

Forecast Skill and Predictability of Observed Atlantic Sea Surface Temperatures

LAURE ZANNA

Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, United Kingdom

(Manuscript received 22 September 2011, in final form 10 January 2012)

ABSTRACT

An empirical statistical model is constructed to assess the forecast skill and the linear predictability of Atlantic Ocean sea surface temperature (SST) variability. Linear inverse modeling (LIM) is used to build a dynamically based statistical model using observed Atlantic SST anomalies between latitudes 20°S and 66°N from 1870 to 2009. LIM allows one to fit a multivariate red-noise model to the observed annually averaged SST anomalies and to test it. Forecast skill is assessed and is shown to be $O(3\text{--}5 \text{ yr})$. After a few years, the skill is greatly reduced, especially in the subpolar region. In the stable dynamical system determined by LIM, skill of annual average SST anomalies arises from four damped eigenmodes. The four eigenmodes are shown to be relevant in particular for the optimal growth events of SST variance, with a pattern reminiscent of the low-frequency mode of variability, and in general for the predictability and variability of Atlantic SSTs on interannual time scales. LIM might serve as a useful benchmark for interannual and decadal forecasts of SST anomalies that are based on numerical models.

1. Introduction

In the Atlantic Ocean, coherent large-scale sea surface temperature (SST) anomalies are present in the observed record. SST fluctuations arise from different mechanisms, such as a local response to stochastic atmospheric heat flux or advection of heat by the ocean circulation (e.g., Battisti et al. 1995; Bjerknes 1964; Halliwell 1998) and can potentially have a large impact on climate on interannual to decadal time scales. The possible dynamical mechanisms involve damped oscillatory modes of the ocean excited by stochastic atmospheric forcing, coupled ocean–atmosphere oscillatory modes randomized by stochastic forcing, or self-sustained ocean and coupled modes and delayed oscillators (e.g., Weaver et al. 1993; Griffies and Tziperman 1995; Chen and Ghil 1995; Czaja and Marshall 2001; Marshall et al. 2001). Despite the large number of studies, it is still difficult to exclude any of the proposed mechanisms, in part because of the relatively short historical record. In numerical model simulations, the time scales and mechanisms explaining the variability in the Atlantic sector differ, and therefore our ability to rely on

these models is somewhat limited. One goal of this study is to understand how the observed record can better guide our numerical models.

The presence of large fluctuations of Atlantic SSTs obviously has many economical and societal consequences because it is associated, for example, with rainfall in the Sahel (Palmer 1986; Biasutti et al. 2008) and hurricane activity (Zhao et al. 2009); therefore providing reliable predictions of Atlantic SST anomalies (SSTa) could be extremely beneficial. The potential predictability of Atlantic SST and ocean circulation variability has been extensively explored in many numerical models (e.g., Griffies and Bryan 1997) and has been found to be on the order of up to a couple of decades; however, potential predictability does not necessarily translate into forecast skill. So far, only one numerical study using a perfect model approach shows that potential predictability of Atlantic SST variations is not a prerequisite for actual forecast skill (Hawkins et al. 2011). While the current general circulation model (GCM)-based tools being developed for regional climate predictions are complex, challenging, and imperfect, the observed record has been mostly ignored when it comes to determining forecast skill and predictability on interannual time scales. The main goal of this study is to extract the statistical properties of the observed record to investigate the dynamics of Atlantic SSTa and their forecast skill.

Corresponding author address: Laure Zanna, Atmospheric, Oceanic and Planetary Physics, Dept. of Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, United Kingdom.
E-mail: zanna@atm.ox.ac.uk

The paper is organized as follows. In section 2, a statistical model of annual mean Atlantic SST anomalies is developed using linear inverse modeling, and its dynamical properties are investigated. In section 3, the forecast skill of the statistical model is evaluated as are the growth events of SST variance, and their relationship to the Atlantic multidecadal variability is examined. A short discussion is provided in section 4.

2. Linear inverse modeling

a. Data

Numerical climate models are extremely useful but often do not reproduce the observed variability properly and introduce structural errors that are difficult to quantify. To investigate the variability and forecast skill of Atlantic SSTa, the Met Office Hadley Centre's SST raw dataset (HadSST2) on a 5° by 5° latitude–longitude grid from 1870 to 2009 (Rayner et al. 2006) is used.

The SST anomalies are constructed by removing the climatology over the entire period considered and also the trend. Separating the natural and anthropogenically forced variability is the subject of active research. Methods attempting to separate the oscillatory-like component and the upward trend are still largely debated (e.g., Ting et al. 2009; DelSole et al. 2011); we therefore opt for a simple solution. The data at each grid cell were detrended using two methods: 1) removing the linear trend as is commonly done in various studies (Enfield et al. 2001; Sutton and Hodson 2005) or 2) using a cubic spline with two breakpoints similar to that of Hawkins et al. (2011). The results presented in this study use the latter (cubic spline); however, the main conclusions remain unchanged if the linear trend is removed. Yet, perhaps a part of the externally forced component of SST variability is still present and could potentially affect the predictions.

