondence serves as some support for the variability seen in this coarser resolution coupled model.

There are only a few patterns from the described fields which manifest the multidecadal oscillatory behavior of interest in this study. For dynamic topography and temperature at 170 m (not shown), EOF-1 and EOF-2 are the relevant patterns; for the meridional stream function it is EOF-2, and for the SSS and SST, it is EOF-1. The other patterns are either nonstationary, which means the 200 y climatology is insufficient to resolve the full extent of its variability (e.g., EOF-1 for meridional stream function) or highly damped (e.g., higher EOF patterns), and they each contribute a neglible amount to the multidecadal power. Hence, we are motivated in the following to focus on

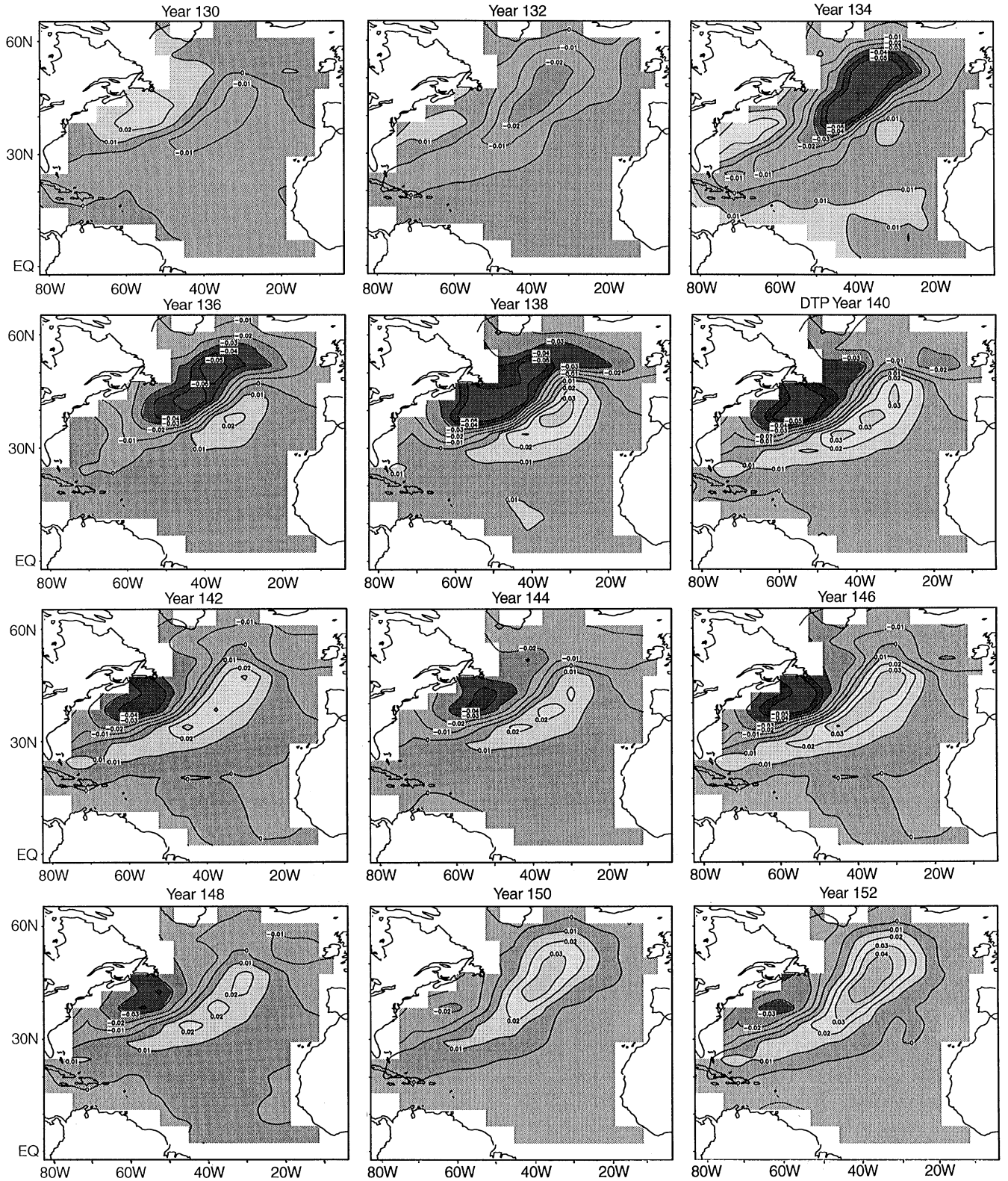


Fig. 7. Combination of EOF-1 + EOF-2 patterns for the North Atlantic dynamic topography over years 130 to 152, which represents roughly one-half of an oscillation. Every *second year* is shown. Units are dimensionless

predictability of only those few patterns which were discussed in this section.

In order to get a sense of the maximum correlations between these patterns, the maximum cross-correlations and the lag at which they occur are given in Table 1. This information provides a rough guide for assessing how useful knowledge of one principle component time series

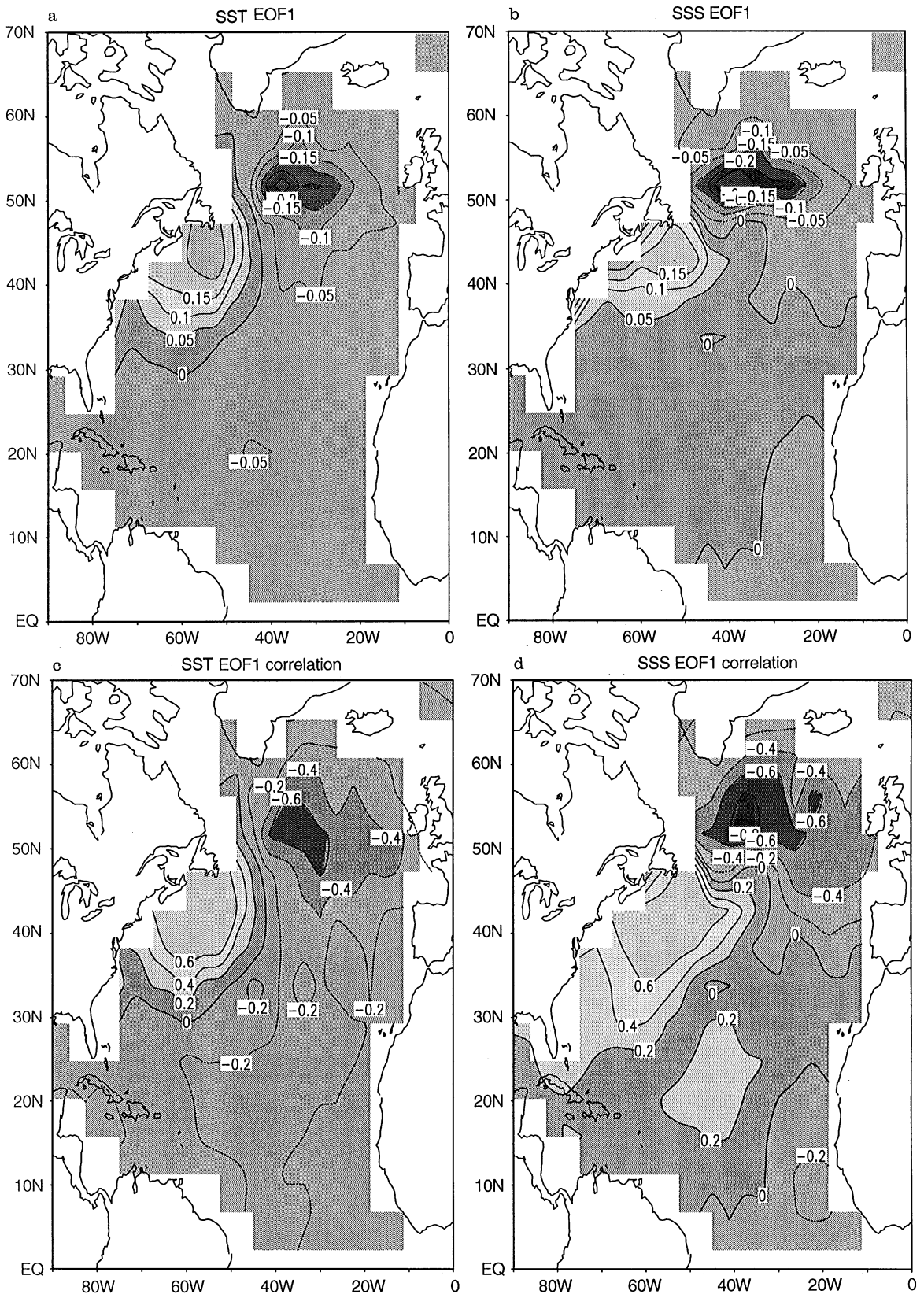


Fig. 8a–d. First EOF for North Atlantic SST **a** and SSS **b** from model years 1–200, and over the same spatial domain as the dynamic topography. **c** and **d** show the maps of the correlation between the principal components and the time series at each grid point. Units are dimensionless

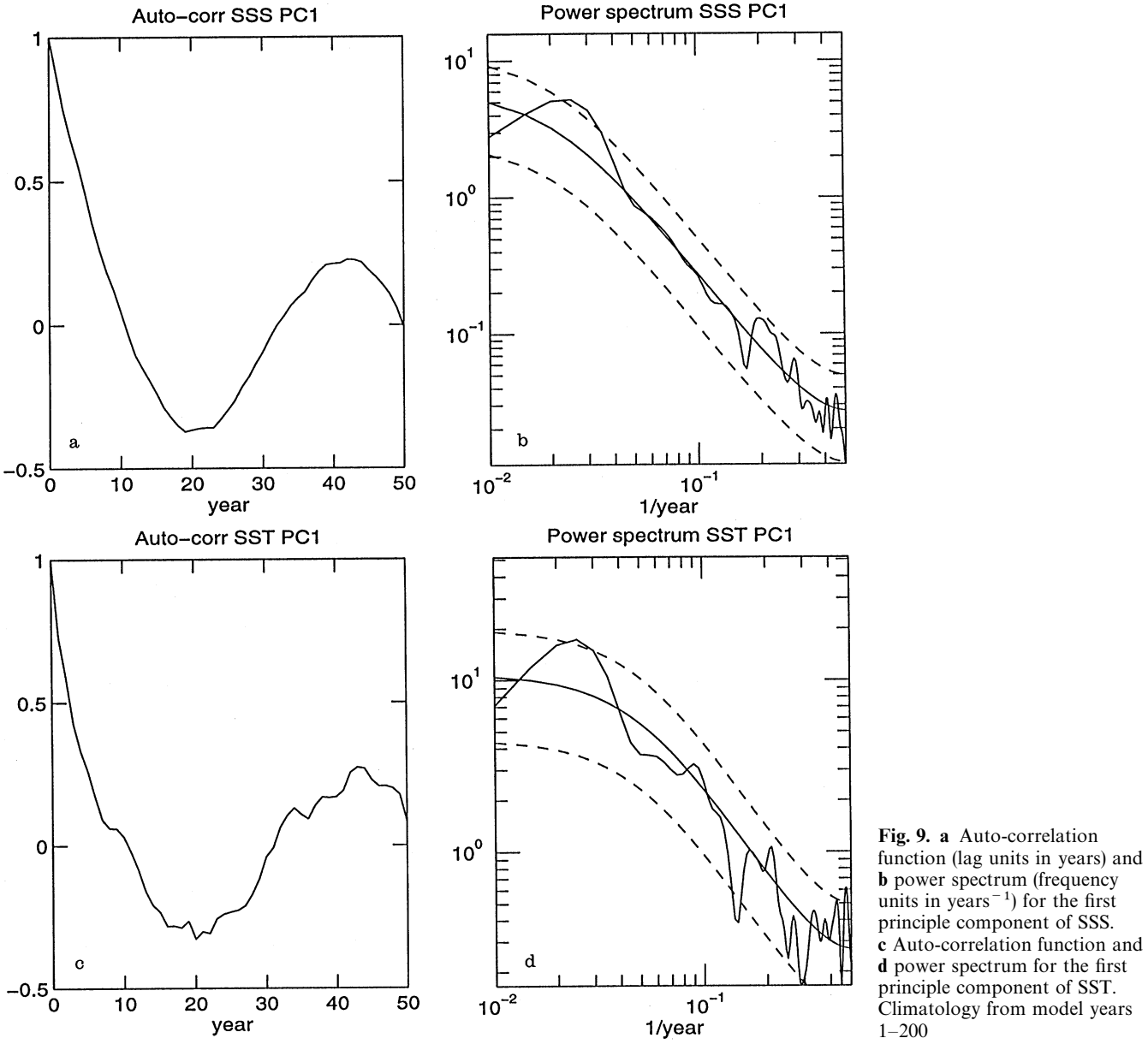


Fig. 9. **a** Auto-correlation function (lag units in years) and **b** power spectrum (frequency units in years^{-1}) for the first principle component of SSS. **c** Auto-correlation function and **d** power spectrum for the first principle component of SST. Climatology from model years 1–200