The domain is confined to the Atlantic basin between latitudes 20°S and 66°N for several reasons: the frequency of SST measurements during 1870–2009 is relatively high when compared with all other regions, the impact of sea ice variability on SSTa is minimized, and the possible modulations of midlatitude SSTa from the tropics due to the large-scale ocean circulation are taken into account. The time series at each grid point are averaged from January to December of each year to obtain annual averages; only years having 10 or more months are retained. Grid cells with less than 75% of the years are omitted. The high latitudes, despite being important in Atlantic variability, are left out because of their poor sampling over the period considered and, as mentioned, the

spurious effects of sea ice. It would have been possible to use interpolated datasets such as the Hadley Centre Sea Ice and SST dataset (HadISST; Rayner et al. 2003) or reduced-space optimal analysis Met Office Historical SST dataset (RSA MOHSST5; Kaplan et al. 1998) instead of the raw data to include the high latitudes (e.g., the Labrador and Greenland Seas); however, these datasets are constructed using empirical orthogonal function (EOF) optimal interpolation methods that could potentially compromise the analysis.

b. Estimation of the linear inverse model

In this study, linear inverse modeling (LIM) is used as a tool to construct a model of the Atlantic Ocean that is based on the statistical properties of observed annual mean SST anomalies. LIM has been a useful tool in seasonal predictions (e.g., Penland and Sardeshmukh 1995; Newman 2007), especially in the tropics, predictions of which are found to be skillful at lead times of several months. While statistical models have their limitations, they can often provide a useful assessment of the variability and forecast skill of the observed record, allowing for a comparison with numerical models, and serve as a benchmark for numerical climate studies.

The main assumption underlying LIM is that the time-dependent annual average SSTa can be separated into a linear deterministic part and a nonlinear part represented by a linearized term that is dependent on the base state and white noise. To reduce the number of degrees of freedom, we construct a reduced space for SSTa using EOFs as the spatial fields and their respective principal components (PCs) as the time evolution of the EOFs. The data are weighted by the surface area of the grid box and their standard deviation (Moore and Kleeman 2001). Figure 1 shows the spatial patterns of the first three EOFs that explain 32%, 10%, and 7.7% of the variability and are reminiscent of many studies (e.g., Deser and Blackmon 1993).

The evolution of the PCs can therefore be written as a linear dynamical system forced with white stochastic forcing η such that

$$d\mathbf{P}/dt = \mathbf{A}\mathbf{P} + \eta. \quad (1)$$

Here $\mathbf{P}(t)$ is the PC state vector and the deterministic matrix \mathbf{A} defines the evolution of \mathbf{P} . The matrix \mathbf{A} is estimated from multiple linear regressions over the lead time $\tau_0 = 1$ yr. The matrix \mathbf{A} is given by

$$\mathbf{A} = \frac{\ln[\mathbf{C}(\tau_0)\mathbf{C}(0)^{-1}]}{\tau_0}, \quad (2)$$

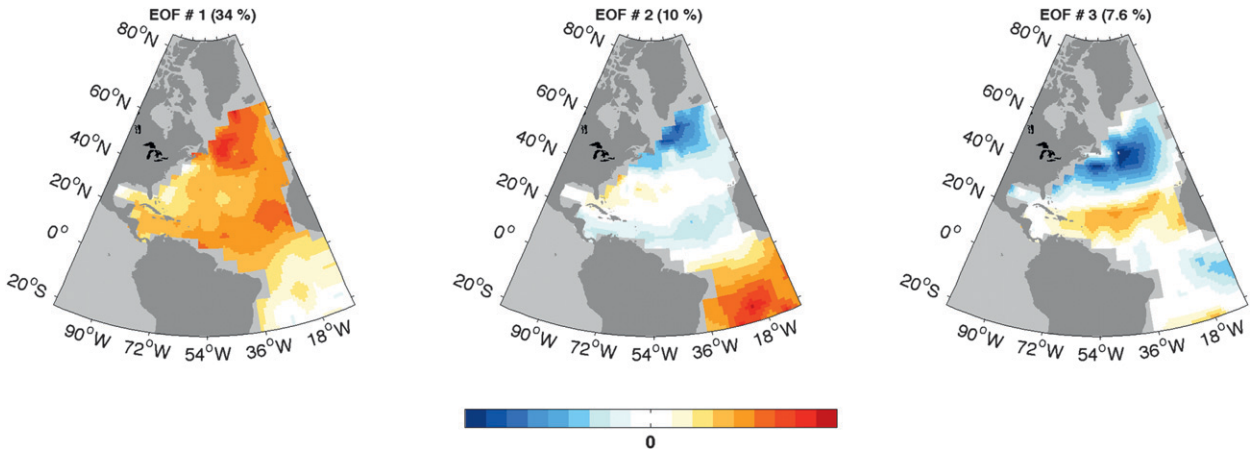


FIG. 1. First three EOFs of annually averaged SST anomalies in the Atlantic basin using the raw dataset HadSST2. The contour intervals are the same for each plot and are arbitrary. The fraction of variance explained by each EOF is indicated in parentheses.