Table 1. Maximum cross correlation for chosen time series from model years 1–200

	THC	msf-2	T3-1	T3-2	dtp-1	dtp-2	SSS-1	SST-1
THC	1.0 (0)	−0.67(0)	−0.63 (4)	0.68 (6)	0.61 (−2)	−0.69 (5)	0.57 (19)	−0.53 (2)
msf-2		1.0 (0)	0.52 (0)	−0.64 (8)	0.59 (14)	0.59 (5)	−0.59 (18)	−0.52 (19)
T3-1			1.0 (0)	0.63 (−10)	−0.83 (−3)	0.73 (3)	0.74 (−2)	0.77 (−1)
T3-2				1.0 (0)	−0.75 (5)	0.76 (−4)	0.64 (9)	0.55 (11)
dtp-1					1.0 (0)	0.65 (−10)	−0.80 (2)	−0.71 (2)
dtp-2						1.0 (0)	−0.55 (13)	0.58 (−3)
SSS-1							1.0 (0)	0.75 (−1)
SST-1								1.0 (0)

Numbers in parenthesis indicate the lag (in years) where the quoted maximum cross-correlation occurred
THC, THC index, msf-2, meridional stream function PC-2; T3-1,2, temperature at level 3 (170 m) PC-1,2; dtp-1,2, dynamic topography PC-1,2; SSS-1, SSS PC-1; and SST-1, SST PC-1

might be for specifying another. For example, this table indicates that if one had complete knowledge of the amplitude of dynamic topography EOF-1, there is a -0.7 correlation with PC-1 of the SST field at two years lag. Hence, information about the dynamic topography's dominant multidecadal pattern transfers into a significant amount of information regarding the SST's dominant multidecadal pattern and thus might be a useful means of predicting this mode of the SST field.

4 Three multidecadal variability regimes

The previous section summarized the dominant patterns of multidecadal variability for the North Atlantic portion of the coupled model as seen during the control experiment years 1–200. This is the same portion of the model dataset analyzed by DMS1. It is a particularly good sample in regards to North Atlantic multidecadal variability since it exhibits coherent North Atlantic oscillations, thus providing a clean separation of the multidecadal variability from shorter and longer time scale variability. As indicated in DMS1, the variability over the course of the control experiment appears to be modulated on a century or multi-century time scale. Figure 10 provides a time series of the yearly mean, unfiltered and undetrended North Atlantic THC index for 1000 y of the control experiment. In particular, note that during the few hundred years centered at year 500, the multidecadal variability is more damped and hence has reduced power relative to the first 200 y. Such characteristics are shown in Fig. 11, which shows statistics for the first and second principle components for dynamic topography taken during years 401–600. These results indicate a highly damped oscillatory (for EOF-1) or red noise (for EOF-2) character. These statistics should be compared to the much more oscillatory behavior in Fig. 6 seen during years 1–200. In general, there is a significant reduction in the power at the multidecadal time scale for all fields, although it should be noted that the total variance is not substantially reduced; i.e., part of the reduction at the multidecadal scale is compensated by an increase in the power at shorter time scales. As the amount of multidecadal variability seen in the years 1–200 is somewhat less than suggested by the analysis of observations by Levitus (1990) (see discussion of dynamic topography in Sect. 3.3), the multidecadal variability during model years 401–600 is even more reduced relative to the observations. One can interpret this result as indicating model deficiencies and/or an insufficiency in the data record (which extends only over the second half of this century). Whatever the interpretation, the presence of secular changes in the model's variability presents an opportunity to characterize the differences in predictability which may occur within the different regimes. For this purpose, one of the four ensemble predictability experiments, ensemble A, was initialized near year 130 of the climatology, which is during the latter half of the initial oscillatory regime. A second ensemble, ensemble B, was initialized near year 500 of the climatology, which is in the middle of the highly damped regime.

There is a third regime of the model's North Atlantic multidecadal variability associated with a large anomalously negative fluctuations in the THC with maximum amplitude near year 950 seen in Fig. 10. Further analysis has revealed that the large fluctuation is associated with a substantial fresh water anomaly (roughly 1 psu spread over the upper few hundred meters and extending throughout most of the Greenland Sea) originating in the

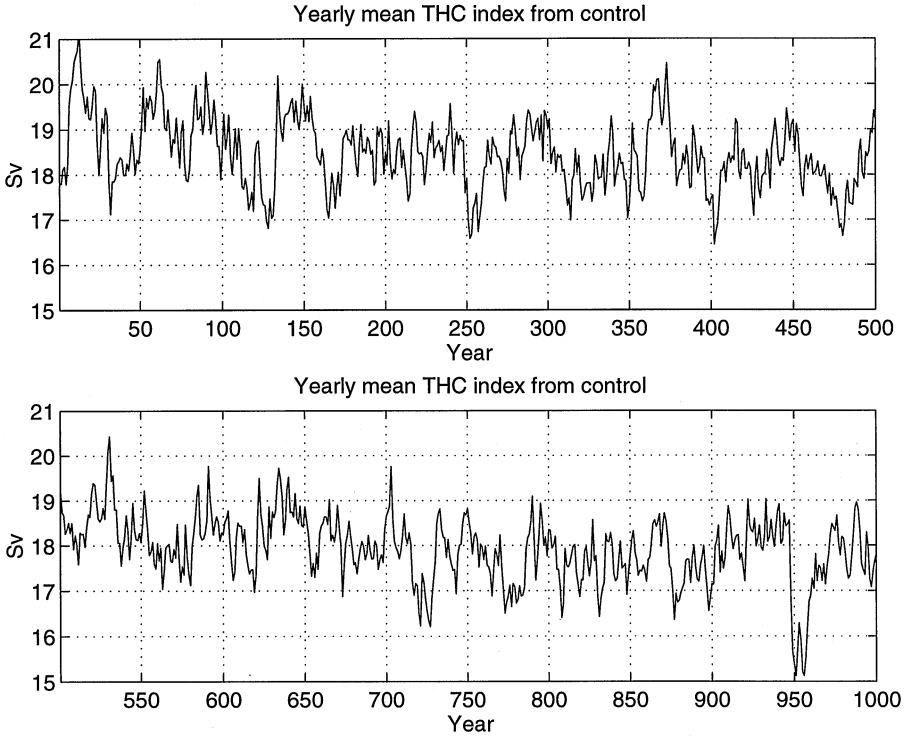


Fig. 10. Yearly mean THC index, which is defined as the maximum of the Atlantic meridional streamfunction within the latitude band of 40.5 °*N* to 72 °*N*. Shown here are the results from the control experiment years 1–1000. No filtering or detrending has been applied. For the three different climatologies, the mean and standard deviation (after a linear detrending is applied to the 200 year climatologies) are 18.7 (0.81) Sv for years 1–200, 18.2 (0.64) Sv for years 401–600, and 17.7 (0.74) Sv for years 801–1000. The ocean initial conditions for ensemble A was year 130, ensemble B used year 500, ensemble C used year 941, and ensemble D used year 951

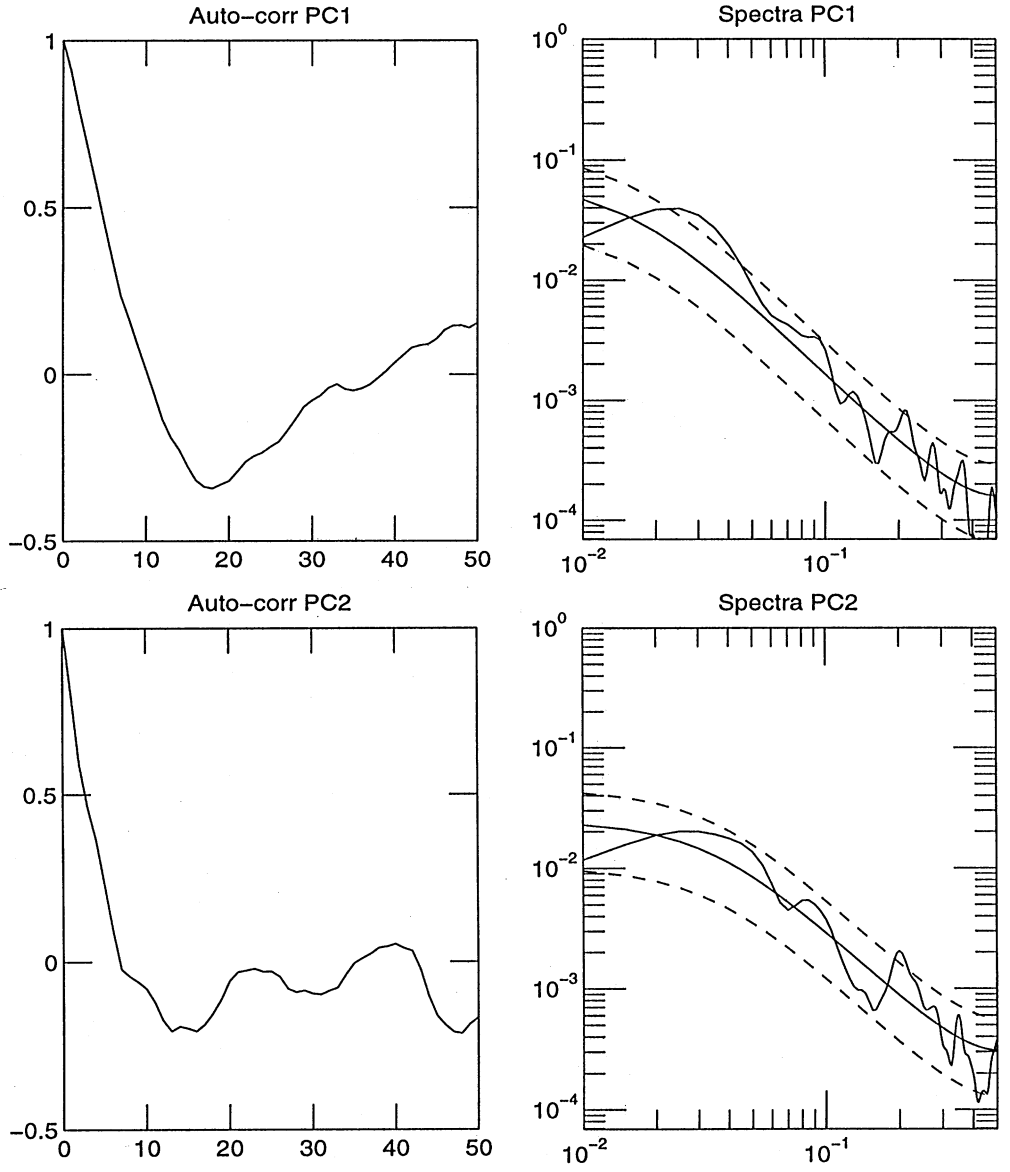


Fig. 11. Auto-correlation functions (lag units in years) and spectra (frequency units in years^{-1}) for dynamic topography principle component using model years 401–600

Arctic which advects southward in the East Greenland Current. As it enters the North Atlantic, the maximum THC decreases by roughly 4–5 standard deviations from the state before the event. The event represents an extreme example of the variability described recently by DMS2, who show such events, at a generally much reduced amplitude, occurring with a near 50 y periodicity as revealed in SST near East Greenland. DMS2 show that this high-latitude variability is associated with variability in the overlying and downstream atmospheric variability. They also indicate that the interaction between the North Atlantic variability and the high-latitude fresh water anomalies is not always as clear as it appears for the extreme event discussed here. Given the interesting dynamics of this variability, and given the extreme nature of the event near year 950 which clearly influences the middle-latitudes of the North Atlantic, we decided to investigate the predictability of the associated North Atlantic variability. For this purpose, two additional ensembles were run: one starting just prior to the effects being seen in the North Atlantic (ensemble C, starting at year 941) and one starting at the extreme phase of the event (ensemble D, starting at year 951). Additionally, in order to assess the predictability of the high-latitude variability as distinct from that determined by the North Atlantic EOF patterns, the predictability of the SST index analyzed by DMS2 will be presented in the following. The SST index represents a spatial average over a region next to East Greenland (15–25 °*W* at 70 °*N*) and was shown in DMS2 to be a good index for high-latitude multidecadal variability.