using the covariance matrices $\mathbf{C}(\tau_0) = \langle \mathbf{P}(t + \tau_0) \mathbf{P}^T(t) \rangle$ and $\mathbf{C}(0) = \langle \mathbf{P}(t) \mathbf{P}^T(t) \rangle$, where the angle brackets denote an ensemble average or an average over all t if the statistics of η are stationary. Only 13 EOFs/PCs, explaining about 79% of the variability, are retained to evaluate \mathbf{A} , ensuring that \mathbf{A} remains well defined with negative eigenvalues and a relatively low condition number.

c. Eigenanalysis of \mathbf{A}

To verify that the LIM adequately describes the Atlantic SSTa variability, several tests can be performed. To explore the properties of the LIM and its spectrum, the eigenvalues $\gamma_j = \alpha_j + i\beta_j$ and corresponding eigenvectors \mathbf{v}_j (or empirical normal modes) of \mathbf{A} are computed (Penland and Sardeshmukh 1995, hereinafter PS95). The eigenmodes are often referred to as principal oscillation patterns (Hasselmann 1988). The \mathbf{A} must be dissipative with negative eigenvalues for the LIM to have stationary statistics and to be a valid model for the SSTa dynamics, which is confirmed by the tests performed on the eigenanalysis of \mathbf{A} . Table 1 shows the decay and oscillatory time scales of the different eigenmodes. Some of the decaying eigenmodes oscillate with periods ranging from 8.5 to 137 yr. Note that perhaps some of the dynamical properties of the system are poorly represented because of the short time series or because of sampling issues. For example, similar results for the eigenanalysis of the matrix \mathbf{A} were obtained when evaluated for lead times up to $\tau_0 = 4$ yr; however, this is not the case for larger lead times. In this event, we run into the Nyquist problem for which the lead time τ_0 is close to one-half the period of one the eigenmodes (eigenmode $2/3$ in Table 1), and the eigenmodes cannot be reliably estimated. Different EOF truncations lead to

qualitatively similar results, with some quantitative changes in the eigenvalues.

d. Multivariate model of SSTa

To further validate that Eq. (1) is an adequate model of the observed Atlantic SSTa, the LIM should in part reproduce the SSTa power spectrum. The noise forcing η and its statistics can be evaluated from the fluctuation–dissipation relationship (PS95), that is, $\mathbf{A}\mathbf{C}(0) + \mathbf{C}(0)\mathbf{A}^T + \langle \eta\eta^T \rangle dt = 0$. The covariance matrix of the forcing, $\langle \eta\eta^T \rangle dt = \mathbf{Q}$, has to be positive definite. All of the eigenvalues of \mathbf{Q} are found to be positive except for one negative with a small amplitude. The higher-order EOFs are possibly not reliably estimated from the (noisy) data, leading to the small negative eigenvalue of \mathbf{Q} since using fewer EOFs leads to a purely positive definite matrix \mathbf{Q} . Equation (1) is integrated forward for 49 000 yr (Penland and Matrosova 1994) with a time step of 1 h, with η generated using the {eigenvectors, eigenvalues} of \mathbf{Q} , $\{\mathbf{q}_j, r_j^2\}$, such that $\eta = \sum_j w_j(t) r_j \mathbf{q}_j$, where $w_j(t)$ are independent Gaussian white noises with unit variance.

TABLE 1. Eigenvalues of the linearized matrix \mathbf{A} constructed using 13 EOFs.

Mode	$\tau_r = \alpha^{-1}$ (yr)	$\tau_i = 2\pi/\beta$ (yr)
Mode 1	1.2	—
Modes 2/3	1.4	± 8.5
Mode 4	3.1	—
Modes 5/6	4.0	± 13.4
Modes 7/8	4.9	± 21.1
Modes 9/10	6.7	± 36.8
Modes 11/12	8.0	± 137.1
Mode 13	15.3	—

The LIM reproduces the main features of the observed power spectrum of the domain-averaged Atlantic SSTa as shown in Fig. 2a. The LIM spectrum is as expected much smoother than the observed spectrum, because of the relatively few degrees of freedom in the truncated EOF space. Some of the irregularities could also be due to the poor sampling and errors from the observations.

3. Forecasts and predictability

a. Linear forecasts

Given that the LIM passes several of the necessary tests, it can be now used to provide forecasts using the estimated propagator

$$\mathbf{B}(\tau) = \exp(\mathbf{A}\tau) = \exp\left[\frac{\tau}{\tau_0} \ln \frac{\mathbf{C}(\tau_0)}{\mathbf{C}(0)}\right].$$

Forecasts of \mathbf{P} (denoted by $\hat{\mathbf{P}}$) are therefore given by $\hat{\mathbf{P}}(t + \tau) = \mathbf{B}(\tau)\mathbf{P}(t)$. The spatial pattern of the forecast can be recovered by multiplying $\hat{\mathbf{P}}$ by the respective EOF patterns. Obviously, the LIM forecasts have to be tested on data that are independent of those used to determine \mathbf{A} , which is done as follows: the data are subsampled by removing one decade, the reduced space and \mathbf{A} are recomputed for the remaining years, the forecasts are evaluated for the independent years, and this procedure was repeated every year.