Predictability deduced from ensembles depends on the climatology used to normalize the statistics and to define the dominant patterns. Based on our interest in understanding the predictability within the different regimes of multidecadal variability discussed, we employ three different 200 y climatologies: years 1–200 for ensemble A, years 401–600 for ensemble B, and years 801–1000 for ensembles C and D. The EOF patterns will likewise be defined from within these different 200 y blocks. The

dominant EOF patterns from each of the climatologies are generally of similar nature, which allows for the direct comparison of the predictability times deduced from the different ensembles.

We note that the physical or numerical mechanisms accounting for the secular variability remain to be understood and will not be addressed here. Furthermore, we will not be addressing the predictability of the shifts between the regimes of ensembles A and B, since to do so would require substantially longer ensemble integrations than that used in this study.

5 Predictability with active THC variability

Ensemble A has 12 coupled model integrations, each of 30 y length. The oceanic initial condition is taken from year 130 of the climatology, while the atmospheric initial conditions are chosen from various January 1's for random years in the control integration. Year 130 occurs during an anomalously weak phase of the circulation, with a meridional circulation index near an extremum roughly twice the standard deviation below the mean (-1.5 Sv anomaly from the 18.7 Sv mean over years 1–200). Such a value of the index is recovered roughly 40 y later, which indicates that this particular realization of the oscillation has period near 40 y. The variability over roughly half of this oscillation are shown in Fig. 7. The climatology used for quantifying the predictability of this ensemble is that of years 1–200, which is relatively oscillatory phase of the circulation's variability.

5.1 Predictability of the dynamic topography

The principle component trajectories for the dynamic topography EOF-1 pattern is shown in Fig. 12. Also shown in this figure is the ensemble variance, normalized by that of the climatology. The trajectories and statistics for EOF-2 are similar. Higher EOF patterns, however, show a significantly faster variance growth, which is consistent with these higher order patterns having negligible predictability. The ensemble trajectories for EOF-1 remain tightly packed out to the decadal time scale, reaching 50% of the climatological variance after 16 y. This result indicates that nearly 1/2 of the 40 y oscillation indicated in Fig. 7 is predictable.

Also shown in Fig. 12 is the 95% red noise null hypothesis. The coupled model ensemble variance is seen to grow somewhat slower than the least damped of the red noise processes (i.e., the red noise process with the longest autocorrelation time) which could potentially account for the variability of the climatology. Investigation of the autocorrelation function and spectrum for this field (Fig. 6) leaves little doubt that the pattern is exhibiting a damped oscillatory behavior at the multidecadal time scale, not red noise. Correspondingly, the ensemble variance growth expected from a damped harmonic process (see Eq. (17) in the Appendix), which has been fit to the climatology, easily matches that of the coupled model, further supporting our interpretation of this variability as being analogous to that of a damped oscillator driven by noise (see also Fig. 3 in Griffies and Bryan 1997).

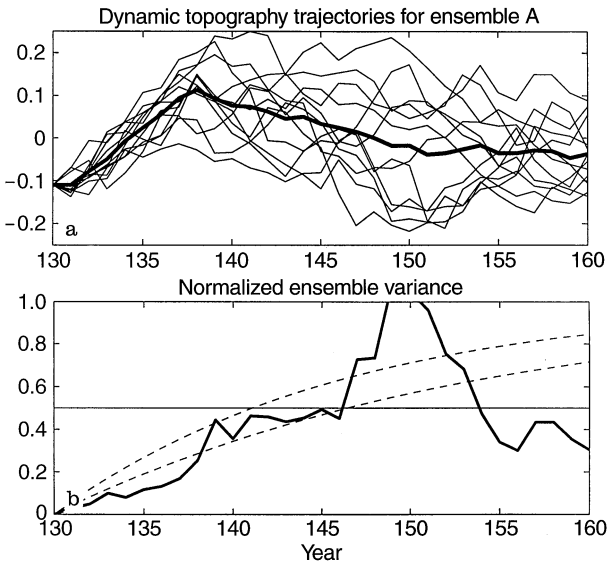


Fig. 12a, b. Ensemble A trajectories for dynamic topography's first principle component. **a** Shows the trajectories, with a *dark solid line* indicating the ensemble mean. Principle component units are dynamic meter. **b** Shows the normalized ensemble variance (*dark solid line*) along with a 95% red noise null hypothesis (*two dashed lines*)

5.2 Predictability of the 170 meter temperature

As mentioned in Sect. 3.5, intermediate depth temperature and salinity fields exhibit the two-mode gyre-like behavior consistent with that seen in the dynamic topography (Figs. 5 and 7). Although not discussed explicitly in Sect. 3, it is interesting to present the ensemble results for temperature at 170 m (model level 3). The time series for the first and second principle components have auto-correlation functions very similar to PC-1 and PC-2 of the dynamic topography, and the EOF patterns likewise look similar to those of the dynamic topography. The ensemble trajectories for the first principle component is shown in Fig. 13. The ensemble variance reaches 50% of the climatological variance only after 20 years, which is half of a full oscillation. Furthermore, the ensemble variance growth is consistently slower than the red noise null hypothesis.

One of the main purposes for presenting the results from this field is to highlight the highly coherent variability seen at modest depths associated with the model's North Atlantic oscillation. Such behavior is what may be expected from an ocean with a long memory that allows for the coherent manifestation of the THC oscillation and the correspondingly high signal to noise ratio. Furthermore, it is encouraging that such a large coherence can be found at the rather modest depths of ≈ 200 m.

5.3 Predictability of the meridional stream function

The results for PC-2 of the meridional stream function and the THC index are given in Fig. 14. Recall that the THC index in based on the maximum value of the meridional transport stream function in the North Atlantic. The ensemble trajectories for PC-2 reach 50% of the

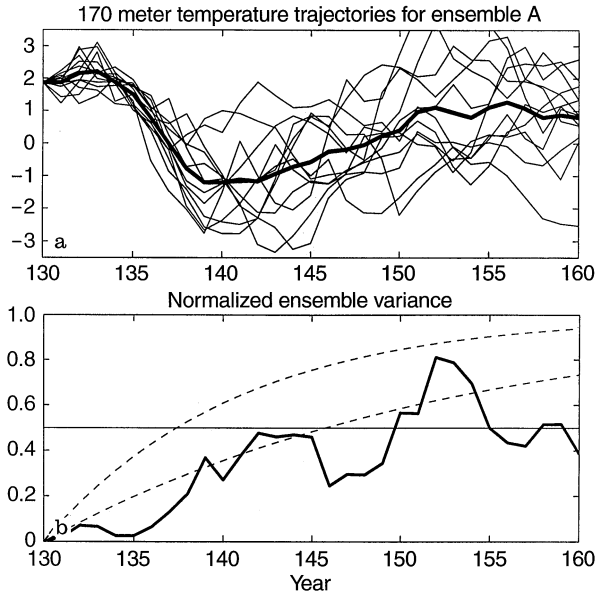


Fig. 13. **a** Level 3 temperature (170 m) first principle component trajectories from ensemble A. Principle component units are C. **b** The normalized ensemble variances along with a 95% red noise null hypothesis

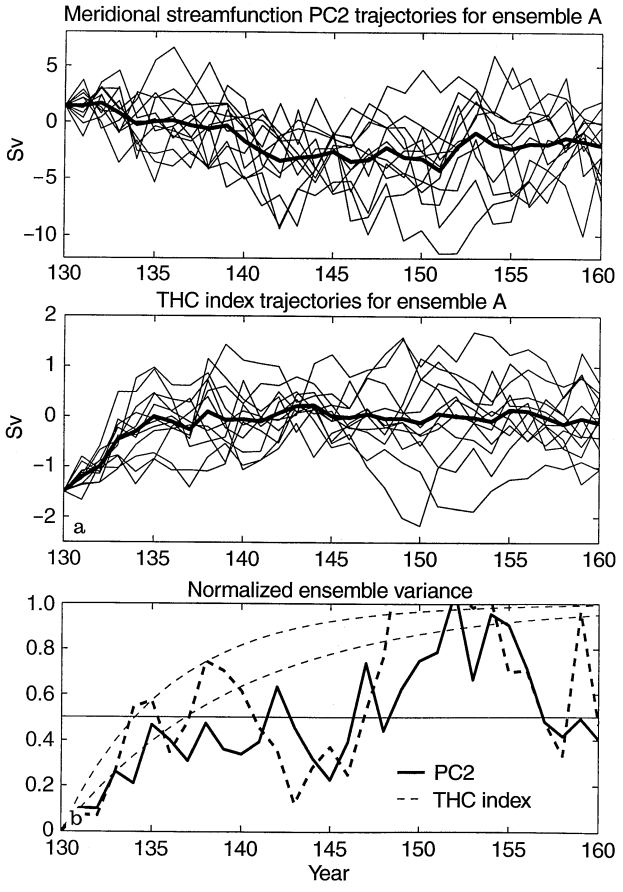


Fig. 14. **a** Meridional stream function principle component 2 and circulation index trajectories from ensemble A. Units are Sv. **b** Normalized ensemble variances along with a 95% red noise null hypothesis

climatological variance after 12–13 y, whereas the stream function index crosses after less than 5 y. Both variances show growth somewhat slower than the red noise null hypothesis at times less than 3 y. However, the THC index variance increases much faster thereafter, crossing the 50% line in about the same time as the red noise null hypothesis would suggest. We interpret this result as an indication that the THC index is significantly affected by the less predictable smaller scale variability described by the higher stream function EOFs. On the other hand, PC-2 represents the amplitude of a large-scale spatially averaged pattern whose variability is decoupled from the higher EOF patterns. Therefore, PC-2 can be expected to maintain more predictability than the THC index. This result highlights the importance of employing a suite of fields for assessing predictability as well as to point out that potential for pessimistic predictability times based on indices representing non-spatially averaged quantities. It should be noted that the reduction of the ensemble variance seen between years 10–15 might be considered significant based on the F-test. In the absence of similar behavior seen in other fields, we will not consider this fluctuation to be of importance.

5.4 Predictability of the North Atlantic SSS and SST

Figures 15 and 16 show the first principle component ensemble trajectories and normalized ensemble variance for North Atlantic SSS and SST. SST PC-1 reaches 50% of the climatological variance after around 5 y (roughly 1/8 of an oscillation period) whereas the SSS EOF-1 pattern reaches this mark after 13 y (roughly 1/4 of a period). The initial 1–2 y ensemble variance growth seen for the SST is consistent with red noise, but it slows

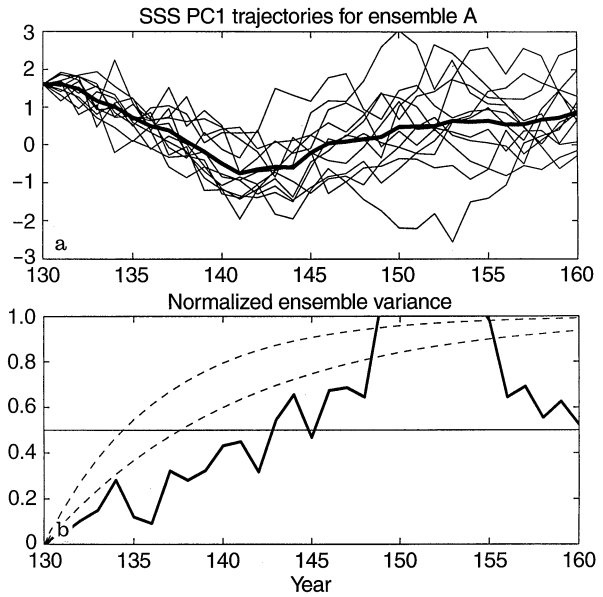


Fig. 15. **a** SSS PC-1 trajectories from ensemble A. Principle component units are psu. **b** the normalized ensemble variance along with a 95% red noise null hypothesis

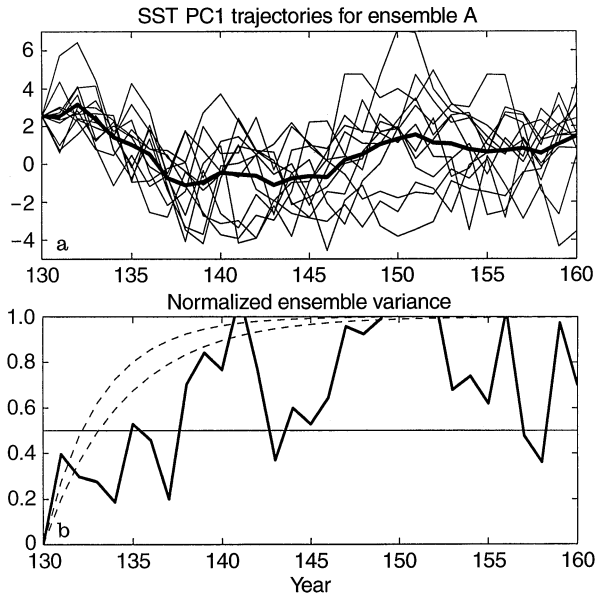


Fig. 16. **a** SST PC-1 trajectories from ensemble A. Principle component units are celcius. **b** The normalized ensemble variances along with a 95% red noise null hypothesis

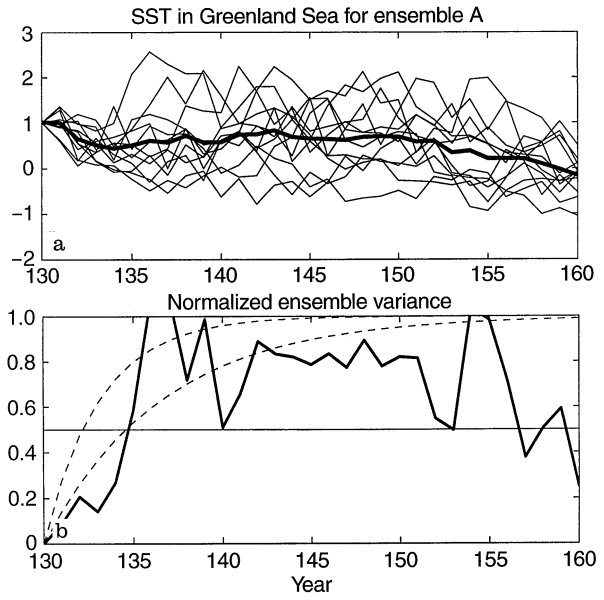


Fig. 17. **a** Ensemble trajectories from ensemble A for SST averaged over 15–25 °*W* at 70 °*N*, which is a region in the Greenland Sea. Units are °C. **b** The normalized ensemble variances along with a 95% red noise null hypothesis

somewhat, making its 50% crossing time later than red noise. This departure from red noise can be interpreted as an indication that the oscillatory component, which dominates the SSS variance growth, is finally able to overcome the fast growth of the red noise, at least for a few years. Although not showing as impressive predictability times as the SSS or other fields, this result for SST is quite important. Namely, it says that a significant amount of information about the model's dominant North Atlantic SST pattern associated with multidecadal variability is potentially predictable for 5 y.

5.5 Predictability of the Greenland Sea SST

The ensemble trajectories for the Greenland Sea SST index analyzed by DMS2 are given in Fig. 17. The ensemble variance is seen to grow rapidly and cross the 50% line at 5 y. Yet the growth has a somewhat slower rate than the red noise null hypothesis, which should be contrasted to the SST in the North Atlantic. As we will see in Fig. 19, the other ensemble experiments support the interpretation that this index is more predictable than the dominant North Atlantic SST pattern, with ensembles C and D showing predictability time scales reaching out to 10–15 y.

5.6 Space-time patterns of predictability

The previous time series analysis provides a quantification of the predictability for various large-scale patterns in the model's North Atlantic. This section briefly presents the space-time structure of the ensemble variance. These maps allow for an understanding of where the ensemble variance grows most rapidly. Besides its intrinsic fundamental interest, such information might be important for the design of observing systems.

Figure 18 shows maps taken every second year for the normalized dynamic topography ensemble variance over the full North Atlantic, including the region which defines the dynamic topography EOF patterns in Fig. 5. Numbers approaching unity indicate regions with large ensemble variance and hence loss of predictability. From early on, ensemble variance is large in the region directly south of Greenland. Soon thereafter, variance saturation occurs off eastern Greenland and then pervades the northern/northwestern part of the Atlantic.

Analysis of convection is the model reveals that the regions of early variance saturation are also regions associated with the model's deep water formation. In particular, most convection in the model's Northern Hemisphere occurs off eastern Greenland and southwestward towards to southern tip of Greenland. Convection acts as a downward pathway into the ocean for atmospheric signals. Hence, it can be expected that the deep convection regions will be the first regions to transmit into the deep ocean the effects of the perturbed atmospheric initial conditions. Furthermore, convection is a highly intermittent process and therefore can be associated with high sensitivity to initial conditions and hence rapid loss of oceanic predictability, even without the added perturbations from the atmospheric initial conditions.

The series of pictures in Fig. 18 suggests that once the strong growth in ensemble variance is initiated in the convection regions, oceanic dynamics act to propagate these high variance or high error regions southward, through the East Greenland Current and into the midlatitudes with a tendency for trapping them along the western boundary as well as in regions of strong eastward

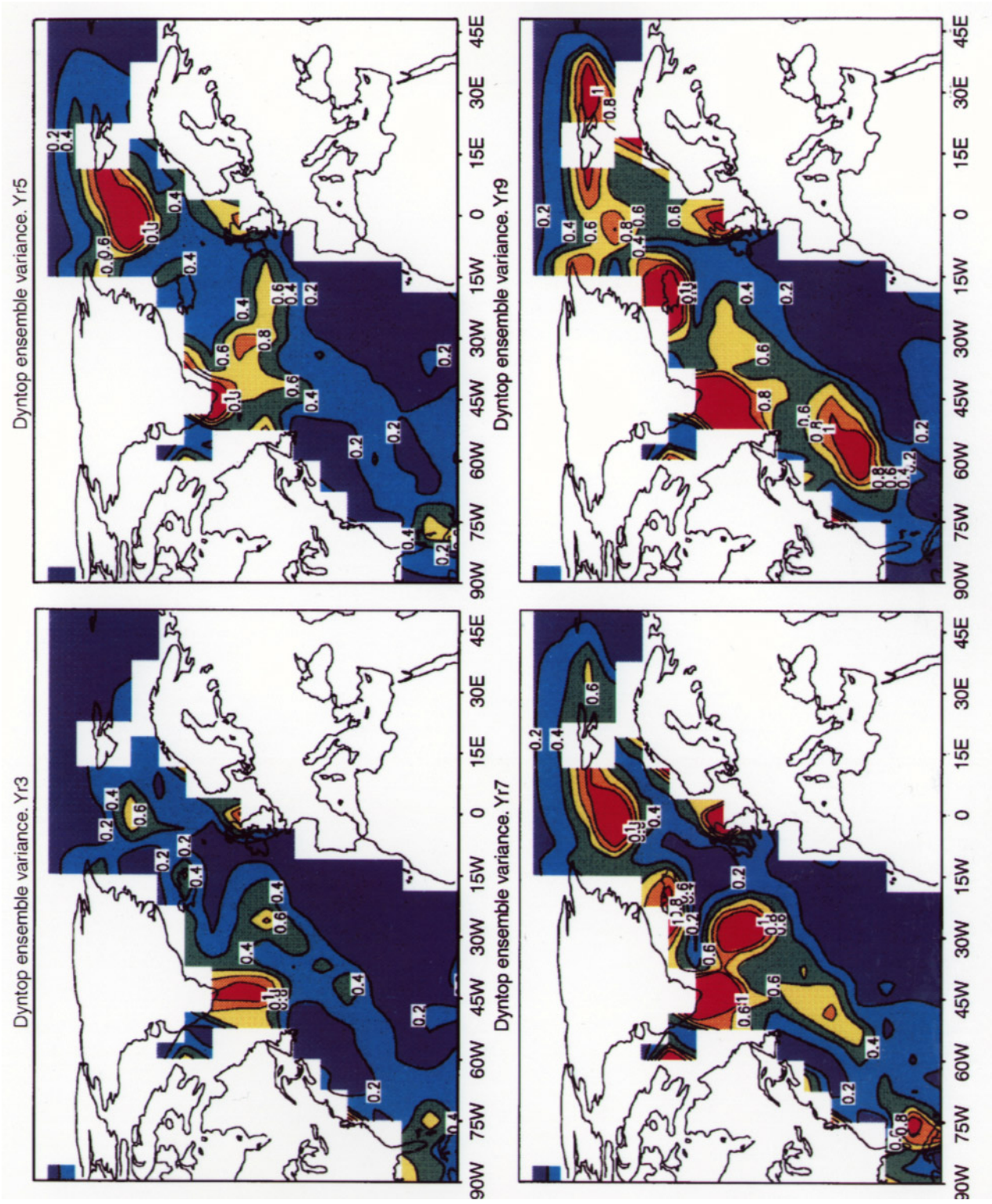
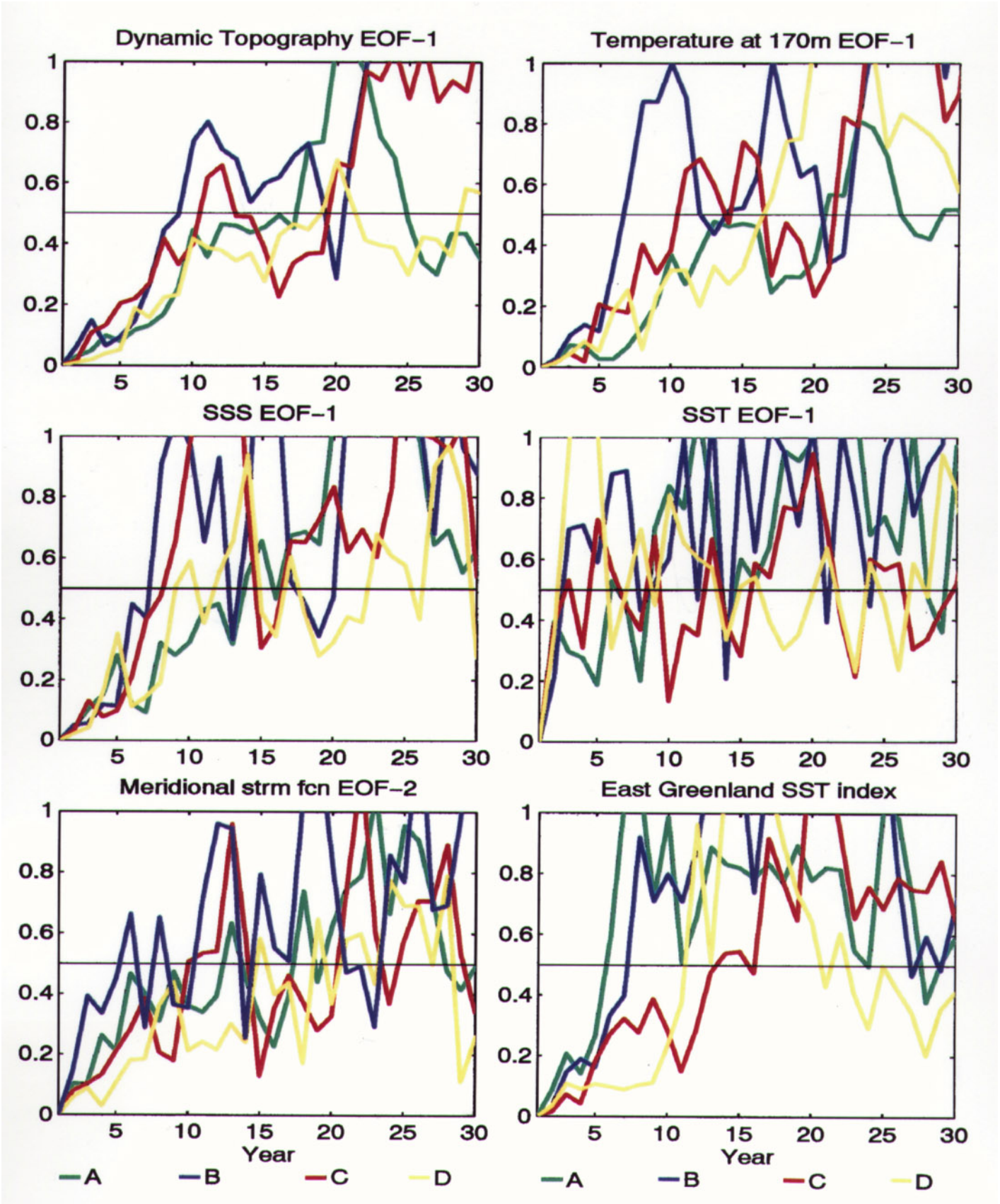


Fig. 18. Maps of the normalized ensemble variance for ensemble A's North Atlantic dynamic topography shown for years *3*, *5*, *7*, and*9* after the start of the ensemble at year 130. The *red regions* indicate regions of saturated ensemble variance, and hence rapid loss of predictability. The regions of large variance in years *3* and *5* are associated with areas of strong ocean convection. The subsequent maps indicate the transfer of this variance southward along the western boundary as well as eastward across the North Atlantic drift. Once convection initiates the strong growth in ensemble variance, ocean dynamics act to propagate these errors southward, through the East Greenland Current and into the mid-latitudes with a tendency to trap the errors along the western boundary due to the β-effect, except in regions of strong eastward flow such as the Northeastward extent of the North Atlantic Current. From Griffies and Bryan (1997)

Fig. 18.