Figure 2b shows the reconstructed domain-averaged Atlantic SSTa time series in the EOF space (black curve) and associated forecasts (gray curves) using the LIM for lead times of up to 10 yr, plotted at 1-yr intervals. The LIM seems to show some skill for a few years, yet it seems to be state dependent as some decades are better predicted than others. We perform a quantitative assessment of the skill of these predictions by comparing the linear model forecast errors with an autoregressive (AR1) process (equivalent to damped persistence), shown in Fig. 2c. Moreover, we compare the forecast error of the LIM with the theoretical error covariance due to the presence of the unpredictable white-noise forcing (referred to as a “perfect” linear model), given by (PS95) as

$$\langle \mathbf{e}(\tau)\mathbf{e}(\tau)^T \rangle = \mathbf{C}(0) - \mathbf{B}(\tau)\mathbf{C}(0)\mathbf{B}^T(\tau). \quad (3)$$

The estimates of the forecast error covariance provide an estimate of the statistics of the noise forcing η responsible for the forecast error \mathbf{e} . The LIM outperforms the AR1 model for all lead times; however, the actual errors using the LIM are larger than those predicted by the theoretical perfect model. The actual errors of the LIM are up to 30% larger than their theoretical estimates

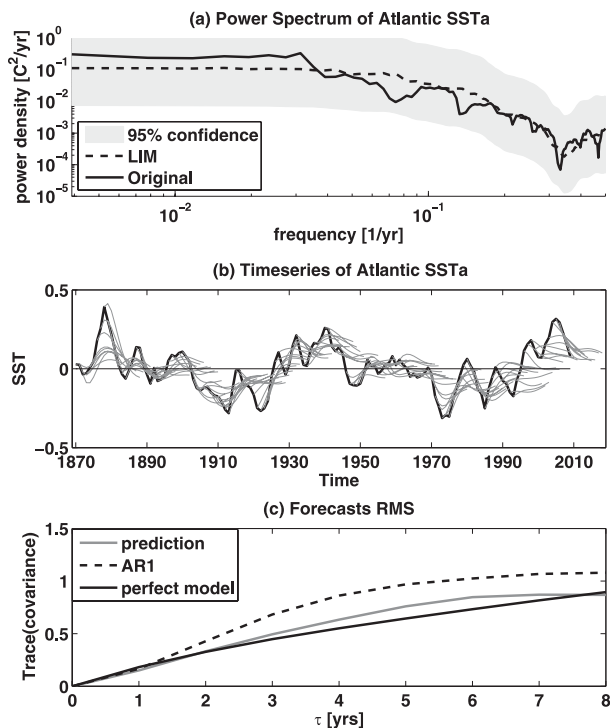


FIG. 2. (a) Power spectrum of Atlantic annual mean SSTa averaged over the domain 20°S–66°N: observations (solid line) and LIM (dashed line). Gray shading denotes the 95% confidence interval evaluated from a 49 000-yr run of the LIM (the model output is separated into 350 segments of 140-yr time series, and the ensemble mean is used to estimate the power spectrum of the model and confidence interval). (b) Time series of Atlantic annual mean SSTa averaged over the domain 20°S–66°N for the original dataset (black line) and for 10-yr segment predictions using the LIM (gray lines). (c) Normalized domain-integrated SST error forecast for the LIM (gray line), a perfect linear model (solid black line), and an AR1 model (dashed black line). The errors are normalized by the trace of the covariance matrix $\mathbf{C}(0)$.

at a lead time of 5 yr. This does not necessarily invalidate the linearity assumption. While the differences between the actual and theoretical errors can be seen as a measure of the nonlinearities of the system, they could also be due to the severe EOF truncation or the nonstationarity of the data.

Maps of mean-square errors relative to climatology [using the climatology from HadSST2 and simply assuming that $\hat{\mathbf{P}}(t + \tau) = 0$], shown in Fig. 3, indicate the regions for which the forecast skill is greatly reduced. For years 1–2, the forecast skill in the Atlantic basin is largely superior to climatology. For years 3–5, the forecast skill decreases south of Greenland, in the subpolar region. For longer lead times, the model fails in the same subpolar region and skill is lost in the tropics, especially near the African coast. In the subpolar region, the variance of SSTa is large and ocean–atmosphere coupling in addition to ocean mixed-layer processes is strong, leading