Dynamic Topography EOF–1
Temperature at 170m EOF–1
SSS EOF–1
SST EOF–1
Meridional strm fcn EOF–2
East Greenland SST index
0
0.2
0.4
0.6
0.8
1
5
10
15
20
25
30
Year
A
B
C
D

flow such as the northeastward extent of the North Atlantic Current. The tendency for western boundary trapping and coastal propagation are another manifestation of the importance of dynamics on predictability. The errors also grow as they propagate due to the effects of further atmospheric forcing. Eventually, there is a basin-wide saturation of variance (beyond year 20) and hence total loss of predictability for the field. It remains a topic for future studies to assess whether mesoscale oceanic eddies, which are absent in this model and which would act as an additional source for rapid error growth, will significantly alter the error growth and propagation scenario suggested from the present model.

5.7 Summary of the ensemble variance results

As a summary of the results from this section, as well as to allow for a direct intercomparison between the different ensemble experiments, Fig. 19 shows the normalized ensemble variance for each of the time series from the four ensembles. The main point about how the oscillation can increase the predictability above red noise has been made with the detailed analysis of ensemble A. Therefore, to keep the figure uncluttered, no red noise bounds are included in Fig. 19. This figure constitutes the main quantitative results from this study and will be referred to further in the subsequent sections.

6 Predictability with weak THC variability

Ensemble B has 12 experiments initialized with the ocean component at year 500, which is a year near the peak of an anomalously strong North Atlantic circulation (+1.5 Sv anomaly from the 18.2 Sv mean over years 401–600). As discussed in Sect. 4, the climatology defined by years 401–600 exhibits highly damped multidecadal variability relative to that associated with years 1–200. The large damping suggests that ensemble B will show a reduced amount of long term predictability relative to ensemble A. Indeed, this is the case as can be seen from the ensemble variance plots shown in Fig. 19. Consistently, the fields from this ensemble show the shortest or close to the shortest predictability time scales for the North Atlantic patterns. However, it shows a slight increase in predictability for the Greenland Sea SST index compared to ensemble A. In general, there is not much multidecadal variability in the North Atlantic of an oscillatory character to predict, and so the predictability times are determined largely by the purely damped dynamics of the Hasselmann (1976) model.

Fig. 19. Ensemble variances for selected principle component time series as well as the Greenland Sea SST index. Shown are the variances (ordinate), normalized by its associated climatology, versus time from the four coupled model ensemble experiments *A*, *B*, *C*, *D*. Given are results for the dominant North Atlantic patterns for dynamic topography, which are summarized by the first and second EOFs, the first EOF for SST and SSS, EOF-2 for the Atlantic meridional stream function of the zonally integrated Atlantic velocity, and SST averaged over the high-latitude region in the Greenland Sea. This figure provides a summary of the main quantitative results in this paper. From Griffies and Bryan (1997)

7 Predictability with extreme THC variability

Ensembles C and D consist of eight members each. Ensemble C was initialized prior to the extreme negative anomalous North Atlantic THC and so provides a measure of the predictability of the event's onset as seen in the North Atlantic. Ensemble D was initialized at the extreme phase of the event and so provides a measure of the recovery of the North Atlantic's circulation to near its mean state. Each of the ensembles show one to two decade predictability times, more in line with that seen in ensemble A than ensemble B. However, the physical mechanism associated with this enhanced predictability is not expected to respect the simple linear oscillator character seen in ensemble A, nor the red noise behavior of ensemble B. Rather, the ensemble variance results in Fig. 19 indicate that if either the large high-latitude fresh anomaly (ensemble C) or the extreme negative phase (ensemble D) are initialized in the coupled model, its effects as seen in the North Atlantic dynamic topography and meridional stream function can be predictable for up to two decades. The multidecadal variability associated with the Greenland Sea oscillation, as measured by the Greenland Sea SST index, also possess a considerable predictability, with 50% of the variance preserved until 10–15 y. Physically, what appears to be responsible for the high predictability seen for this event is the advective mechanism bringing about the large reduction of the North Atlantic circulation. As noted earlier in the discussion of ensemble A's dynamic topography variance maps (Fig. 18), convection ordinarily leads to a rapid loss of predictability. However, for this particular event, the large negative surface salinity anomaly advecting southward in the East Greenland current is able to suppress convection and so reduces the speed at which ensemble C's trajectories spread; i.e., it reduces the sensitivity to the perturbed atmospheric initial conditions. Ensemble D, which quantifies the predictability of the recovery from the negative event, is perhaps predictable since the salinity anomaly initiating the reduced overturning is absent in the high latitudes, which means that the mean stratification relaxes the system towards its mean state within some adjustment time scale. In contrast to the results from ensembles A and B, the behavior surrounding the extreme event, especially the onset of the event, would be difficult to forecast without a properly initialized coupled model. The reason for this difficulty is that the event lies far from the typical climatological fluctuations, and so a statistical model based on typical events will necessary have a difficult time predicting extreme events.

8 Summary and discussion

Observations indicate the presence of SST variability on decadal and multidecadal time scales in the North Atlan-

tic (see CLIVAR 1995 for a recent review). Based on the original ideas of Bjerknes (1964), investigators have attributed the longer term SST variability to fluctuations in the ocean's thermohaline circulation (e.g., Kushnir 1994). Such variability is potentially important for its influence on the North Atlantic Oscillation (CLIVAR 1995), which is a mode of atmospheric circulation important for climate over the North Atlantic and Northern Europe. The North Atlantic Ocean is special in that its characteristics (e.g., geography, temperature and salinity characteristics, surface forcing, etc.) allow for multidecadal variability associated with variations of the thermohaline circulation. The coupled ocean-atmosphere model used in this study is one of the few coupled ocean-atmosphere general circulation models that simulate this type of large-scale Atlantic air-sea interaction which also includes fluctuations in the thermohaline circulation. This study quantified the predictability of simulated patterns of North Atlantic multidecadal variability using four ensemble experiments with the coupled model. The ensemble experiments used the same initial state of the ocean model yet randomly chosen atmospheric initial states as starting conditions. This design was based on an ocean-centered perspective. Namely, we wanted to see just how long the dominant oceanic patterns would maintain predictability after subjecting the same ocean state to a randomly chosen atmosphere. We know *a priori* that the synoptic scale 'memory' of the atmosphere is very short. Our experimental design is equivalent to neglecting it altogether. The perturbations implied by this change in atmospheric state are not small, especially over the storm track regions of the North Atlantic most associated with the forcing of the thermohaline circulation. It is solely because of the memory built into the large-scale ocean circulation patterns that any nontrivial yearly to decadal predictability was realized.

The first part of this work presented the model's North Atlantic variability as summarized by the dominant EOF patterns for the meridional circulation stream function, zonally averaged density, dynamic topography, sea surface salinity and sea surface temperature. The decision to focus on EOF patterns was motivated by their ability to average over space and to separate that part of the variability most relevant for the multidecadal time scales of interest. In general, we found that the dominant EOFs were also the most predictable, and so the EOFs provided a natural sorting of predictability time scales as well.

Averaging zonally in the Atlantic sector of the model, we compared the dominate patterns for the overturning transport stream function with the dominate patterns of the north–south density gradient. The results are generally consistent with the paradigm suggested by the Stommel (1961) box model of the thermohaline circulation as elaborated by Griffies and Tziperman (1995). Basically, this highly simplified description suggests an association between the model's north-south density or pressure gradients and fluctuations of the North Atlantic circulation. Fluctuations in salinity act as positive feedbacks and temperature anomalies, as well as atmospheric restorative effects, act as negative feedbacks. Atmospheric forcing, with a very small correlation time relative to that of large-scale ocean processes, provides the noisy forcing with power over a very broad time scale. It should be noted that the analysis of the zonally averaged density was presented as a conceptual addition to the study to help motivate the box model perspective. It turned out that these patterns show only modest predictability (analysis not shown) and so they were not discussed in the subsequent portion of the study.

Useful as the meridional-depth plane perspective is, it is quite incomplete. For example, it does not take into account the effects of rotation. As a complement, a horizontal perspective was presented which pointed to the relevance of a gyre-like circulation clearly seen in the dynamic topography as a cyclonically rotating dipole. This pattern is concentrated in the western part of the North Atlantic and is well captured by the first and second dynamic topography EOF patterns. The surface manifestation of the multidecadal variability was seen in a single SST and SSS EOF mode. Table 1 provided the maximum cross-correlations between the various patterns, which presents a rough indication of how useful knowledge of one time series would be for determining the other. Further study is required to fully bring out such relationships.

The predictability time scales for ensembles A and B differed significantly. The first ensemble was initialized during a period when the multidecadal thermohaline oscillations were particularly active. The predictability for the dominant EOF patterns were well described by either the damped harmonic process, or a combination of the damped harmonic and red noise processes. Ensemble B, on the otherhand, showed a highly damped response, which is consistent with its initial conditions being within a portion of the model's climatology with smaller power at the multidecadal time scale as well as more damped thermohaline circulation variability. Its variability is therefore better summarized by the Hasselmann (1976) red noise model. It has predictability, but the predictability is modest compared to ensemble A. The results show that when the multidecadal power is reduced, the predictability is also reduced. These conclusions regarding the different predictability regimes were supported by a red noise null hypothesis applied to the ensemble variance. Ensemble variances from ensemble A grew more slowly than the red noise hypothesis, thus indicating greater than red noise predictability, whereas those for ensemble B (not shown) were consistent with red noise. It should be noted that since the initial oceanic condition for ensemble A was taken at a negative THC anomaly, whereas that for ensemble B was positive, we might interpret the differences in predictability as indicating some nonlinear mechanism contributing to a dependence of predictability on the phase of the fluctuation. Although further ensemble experiments will be required to test this hypothesis, the linear arguments given here do not suggest that such a phase dependence is relevant for model regimes A and B.

We have argued for the relevance of a linear stochastic perspective for summarizing the statistics from ensembles A and B and their associated climatologies. In particular, the red noise model of Hasselmann (1976) and a linear damped oscillatory model provided a good fit to the temporal behavior of the EOF patterns. Building on this description, salient mathematical details and the theoretical predictability related to these two processes are given

in Sect. A of the Appendix. The linear behavior implies that it is sufficient to use the ensemble variance to quantify predictability. The relevance of the linear models is consistent with the amplitude of the multidecadal variability in these regions of the model's integration being rather modest and, thus within the model's "stable linear regime" (Griffies and Tziperman 1995). Therefore, it is probably sufficient to study the linear stochastic dynamics of the climatological patterns in order to assess their predictability in the coupled model from climatologies alone without making ensemble experiments. This result is highly favorable for future predictability studies using more detailed models, assuming they also exhibit such linear behavior, for which the amount of calculation may preclude running the extensive ensemble experiments conducted here. Additionally, it points to the relevance of linear stochastic models for providing a reasonable forecast of non-extreme forms of multidecadal variability.

In addition to quantifying the predictability seen in two regimes of modest thermohaline variability, this study presented the results from two ensembles initialized near an extremely large negative fluctuation in the North Atlantic THC. This fluctuation is associated with a large fresh water anomaly advecting southward in the East Greenland Current. When it enters the North Atlantic, the maximum THC decreases by roughly 4–5 standard deviations from the pre-event state. It was found that the variability surrounding both the initiation of the reduction of the North Atlantic circulation (ensemble C) and the recovery from the minimum (ensemble D) possess predictability time scales on the order of 10–20 years depending on the particular field. The linear behavior seen in the climatologies and ensemble members from the earlier portions of the control integration were not seen in the statistics from this event. Namely, the auto-correlation function for the 801–1000 y climatology (not shown) is much more damped than that from the more oscillatory period of years 1–200. Therefore, predictability times based on the linear theory applied to ensembles C and D would be significantly smaller than those actually realized. Physically, what appears to be responsible for the good predictability seen for this event is the advective mechanism associated with its cause. Namely, the convective activity, which is the central source of rapid predictability loss, is suppressed by the large fresh anomaly and so the sensitivity to the perturbed atmospheric initial conditions is reduced. In contrast to the results from ensembles A and B, the behavior surrounding the extreme event cannot be forecast using a statistical model based on climatology. Rather, a properly initialized coupled model would be necessary.