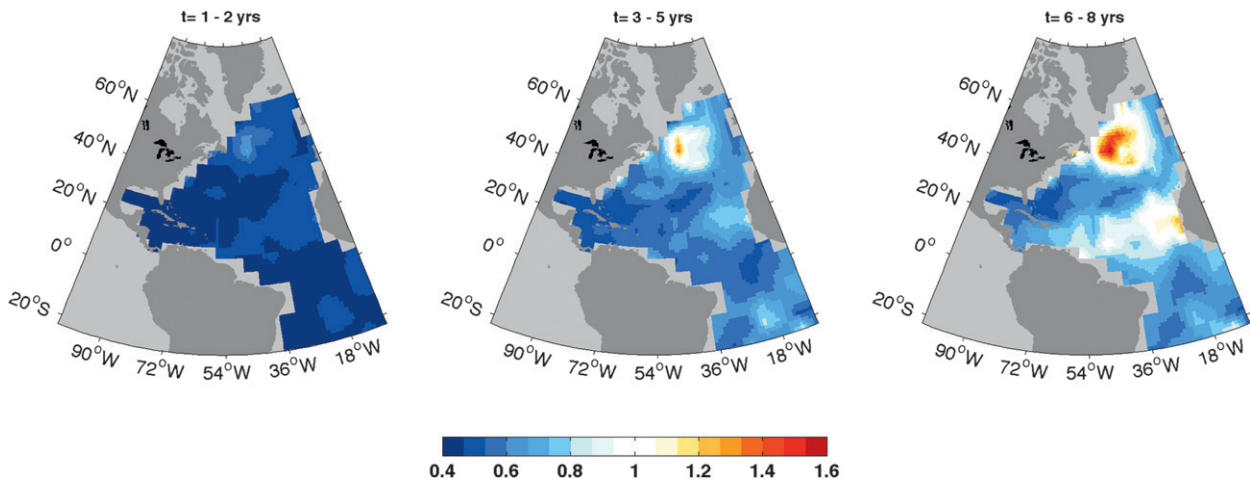


FIG. 3. Mean-square error relative to climatology [the climatology is simply determined from the HadSST2 dataset with $\hat{\mathbf{P}}(\tau) = 0$] for (a) $t = 1-2$ yr, (b) $t = 3-5$ yr, and (c) $t = 6-8$ yr.

to the possible reduction of the forecast skill of the LIM. If the anomalies were constructed using averages running from July to June (instead of from January to December), the main conclusions would still hold but the forecast skill would be reduced by roughly 10%. Including the full winter seems to have a negative effect on the forecast skill, again potentially being due to the large air–sea interaction or sporadic ocean variability during this season.

b. Transient growth, optimal initial conditions, and predictability

Similar to most geophysical systems, \mathbf{A} is nonnormal ($\mathbf{A}\mathbf{A}^T \neq \mathbf{A}^T\mathbf{A}$) with nonorthogonal eigenvectors (Farrell 1982). Through interference of the decaying nonorthogonal eigenmodes \mathbf{v}_j of \mathbf{A} , large transient amplification of variance is possible. Such transient amplification can be examined for two purposes: 1) to understand error variance growth and to correctly sample and interpret errors in initial conditions [such an interpretation of transient growth is common in the atmospheric, numerical weather prediction, and seasonal forecast communities (e.g., Buizza and Palmer 1995; Moore and Kleeman 1996)] and 2) to explain the actual growth of the variance in the observed record or in numerical models (e.g., Farrell 1989; PS95).

In several GCM experiments, significant transient amplification of annual mean Atlantic SSTa is found on time scales of 4–20 yr (Zanna et al. 2012; Tziperman et al. 2008; Hawkins and Sutton 2009). For the current LIM that is based on observations, the largest possible transient growth of the domain-integrated Atlantic SST variance without forcing is given by $\lambda(\tau) = \mathbf{P}(\tau)^T \mathbf{P}(\tau) / \mathbf{P}(0)^T \mathbf{P}(0)$ and is shown in Fig. 4a as a function of the

lead time τ . The $\lambda(\tau)$ is the leading singular value of $\mathbf{B}(\tau)$ such that the maximum amplification of SST variance is found at about $\tau = 3$ (to 4) yr with an amplitude of roughly 8. Further exploration of the singular values of $\mathbf{B}(\tau = 3$ yr) reveals that three additional singular values are larger than unity, therefore exhibiting possible linear growth; those three additional singular values are ranging between 1.5 and 4, therefore leading to significantly less amplification than the leading singular value. The maximum amplification curve in Fig. 4a implies that the actual growth of variance from the interference of nonorthogonal eigenmodes can be sustained for more than 10 yr, or actually up to 18 yr (the time at which λ decreases to unity). This time scale can be viewed as an optimistic estimate or an upper bound on the predictability time of the linear events without forcing. However, given that the system is not a perfect linear system and that forcing (noise) is present, predictability is lost as a result of the white-noise forcing. In fact, for the domain-averaged SSTa to be predictable, $\lambda(\tau)$ has to be larger than the error covariance $1 + \langle \mathbf{e}(\tau)\mathbf{e}(\tau)^T \rangle$, and the limit for predictability using this criterion is slightly under 10 yr. The other way to think about the growth, as mentioned earlier, is in terms of error variance growth such that the rapid increase in SST variance over 3–4 yr could lead to a limited predictability as a result of errors in initial conditions and could potentially explain the forecast skill of 3–5 yr, especially in the subpolar regions.