The analysis of the climatology for the Greenland Sea variability, of which the extreme event near year 950 is a particularly interesting example, was provided recently by DMS2. For all four ensemble experiments, the predictability time scales of the Greenland Sea SST index used in DMS2 showed an increase over that of the first principle component for SST in the North Atlantic domain. The time scales for the Greenland Sea SST index ranged between 5–15 y whereas the North Atlantic SST EOF-1 was predictable for roughly 5 y for all ensembles. This enhanced predictability in the high latitudes may prove to be important for enhancing the predictability of atmospheric variability since this high-latitude SST variability provides a somewhat more significant influence on the model's atmosphere (DMS2) than do the mid-latitude SST patterns (DMS1).

We also studied predictability of the Atlantic's meridional stream function in the model. Multidecadal variability in this field is shown most clearly in the second EOF pattern. This pattern has a two-cell structure with most variability in the North Atlantic part of the Atlantic basin. Additionally, this two-cell pattern's temporal behaviour was closely correlated to the THC index, which is based on the maximum overturning at any latitude in the North Atlantic sector. The predictability of the THC index, however, proved to be less than half that of EOF-2 since the index incorporated a fair amount of variability associated with small-scale 'noisy' EOF patterns. This result is not unexpected due to the definition of the THC index as a point maximum over the whole meridional-depth plane in the North Atlantic, rather than as the amplitude of a spatially filtered pattern such as that of EOF-2. The result also highlights the importance of studying a suite of patterns and indices when quantifying predictability.

8.1 Relevance for monitoring

The results of this study indicate that one of the cleanest low frequency signals of the model's North Atlantic multidecadal variability is seen in the dynamic topography. This result is relevant from an observational perspective, since dynamic topography is closely related to sea surface height (Fu and Smith 1996). Sea surface height is a field now measured directly by satellite altimeters (TOPEX/POSEIDEN, ESA-1). Indeed, satellite altimeters and hydrographic measurements by ships of opportunity already provide the foundations of a system for monitoring the sea surface height/dynamic topography of the North Atlantic. Hence, the model results suggest that dynamic topography and/or sea surface height measurements will play a central role in the diagnosis of North Atlantic multidecadal variability as well as for initializing North Atlantic climate forecasting systems. Indeed, within the limitations of the GFDL coupled climate model to simulate aspects of air-sea interaction on multidecadal time scales, this predictability study shows that monitoring of the North Atlantic ocean could be the basis for projecting climate variability in the ocean over much longer periods than previously considered possible. It is a fortunate circumstance from an observational perspective that the model shows dynamic topography has a somewhat more extended predictability time than the meridional stream function, since the stream function is not readily accessible to direct measurements.

A brief study of the spatial patterns of ensemble variance growth for the dynamic topography implicated the model's centers of convective activity as sources of rapid ensemble variance growth and consequent breakdown of predictability. Oceanic dynamics then act to propagate these errors, with a tendency to trap them along the western boundary and within the northeastward extent to

the North Atlantic Current. The exception to this scenario was seen in ensemble C, whose coupled model members contained the extreme fresh water anomaly advecting southward in the East Greenland Current. By greatly reducing the convective activity, this anomaly slowed the growth of ensemble variance and hence increased predictability even for a situation in which the North Atlantic thermohaline circulation was not as strongly oscillating as it was during years 1–200. A future predictability study which determines the model's sensitivity to oceanic initial conditions in these convective regions would lend further insight into the role convection plays in the growth of ensemble variance. In addition to convection, it should be kept in mind that oceanic mesoscale eddies, which are not resolved by this model, add yet another source of stochastic variability (in addition to the oceanic convection and atmospheric synoptic activity already contained in the model) whose effects on multidecadal variability and predictability have yet to be thoroughly investigated.

8.2 Some limitations

The ensemble experiments studied here only perturbed the initial conditions for the atmospheric component of the coupled model. We believe that this design is useful as a first step for quantifying the predictability times of the model's multidecadal variability. However, it is an incomplete study in that the importance of initializing oceanic fields was not addressed. The effect of errors in the initial state of the ocean component of the model are important from the perspective of designing observational networks. Experiments with such a focus are planned for future studies.

This study implicitly assumes that an assessment of the long-range predictability of surface and subsurface ocean circulation changes is an important goal in itself. It does not directly address the related predictability of low-frequency atmospheric variability. This is a much more difficult problem for which the present study is not entirely suitable. The difficultly is related to two issues, both of which were taken into account in our ensemble experimental design. First, there is the dominance of the very unstable and unpredictable (beyond a few days) synoptic activity on mid-latitude atmospheric variability. Secondly, it is not completely clear how sensitive low-frequency modulation of this atmospheric variability is to low-frequency variability in mid-latitude SST anomalies. In the R15 atmospheric model used here it is thought that such sensitivity is small (Held and Philips 1993; Kushnir and Held 1996; Palmer 1996) and so one may have difficulty detecting much of a predictable low-frequency signal in the atmosphere from the low-frequency oceanic forcing attributable to fluctuations in the THC. This weak coupling has also been indicated in the context of the model's ENSO-like variability seen in the tropical Pacific (Knutson et al. 1997; Chang et al. 1996). It should be noted that DMS1 did document a low-frequency atmospheric signal, albeit quite mild, associated with the model's North Atlantic THC variability as well as an atmospheric association to the SST variability in the Greenland Sea (DMS2). Future work will be aimed at assessing the possibilities for extending the oceanic predictability studied here to the overlying atmospheric fields.

It should be noted that the work of Latif and Barnett (1994, 1996) and Latif et al. (1996) indicate a more substantial atmospheric sensitivity from their coupled model, which uses the higher resolution T42 atmospheric component. Indeed, they describe fully coupled mid-latitude interdecadal ocean-atmosphere patterns of variability. Furthermore, in a coarser resolution model with a T21 atmosphere, Timmermann et al. (1996) discuss a multidecadal coupled air-sea mode of variability involving the thermohaline circulation. It would be very interesting to see what sort of ensemble predictability arises from their model and whether the linear theory employed in the current study is applicable to their variability and its predictability.

In addition to uncertainties regarding the sensitivity of the atmospheric component to mid-latitude SST anomalies, there are certain oceanic processes which are not realistically simulated in this model. For example, the model is unable to produce deep-water in a completely realistic manner through overflows from marginal seas (Dickson and Brown 1994). Also, to stabilize the model's isopycnal diffusion scheme, a nontrivial background horizontal diffusion is added which acts to cause unphysical diapycnal diffusion in regions of strong fronts (Griffies et al. accepted 1997). In particular, this diffusion acts to create a large amount of upwelling in regions of strong western boundary currents (Böning et al. 1995). Variability in this upwelling region in the North Atlantic is related to variability in the meridional stream function as manifested in its EOF-2 pattern. The absence of mesoscale eddies, with their mixing properties and added source of stochastic variability, add to the list of unresolved phenomena. Finally, uncertainties surrounding the use of flux adjustments, which are largest in the model's North Atlantic, remain a problem which coupled model developers are actively addressing.

From a climate variability perspective, the model deficiencies, including overly coarse resolution, are unfortunate since in many ways they require a large amount of dissipation in order to supply numerically stable solutions. This extra dissipation results in rather sluggish dynamics which in turn, as mentioned previously, may be responsible for the reduction in amplitude in the model's variability relative to observations (Levitus 1990). It is unclear what such an underestimate of amplitude may imply for predictability. If the lesson learned from comparing ensembles A and B can be extrapolated, then a model exhibiting larger amplitude variability might also have enhanced predictability. This is but one of the fundamental issues that must be addressed with the next generation of coupled models.

Given these caveats, within the current generation of coupled models which are capable of being integrated for the many years necessary to undertake the current study, we feel that the GFDL model is quite good at capturing certain aspects of the North Atlantic variability. It, therefore, remains our belief that the patterns described here, and their predictability, will provide a benchmark on which future studies can be based.

9 Main conclusions

The North Atlantic portion of the GFDL coupled model used here exhibits multidecadal variability related to fluctuations of the thermohaline circulation (Delworth et al. 1993). This variability is characterized by distinct spatio-temporal patterns which can be summarized by the first few empirical orthogonal functions for the meridional stream function, dynamic topography, SST, and SSS. There were three main regimes of multidecadal North Atlantic variability characterized in the model's climatology: a regime in which the multidecadal variability was highly damped, one in which it has more oscillatory, and one in which it experienced an extremely large negative anomalous fluctuation. The predictability of the dominant patterns for these regimes were quantified using ensemble experiments whose initial oceanic conditions were taken from some point within the regimes and whose atmospheric initial conditions were randomly chosen from the same day of different years within the climatology. The temporal variability of the climatology and ensembles for the damped and oscillatory regimes were well modeled by one of two linear stochastic models: red noise or the noise-driven damped harmonic oscillator. Mathematical details for both models, as applied to the predictability problem, are described in the Appendix A, and a method for bounding the ensemble variance with that expected from red noise processes is given in Appendix B.

For the oscillatory regime, the dominant pattern of dynamic topography variability showed predictability to 16 y, which is nearly half of the oscillation period; the dominant multidecadal meridional stream function pattern was predictable for 12 y, the SSS for 13 y, the SST for 5 y, and the temperature at 170 m for 20 y. The THC index discussed in DMS1 was seen to be a pessimistic index of predictability since it encompassed a nontrivial amount of variability from higher EOFs of less predictability. The predictability times for the damped regime was roughly 1/2 that of the oscillatory regime. The predictability of the onset of the extreme event was good, with patterns in ensemble C showing predictability times roughly that from ensemble A taken during the oscillatory regime. Ensemble D, which measured the predictability of the recovery from the extreme negative anomalous THC, also showed good predictability on the order of 10–15 y for the dynamic topography. The predictability surrounding the extreme event was notable as deviating significantly from that expected from the linear stochastic theory, which can be expected since the mechanism for the event relate to an extremely large fresh water anomaly advecting southward from the Greenland Sea. It represents an anomaly roughly 4–5 times the standard deviation, which is therefore far on the extremes of the climatology.

In general, the coupled model results suggest that predictability of the multidecadal North Atlantic variability is directly related to the amount of power present at the multidecadal time scale. The larger the power, the more oscillatory the variability and the more predictable it is. With damped multidecadal variability, the red noise model of Hasselmann is sufficient to model the variability, and the predictability is greatly reduced. The relevance of the linear stochastic theory for the first two ensembles, and their associated climatologies, allows for the rough quantification of predictability in other models without running ensemble experiments, assuming they exhibit variability that is well modeled by linear stochastic dynamics. This result is very important since future models, with added realism, are likely to be too expensive to be run to the extent used here. It also points to the relevance of linear stochastic models of added sophistication to the models employed here for making real-world forecasts of this variability. As a caveat to the utility of the linear models, the variability associated with the extreme event points to the importance of suitably initialized coupled models for providing useful forecasts of events lying at the extremes of the climatology. These events, as suggested by the model, present a large amount of predictability which extends towards two decades for the dynamics topography and so could be of good practical value. An analysis of spatial patterns for ensemble variance in the dynamic topography made the association between rapid loss of predictability and areas of active convection.

Analysis of the very high-latitude variability as indicated by a spatially averaged SST index in he Greenland Sea (DMS2) suggested that this variability has predictability time scales on the order of 5–10 y for modest variability regimes (ensembles A and B), and 10–15 y for larger variability regimes (ensembles C and D). These time scales are longer than the roughly 5 y for the North Atlantic EOF-1 SST. The interaction between the high-latitude variability and the North Atlantic variability is an ongoing study. For the extreme event, however, it is clear that the large fresh anomaly coming out from the Arctic was followed by a large reduction in the North Atlantic THC, hence the predictability for the high latitudes and the North Atlantic are related for this event.

Throughout this study, the ability of the model's dynamic topography to represent in a clean manner the multidecadal variability seen in the North Atlantic strongly suggests that this field, or its analog in sea surface height, such play a central role in observing multidecadal variability, and ultimately in providing practical means for its prediction. This result therefore strongly supports the ongoing efforts to develop a long-term climatology for sea surface height.