While the system seems to be able to sustain SST variance growth, it is interesting to explore whether these potentially predictable linear growth events exist in the record and to determine their spatial structure. The initial conditions leading to the maximum growth at a given time τ are given by the leading singular vector of $\mathbf{B}(\tau)$,

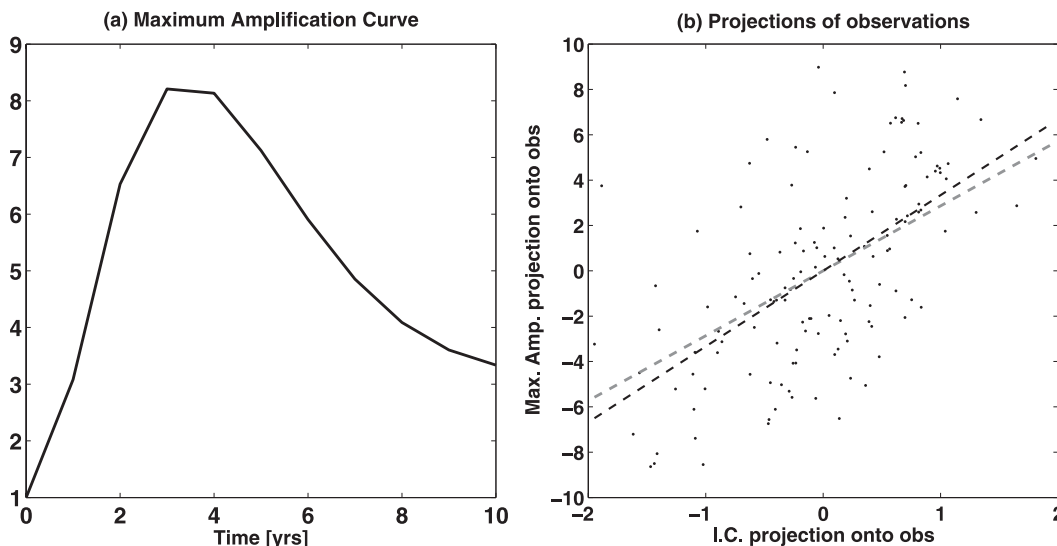


FIG. 4. (a) Maximum amplification curve, showing maximum amplification λ as function of lead time τ . (b) Projection of observations upon the optimal initial condition for amplification of SST anomalies in the LIM over a 3-yr interval vs the optimal evolved SST state 3 yr later. The gray dashed line is the expected growth $\lambda^{0.5}(\tau = 3 \text{ yr}) = 2.45$; the black dashed line is the linear fit with linear correlation $r = 0.6$.

often termed optimal initial conditions (Farrell and Ioannou 1996). Optimal perturbations can be relevant to the variability of the observed system; for example, PS95 showed that the leading singular vector amplifying the tropical Pacific SST corresponds to the most efficient initial condition for the development of ENSO.

The spatial pattern of the optimal initial perturbations corresponding to the largest growth [obtained for $\mathbf{B}(\tau = 3 \text{ yr})$] is shown in Fig. 5a. The initial state of the optimal perturbations (Fig. 5a) exhibits several signals in the Northern Hemisphere (in the vicinity of the Gulf Stream and its extension, the southern tip of Greenland, the Gulf of Mexico, and the Canary Islands) and an additional anomaly in the Southern Hemisphere around 20°S . The shape of the leading singular vector at different τ does not change very much, with always a relatively pronounced anomaly in the middle of the basin, in the vicinity of the Gulf Stream's extension. Therefore the LIM has only one main leading optimal initial structure. The evolution of the perturbations (Fig. 5) over the optimal growth time scale hints at a nonlocal growth of the perturbations (at least over the first year) with possible mechanisms involving ocean-atmosphere interaction, mixed-layer processes, re-emergence of SST, and advection by the large-scale ocean circulation. The optimally evolved pattern (Fig. 5d), reminiscent of the first EOF and the low-frequency mode of variability in the Atlantic (Schlesinger and Ramankutty 1994), has a single sign over the Northern Hemisphere with the strongest anomalies in the subpolar

gyre region and just north of the tropics. In the Southern Hemisphere, a small anomaly of opposite sign has developed and is centered around 20°S , 27°W . The optimal perturbation found can be seen as a precursor event to the large warming (or cooling) over the basin as shown in Fig. 5d. The initial anomalies along 30°N (Fig. 5a) and the evolved pattern after one year (Fig. 5b) resemble the optimal initial and final patterns after 9 months found in Vimont (2012) in which LIM was constructed with monthly SST data.

When analyzing optimal structures, it is important to consider whether the growth is actually observed to occur as predicted by the LIM. Figure 4b shows the projection of observed anomalies onto the optimal initial conditions versus their projection onto the predicted evolved structure 3 years later. The remarkable agreement between the slopes of least squares fit and the linear amplification of SST [$\lambda^{1/2}(\tau = 3 \text{ yr})$] and the correlation of 0.6 are indications that optimal SST growth is relatively well captured by the LIM and is relevant for observed Atlantic variability. As expected, some scatter is observed about the least squares fit due to the noise η reducing the correlation.

c. Dynamical eigenmodes of Atlantic variability

A vast number of studies have been looking at the temporal and spatial variations of the Atlantic multidecadal variability or Atlantic multidecadal oscillation (AMO) and their impact on climatic phenomena such as rainfall or hurricane activity (Enfield et al. 2001; Sutton

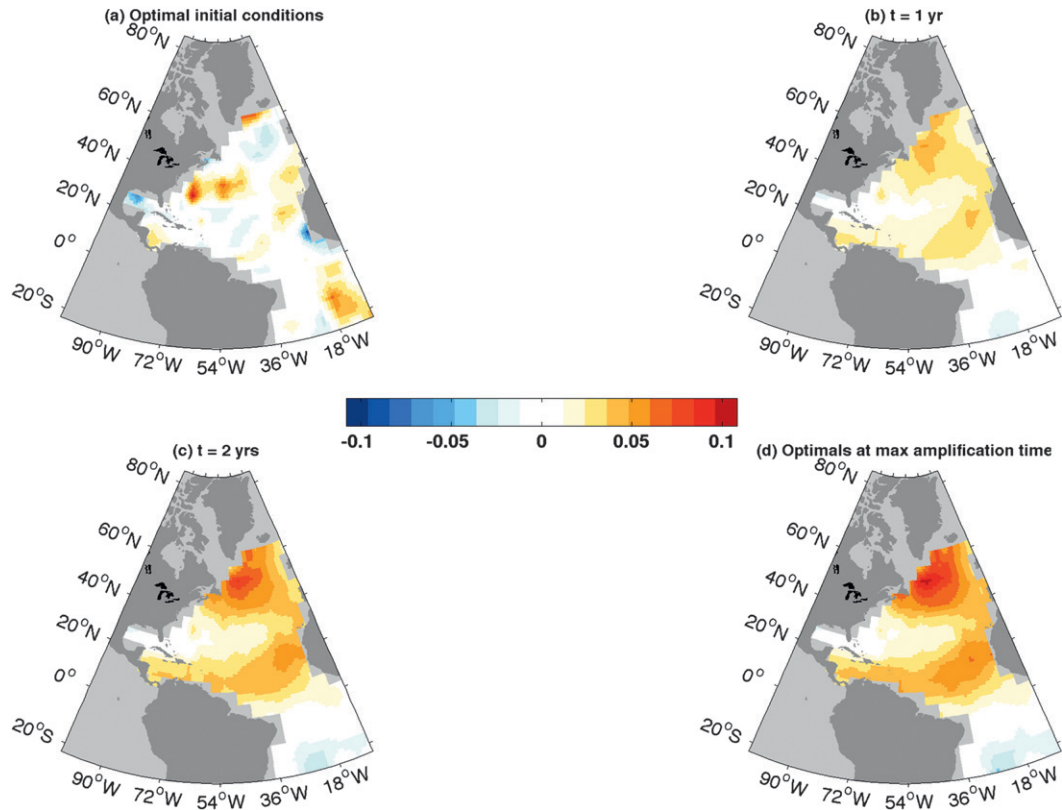


FIG. 5. Time evolution of the optimal perturbations: (a) optimal initial conditions, with $t = 0$, (b) optimals at $t = 1$ yr, (c) optimals at $t = 2$ yrs, and (d) optimals at the maximum amplification time, $t = 3$ yr. Contour intervals are the same for all panels, and the amplitude is arbitrary (due to the linearity of the optimal growth problem).

and Hodson 2005). The Atlantic multidecadal variability is characterized by basinwide fluctuations in North Atlantic SST anomalies; this SSTa pattern of variability exhibits some similarities with the optimally excited pattern shown in Fig. 5d and is found in several numerical models.

In this section, we consider how the unsmoothed AMO index time series and the optimal patterns project onto the eigenmodes of the linearized matrix \mathbf{A} to gain some insight about the source of predictability of Atlantic SSTa. The AMO index time series are calculated from the monthly Kaplan SST dataset. The index is an area-weighted average of linearly detrended SST anomalies over the North Atlantic between 0° and 70°N . The index was obtained online (<http://www.cdc.noaa.gov/data/timeseries/AMO/>). The AMO index is simply a measure of the domain-averaged annually averaged SST variability (and is not defined as the leading EOF/PC explaining most of the variance, unlike the Pacific decadal oscillation in the North Pacific Ocean). The 5-yr running mean of the temporal evolution of the AMO index is often used and seems to suggest shifts from one sign to the other in Atlantic SSTa on multidecadal time

scales [70–80 yr as in Kushnir (1994)]. Because of the presence of the external anthropogenic forcing and our difficulty in adequately removing the forced component of SST fluctuations to isolate the internal variability, different studies define different periods as a shift in the sign of the AMO index. We will not worry about such issues in this work and only examine year-to-year variations of the unsmoothed time series.

The regression coefficients \mathbf{R}_{AMO} between the PCs and the AMO index are found in a way that is similar to the procedure described in Tziperman et al. (2008). Therefore, the AMO index reconstructed from the PCs is given by $\widehat{\text{AMO}} = \mathbf{R}_{\text{AMO}}^T \mathbf{P}$. The correlation between the original AMO index time series and AMO is roughly 0.91. The time-dependent PCs, $\mathbf{P}(t)$, can be written as the sum of the contributions of the different eigenmodes of \mathbf{A} such that