Acknowledgements. This work would not have been possible without the generous support and guidance of Tom Delworth, Ron Stouffer, and Eli Tziperman. Further beneficial discussions and advice were provided by Jeff Anderson, Bhupendra Goswami, Alex Hall, Isaac Held, Mojib Latif, Jerry Mahlman, Syukuro Manabe, Cecile Penland, and Tapio Schneider. Thanks also go to GFDL, and especially its Ocean Group, who kindly provided us with the Cray-YMP time necessary to conduct this research. Funding for SMG was provided by a fellowship from the NOAA Postdoctoral Program in Climate and Global Change and NOAA's Geophysical Fluid Dynamics Laboratory. Support for SMG and KB was also provided by Atlantic Climate Change Program funding from NOAA's Office of Global Change.

Appendix A
Predictability for two univariate processes

This section of the appendix presents a mathematical discussion of predictability realized in the red noise and

noise-driven damped harmonic oscillator processes. Although some of the material presented here (such as the auto-correlation functions and spectra) can be found in standard texts (e.g., Reif 1965; Jenkins and Watts 1968; Gardiner 1985), we know of no reference which provides the focus on predictability presented here.

A.1 Red noise process

The *red noise* or Ornstein-Uhlenbeck process $v(t)$ is described by the linear stochastic differential equation

$$\dot{v}(t) = -\alpha v(t) + \xi(t), \tag{1}$$

where the overdot indicates a time derivative. $\alpha \geq 0$ represents the effects of dissipation or damping which acts to relax $v(t)$ back to its equilibrium value $v_{eq}(t) = 0$ thus providing a negative feedback. $\xi(t)$ represents the rapidly de-correlating forces acting on the system. This noise term is modeled as a Gaussian white noise process with zero mean and auto-covariance $\langle \xi(t)\xi(s)\rangle = \sigma_\xi^2 \delta(t-s)$, where $\delta(t)$ is the Dirac delta function and σ_ξ^2 determines the power of the noise. The expectation operator $\langle \bullet \rangle$ can be considered an ensemble average where the ensemble size is infinite. Hasselmann (1976) introduced the red noise process into the climate literature as a means for describing the integrative response of a slow climate sub-system, such as the ocean, to a shorter time scale forcing, such as that provided by the synoptic scale atmosphere. More recently, such a process has been considered a useful "null hypothesis" for characterizing observed and simulated decadal to centennial climate variability (Wunsch 1992; Sarachik et al. 1996).

A.1.1 Stationary statistics. A particular solution for the red noise process can be formally written

$$v_p(t) = \int_{-\infty}^{t} e^{-\alpha(t-u)} \xi(u)\, \mathrm{d}u. \tag{2}$$

Choosing the lower limit at $-\infty$ for the particular solution allows the stationary statistics of the process to be found straightforwardly from this expression. The stationary mean $\langle v \rangle$ of the process vanishes. The stationary auto-covariance function can be found by multiplying $v_p(t)$ by $v_p(s)$ and taking an expectation. The result is

$$\langle v(t)v(s)\rangle = (\sigma_\xi^2/2\alpha)\, e^{-\alpha|t-s|}, \tag{3}$$

where the subscript "p" can be dropped since the homogeneous part of the solution does not contribute to the auto-correlation function. The covariance between adjacent points in time falls off exponentially with e-folding time $1/\alpha$. Setting $t = s$ gives the zero lag variance $\langle v^2 \rangle = (\sigma_\xi^2/2\alpha)$. The normalized auto-covariance

$$\langle v(t)v(s)\rangle \langle v^2\rangle^{-1} = e^{-\alpha|t-s|} \tag{4}$$

is called the auto-correlation function. For positive lags ($\tau = t - s > 0$), it has a slope of

$$\frac{\partial}{\partial \tau} \langle v(0)v(\tau)\rangle = -\alpha e^{-\alpha\tau}. \tag{5}$$

The slope is $-\alpha$ at zero lag, which indicates an increasing initial fall-off of the auto-correlation as the negative feedback increases.

The Fourier transform of Eq. (1) yields the frequency space solution $v(\omega) = \xi(\omega)/(\alpha + i\omega)$. Since the noise forcing is white, $\xi(\omega) = \sigma_\xi$ is a constant. The squared modulus $|v(\omega)|^2$ gives the spectral density function $S(\omega) = 2\alpha\langle v^2\rangle/(\alpha^2 + \omega^2)$. The e-folding time α^{-1} of the auto-correlation function, beyond which there is basically no correlation or memory remaining in the process, corresponds to the angular frequency $\omega = \alpha$, below which the spectrum flattens out to approximate that of a white noise process with power $2\langle v^2\rangle\alpha^{-1} = \sigma_\xi^2/\alpha^2$.

A.1.2 Optimal forecast. In the context of forecasting the red noise process, the question arises as to what is the most probable future state at times $t = \tau > 0$ given knowledge of the state at $t = 0$. As discussed by Penland (1989) in the meteorological context, since the process is Gaussian, the most probable future state of the system is also that state which is optimal in a least squares sense. This state defines the *damped persistence* forecast (see also Lorenz 1973)

$$v_{dp}(\tau) = v(0)\exp(-\alpha\tau). \tag{6}$$

This forecast damps a persistence forecast $v_{\text{persisr}}(\tau) = v(0)$ back to a zero anomaly with a relaxation time given by the system's auto-correlation time. This forecast is the same as the mean of an infinite sized ensemble whose elements each start at $v(0)$. This result supports the intuitive choice of using the ensemble mean for a forecast.

An expression for a particular ensemble element which starts at $v(0)$ can be written

$$\begin{aligned} v(\tau) &= v(0)e^{-\alpha\tau} + \int_0^\tau e^{-\alpha(\tau-u)}\xi(u)\mathrm{d}u \\ &\equiv v_{dp}(\tau) + \int_0^\tau e^{-\alpha(\tau-u)}\xi(u)\mathrm{d}u, \end{aligned} \tag{7}$$

where the white noise process $\xi(u)$ differs for each element and the damped persistence forecast (6) was identified. Hence, the mean square difference between the forecast $v_{dp}(\tau)$ and an infinite number of realizations of the process, each of which starts at $v(0)$, provides a natural definition of the error in the forecast:

$$\sigma_{dp}^2(\tau) \equiv \langle (v(\tau) - v_{dp}(\tau))^2\rangle = \langle v^2\rangle(1 - e^{-2\alpha\tau}), \tag{8}$$

where $\langle v^2\rangle = \sigma_\xi^2/2\alpha$ is the climatological or stationary variance of the process. The error is also understood as the variance in the infinite sized ensemble about its mean state, as well as the covariance between the forecast and an infinite number of realizations [again, each of which starts at $v(0)$]. It is this measure of the forecast error or its generalizations to higher order linear processes, and its approximation from finite ensembles, which is appropriate for measuring the predictability times for an ensemble forecast of linear Gaussian noise driven systems.

The same error calculation for the persistence forecast $v_{\text{persist}} = v(0)$ yields the mean squared error

$$\sigma_{\text{persist}}^2(\tau) = \langle (v(\tau) - v(0))^2\rangle = 2\langle v^2\rangle(1 - e^{-\alpha\tau}), \tag{9}$$

which saturates to twice the climatological variance yet at a slower rate than the optimal forecast. It is straightforward to show that the error in the persistence forecast is always larger than that of damped persistence, as we would expect since pure persistence uses less information than damped persistence. The slope of the optimal forecast error is given by

$$\frac{d\sigma_{dp}^2}{d\tau} = 2\alpha\langle v^2\rangle e^{-2\alpha\tau}, \tag{10}$$

which indicates an exponentially decaying and positive slope starting at $2\alpha\langle v^2\rangle$. It should be noted that filtering a red noise process will reduce this initial error slope, as expected from there being an increase in persistence for filtered signals (Munk 1960). Thus, care must be exercised when assessing predictability of time filtered processes.

Some numbers are illustrative. With a subjectively defined level of forecast skill $\gamma \equiv \sigma_{dp}^2/\sigma_{\text{climate}}^2 = [1 - \exp(-2\alpha\tau_\gamma)]$, the time τ_γ beyond which the optimal forecast loses its skill or useful information, which serves to define a *predictability time*, is given by $2\alpha\tau_\gamma = -\ln(1-\gamma)$. For example, with $\gamma = 0.5$, $1/\alpha = 5$ y, $\tau_\gamma = 1/(2\alpha)\ln 2 = 1.7$ years. For the less stringent level $\gamma = 0.75$, the useful lead time is $\tau_\gamma = 3.5$ y. The equivalent calculation for the persistence forecast, whose error is given by Eq. (9), yields $\alpha\tau_\gamma = -\ln(1-\gamma/2)$, or a 0.5 error crossing time of 1.4 y and a 0.75 error crossing time of 2.4 y.

A.1.3 Comments. In general, for the linear noise driven systems discussed in this appendix, the absolute amplitude of the white noise forcing does not affect the time at which the system loses its predictability. Rather, for the red noise process, it is the de-correlation time α^{-1} which is relevant. The equivalence of the damping time for the ensemble mean with the correlation time for a climatology indicates the extent to which an arbitrary system is acting like a red noise process. Likewise, the ensemble variance, or equivalently the optimal forecast error, saturates to the climatological variance with e-folding time half the de-correlation time (eq. 8).

A.2 Damped harmonic process

The equation describing the *damped harmonic process* is given by

$$\ddot{x}(t) + \omega_0^2 x(t) = -2\beta\dot{x}(t) + \xi(t). \tag{11}$$

The parameter $\beta \geq 0$ represents the contribution of frictional or viscous dissipation, the frequency ω_0 is the natural frequency of the noise free and dissipation free system ($\xi(t) = \beta = 0$), and the white noise forcing $\xi(t)$ is like that in the red noise example. Setting $\beta^2 < \omega_0^2$ allows the noise free system to exhibit a damped harmonic response. For the opposite inequality, the motion is overdamped. Solutions for the overdamped case can be found from the following underdamped solutions through analytic continuation.

A.2.1 Stationary statistics. A particular solution to (11) is

$$x_p(t) = \int_{-\infty}^{t} \frac{e^{-\beta(t-u)}}{\Omega} \sin[\Omega(t-u)]\, \xi(u)\,du, \tag{12}$$

where $\Omega^2 = \omega_0^2 - \beta^2$. The auto-covariance function is given by

$$\langle x(t)x(s)\rangle = \langle x^2\rangle e^{-\beta|t-s|}\left\{\cos(\Omega|t-s|) + \frac{\beta}{\Omega}\sin(\Omega|t-s|)\right\}. \tag{13}$$

Setting $t = s$ yields the zero lag variance for the process $\langle x^2\rangle = \sigma_\xi^2/(4\beta\omega_0^2)$. For positive lags $\tau > 0$, the slope of the auto-correlation function takes the form

$$\frac{\partial}{\partial\tau}\langle x(\tau)x(0)\rangle = -\langle x^2\rangle\omega_0^2\Omega^{-1}e^{-\beta\tau}\sin(\Omega\tau), \tag{14}$$

which vanishes at zero lag. This slow fall-off should be contrasted to that seen in the red noise process in Eq. (5). It will play a role in determining the behavior of the ensemble variance for this process, and hence for determining its predictability time.

The Fourier transform of the auto-covariance function gives the spectrum

$$S_{ho}(\omega) = \langle x^2\rangle \frac{4\beta\omega_0^2}{4\beta^2\omega^2 + (\omega_0^2 - \omega^2)^2}. \tag{15}$$

The spectrum has a peak at the angular frequency $\omega^2 = \omega_0^2 + 2\beta^2$, becomes the constant $4\beta\langle x^2\rangle\omega_0^{-2}$ at zero frequency, and falls off as ω^{-4} for frequencies $\omega \gg \beta$.

A.2.2 Optimal forecast. Starting from some known initial state $\dot{x}(0)$ and $x(0)$, the *damped harmonic persistence* forecast

$$x_{dhp}(\tau) = \frac{e^{-\beta\tau}}{\Omega}\left\{x(0)\Omega\cos(\Omega\tau) + [\dot{x}(0) + \beta x(0)]\sin(\Omega\tau)\right\}, \tag{16}$$

is the optimal forecast of the harmonic process. Note the need for two pieces of initial state information, $x(0)$ and $\dot{x}(0)$, as well as two pieces of climatological information, the damping coefficient β and natural frequency ω_0, to construct this forecast. The mean squared error in the damped harmonic persistence forecast $\sigma_{dhp}^2 = \langle(x(\tau) - x_{dhp}(\tau))^2\rangle$ is given by

$$\sigma_{dhp}^2(\tau) = \langle x^2\rangle\left\{1 - \frac{e^{-2\beta\tau}}{\Omega^2}[\omega_0^2 - \beta^2\cos(2\Omega\tau) + \beta\Omega\sin(2\Omega\tau)]\right\}. \tag{17}$$

The error σ_{dhp}^2 oscillates as it exponentially saturates to the climatological variance $\langle x^2\rangle = \sigma_\xi^2/(4\beta\omega_0^2)$. The corresponding error for the persistence forecast $x_{\text{persist}} = x(0)$ is

$$\sigma_{\text{persist}}^2(\tau) = 2\langle x^2\rangle\left\{1 - e^{-\beta\tau}\left[\cos(\Omega\tau) + \frac{\beta}{\Omega}\sin(\Omega\tau)\right]\right\}, \tag{18}$$

which saturates to twice the climatological variance and is always larger than the damped harmonic persistence

forecast error. The slope of the optimal forecast error is given by

$$\frac{d\sigma^2_{dhp}}{d\tau} = \frac{4\beta\omega_0^2\langle x^2\rangle}{\Omega^2}\, e^{-2\beta\tau}\sin^2(\Omega\tau), \tag{19}$$

which vanishes at zero lead time. Recall the slope (10) for the error made in damped persistence forecast is $2\alpha\langle v^2\rangle$ at zero lead time. These differences in initial slopes are fundamental to why the harmonic process can provide for substantially increased predictability over the red noise process, especially at short to modest lead times. Further discussions of this point are given in Appendix B.

Some numbers are useful. Consider a harmonic process of period 40 y and decay time 40 y. Such time scales are implied by the sample auto-correlation function for the dynamic topography's EOF-1 and EOF-2 during model years 1–200 (see Fig. 6). The error in the damped harmonic persistence forecast (17) of this process takes 13 y to reach 50% of the climatological variance whereas the pure persistence forecast error (18) reaches 50% after only 6 y.

Appendix B
Red noise null hypothesis for ensemble variance

As seen in Eq. (10), the growth of the normalized red noise variance at zero lead time is 2α, where again α is the red noise negative feedback parameter. Therefore, as expected, with increased damping, the red noise ensemble variance grows faster and the process becomes less predictable. In contrast, for the damped harmonic oscillator process, Eq. (19) indicates that the ensemble variance growth rate vanishes at zero lead time which indicates the potential for enhanced predictability over the red noise process. Both of these behaviors of the ensemble variance correspond to the behaviors of the respective auto-correlation functions, with the red noise having a nonzero fall-off at zero lag (Eq. 5) and the harmonic process having zero fall-off (Eq. 14). It is relevant to pay particular attention to these short lead time behaviors since it is for such times that the F-test indicates there is significant statistical differences between ensemble variances and climatology. It follows that although a significant negative trough in the auto-correlation function is an indicator of an oscillatory component in the variability, if the small lag response is damped much as for a red noise process, the ensemble variance growth at the early lead times will be fast and there is little useful predictability above the corresponding red noise case. Therefore, the most relevant contribution that an oscillatory component can provide to the predictability problem is to slow the initial growth of the ensemble variance by enhancing the auto-correlation or memory.

These considerations motivate a red noise null hypothesis which can be tested against the ensemble variances. Namely, what is tested is whether the realized coupled model sensitivity to the initial conditions is significantly different from the best fit of the climatology to a red noise process. To form the 95% null hypothesis bounds, we estimate the decay time α^{-1} using the lag-one auto-correlation time take from the climatology. The ability of a red noise auto-correlation function to fit that from a damped oscillatory process clearly degrades as the oscillatory component becomes more dominant. As this occurs, it turns out that the fit of the red noise ensemble variance to that of the damped oscillator actually improves. What this simply means is that the predictability arising from the highly oscillatory process can be matched by that of very weakly damped red noise process. Section 5.4 of Jenkins and Watts (1968) provides formulae for obtaining confidence bounds on this estimate of α, which provides an upper and lower limit to the estimated negative feedback parameter $\alpha_{\pm}$. This range of feedback parameters then imply two different variance growth curves $1 - \exp(-2\alpha_{\pm} t)$ which can be used to place a 95% red noise null hypothesis on the coupled model ensemble variance. These bounds are placed on the ensemble variances plotted in this study.

References

Anderson JL, Stern WF (1996) Evaluating the potential predictive utility of ensemble forecasts. J Clim 9:260–269

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

Böning C, Holland WR, Bryan FO, Danabasoglu G, McWilliams J (1995) An overlooked problem in model simulations of the thermohaline circulation and heat transport in the Atlantic Ocean. J Clim 8:515–523

Bryan K, Griffies SM (1996) North Atlantic climate variability on decadal time-scales: Is it predictable? J Atmos Ocean Phys, Russian Acad Sci, Moscow (accepted for publication)

Bryan K, Hansen FC (1995) A stochastic model of North Atlantic climate variability on a decade to century time-scale. Proc Workshop on Decade-to-Century Time Scales of Climate Variability, National Research Council Board on Atmospheric Sciences and Climate, National Academy of Sciences, Irvine, CA

Bryan K, Stouffer RJ (1991) A note on Bjerknes' hypothesis for North Atlantic variability. J Mar Syst 1:229–241

Chang P, Ji L, Li H, Flugel M (1996) Chaotic dynamics versus stochastic processes in El Niño-Southern Oscillations in coupled ocean-atmosphere models. Physica D 98:301–320

Chatfield C (1989) The analysis of time series: an introduction. Chapman and Hall, 241 pp

CLIVAR (1995) CLIVAR Science Plan. World Climate Research Program. 89, 157 pp

Delworth T, Manabe S, Stouffer RJ (1993) Multidecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. J Clim 12:1993–2011

Delworth T, Manabe S, Stouffer RJ (1997) Multidecadal climate variability in the Nordic Seas and surrounding regions: a coupled model simulation. Geophys Res Lett 24:257–260

Dickson RR, Brown J (1994) The production of North Atlantic deep water: sources, rates, and pathways. J Geophys Res 99: 12 319–12 341

Dickson RR, Meinke J, Malmberg SA, Lee AJ (1988) The "Great Salinity Anomaly" in the northern North Atlantic, 1968–1982. Progr Oceanogr 20:103–151

Döscher R, Böning CW, Herrmann P (1994) Response of circulation and heat transport in the North Atlantic to changes in thermohaline forcing in northern latitudes: a model study. J Phys Oceanogr 24:2306–2320

Fu L-L, Smith RD (1996) Global ocean circulation from satellite altimetry and high-resolution computer simulation. Bull Am Meteorol Soc 77:2625–2636

Gardiner CW (1985) Handbook of stochastic methods for physics, chemistry, and the natural sciences. Springer-Verlag Berlin, Heidelberg, New York

Gerdes R, Köberle C (1995) On the influence of DSOW in a numerical model of the North Atlantic general circulation. J Phys Oceanogr 25:2624–2642
Greatbatch R, Peterson KA (1996) Multidecadal variability and oceanic thermohaline adjustement. J Geophys Res 101: 20467–20482
Griffies SM, Bryan K (1997) Predictability of North Atlantic multidecadal climate variability. Science 275:181–184
Griffies SM, Tziperman E (1995) A linear thermohaline oscillator driven by stochastic atmospheric forcing. J Clim 8:2440–2453
Hasselmann K (1976) Stochastic climate models, part 1-theory. Tellus 18:473–484
Held I, Phillips PJ (1993) Sensitivity of the eddy momentum flux to meridional resolution in atmospheric GCMs. J Clim 6:499–507
Huang RX, Luyten JR, Stommel HM (1992) Multiple equilibria states in combined thermal and saline circulation. J Phys Oceanogr 22:231–246
Hughes TMC, Weaver AJ (1994) Multiple equilibria of an asymmetric two-basin ocean model. J Phys Oceanogr 24:619–637
Jenkins GW, Watts DG (1968) Spectral analysis and its applications. Holden-Day, San Francisco, USA, pp 525
Knutson T, Gu D, Manabe S (1997) Simulated ENSO in a global coupled ocean-atmosphere model: multi-decadal amplitude modulation and CO2 sensitivity. J Clim 10:138–161
Kushnir Y (1994) Multidecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. J Clim 7:141–157
Kushnir Y, Held I (1996) Equilibrium atmospheric response to North Atlantic SST anomalies. J Clim 9:1208–1220
Latif M, Barnett TP (1994) Causes of decadal climate variability over the North Pacific and North America. Science 266:634–637
Latif M, Barnett TP (1996) Decadal variability over the North Pacific and North America: dynamics and predictability. J Clim 9:2407–2423
Latif M, Groetzner A, Barnett TP (1996) A mechanism for decadal climate variability. ACCP Notes: the newsletter of the Atlantic Climate Change Programm, III, April 1996, 1–3
Levitus S (1990) Multipentadal variability of steric sea level and geopotential thickness of the North Atlantic Ocean, 1970–1974 versus 1955–1959. J Geophys Res 95:5233–5238
Lorenz E (1969) Three approaches to atmospheric predictability. Bull Am Meteorol Soc 50:345–349
Lorenz E (1973) On the existence of extended range predictability. J Appl Meteorol 12:543–546
Manabe S, Stouffer RJ (1988) Two stable equilibria of a coupled ocean-atmosphere model. J Clim 1:841–866
Manabe S, Stouffer RJ (1995) Simulation of abrupt climate change induced by freshwater input to the North Atlantic Ocean. Nature 378:165–167
Manabe S, Stouffer RJ (1996) Low frequency variability of surface air temperature in a 1000 year integration of a coupled ocean-atmosphere-land surface model. J Clim 9:376–393
Manabe S, Stouffer RJ, Spelman MJ, Bryan K (1991) Transient response of a coupled ocean-atmosphere model to gradual changes of atmospheric CO_2: Part 1: annual mean response. J Clim 4:785–818
Manabe S, Spelman MJ, Stouffer RJ (1992) Transient response of a coupled ocean-atmosphere model to gradual changes of atmospheric CO_2: Part II: seasonal response. J Clim 5:105–126
Munk WH (1960) Smoothing and persistence. J Meteorol 17:92–93
Neelin DJ, Latif M, Jin FF (1994) Dynamics of coupled ocean-atmosphere models: the tropical problem. Ann Rev Fluid Mech 26:617–659
Palmer T (1996) Predictability of the atmosphere and oceans: from days to decades. In: Anderson D, Willebrand J (eds) NATO Advanced Study Institute. Decadal climate variability: dynamics and predictability, Springer-Verlag Berlin, Heidelberg, New York, pp 83–156
Penland C (1989) Random forcing and forecasting using principal oscillation pattern analysis. Mon Weather Rev 117:2165–2185
Rahmstorf S (1995) Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. Nature 378:145–149
Reif F (1965) Fundamentals of statistical and thermal physics. McGraw-Hill, New York, USA, pp 651
Sarachik ES, Winton M, Yin FL (1996) Mechanisms for decadal-to-centennial climate variability. In: Anderson D, Willebrand J (eds) NATO Advanced Study Institute. Decadal climate variability: dynamics and predictability, Springer-Verlag Berlin, Heidelberg, New York, pp 157–210
Shukla J (1981) Dynamical predictability of monthly means. J Atmos Sci 38:2547–2572
Stommel H (1961) Thermohaline convection with two stable regimes of flow. Tellus 13:224–230
Stouffer RJ, Manabe S, Bryan K (1989) Multihemispheric asymmetry in climate response to a gradual increase of atmospheric CO_2. Nature 342:660–662
Stouffer RJ, Manabe S, Vinnikov KY (1994) Model assessment of the role of natural variability in recent global warming. Nature 367:634–636
Timmermann A, Latif M, Voss R, Grötzner A (1996) North Atlantic interdecadal variability: a coupled air-sea mode. Max-Planck-Institut für Meteorologie, Report 223
Tziperman E, Toggweiler JR, Feliks Y, Bryan K (1994) Instability of the thermohaline circulation with respect to mixed boundary conditions: is it really a problem for realistic models? J Phys Oceanogr 24:217–232
Weaver AJ, Hughes TMC (1992) Stability and variability of the thermohaline circulation and its link to climate. Trends Phys Oceanogr 1:15–70
Wunsch C (1992) Decade-to-century changes in the ocean circulation. Oceanography 5:99–106