$$\mathbf{P}(t) = \sum_{j=1}^{j=13} \phi_j(t) \mathbf{v}_j. \tag{4}$$

The projection time series for the j th eigenmode, $\phi_j(t)$, are given by

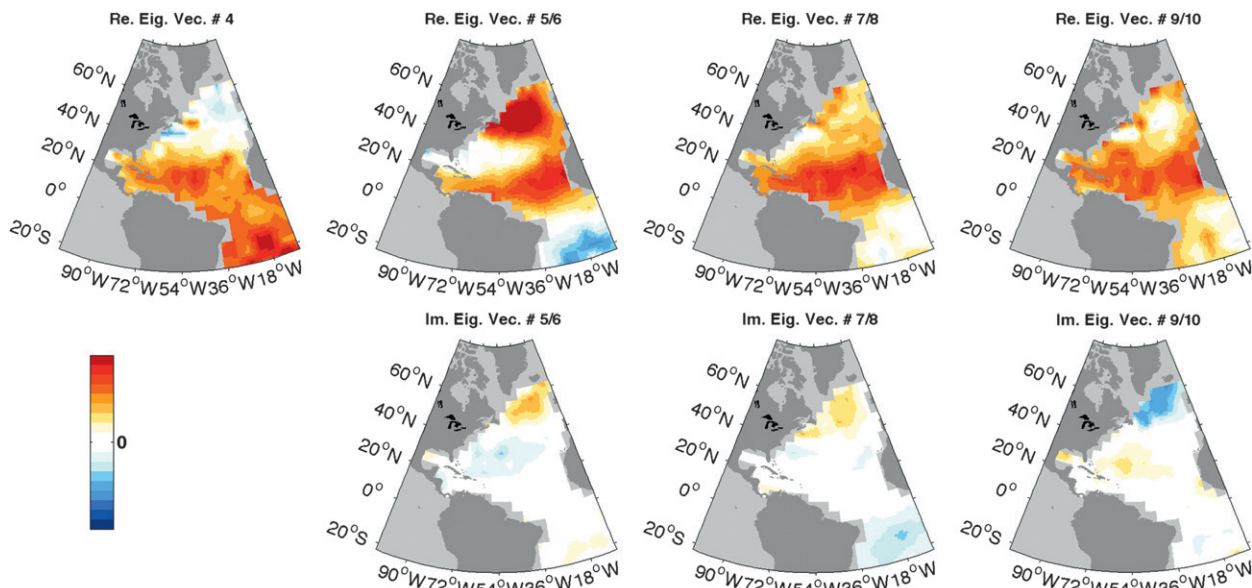


FIG. 6. Spatial patterns [(top) real and (bottom) imaginary parts] associated with the four eigenmodes (4, $5/6$, $7/8$, and $9/10$) necessary to explain the optimal growth and the AMO index time series. The contour intervals are the same for all panels.

$$\phi_j(t) = \frac{\mathbf{g}_j^* \mathbf{P}(t)}{\mathbf{g}_j^* \mathbf{v}_j}, \quad (5)$$

with \mathbf{g}_j being the eigenvectors of the adjoint matrix of \mathbf{A} satisfying the biorthogonality relation (Farrell and Ioannou 1996). Given that the eigenmodes are not orthogonal, the variance cannot simply be partitioned into individual eigenmodes. Multiple eigenmodes are necessary to explain the variance of the AMO time series as well as the evolution of the optimal growth of anomalies found in section 3.

The leading optimal structure and growth can be well approximated by four nonnormal eigenmodes of \mathbf{A} : the combination of three complex pairs of oscillatory modes (modes $5/6$, $7/8$, and $9/10$ in Table 1) and one stationary eigenmode (mode 4 in Table 1), which are acting on different temporal and spatial scales. Their real and imaginary spatial patterns are shown in Fig. 6. Using these four eigenmodes with decay and oscillatory time scales ranging between 3.1 and 6.7 yr and 13.4 and 36.8 yr, respectively, we can then reconstruct and explain the linear growth events. The least-damped eigenmode does not participate in the optimal growth.

In looking at Fig. 6, one could assume that the oscillation of eigenmode $5/6$ would lead the optimal growth of the anomalies and possibly define the evolution of the AMO index. However, the variance explained by eigenmode $5/6$ is only one-half of the observed AMO variance, and the time series of the eigenmode $5/6$ correlates only at 0.43 with the AMO index. Similar results are obtained for the evolution of the optimal pattern.

The projection of the AMO index time series onto the eigenmodes reveals that the same four eigenmodes are necessary to reproduce the evolution of the AMO index. This is perhaps not surprising since the AMO index is closely related to the SST variance norm used to maximize the optimal growth. The reconstructed AMO with the four eigenmodes and AMO time series are shown in Fig. 7 and are found to have a correlation of approximately 0.89. It therefore appears that the skill and predictability of Atlantic SSTa arise from the decaying time scales of the participating eigenmodes and not from their oscillatory time scales, similar to the PDO variability (Newman 2007).

4. Discussion and conclusions

Using the linear inverse modeling approach, we constructed a statistical model of annually averaged Atlantic SST anomalies. The main results can be summarized as follows:

- SST variability in the Atlantic domain can be well captured by linearized dynamics forced with white noise.
- The LIM captures the main features of Atlantic SST anomalies, including their power spectrum and optimal growth events.
- Linear forecasts of SSTa are successful for up to 5 yr after which the skill deteriorates; using a 16-yr-long time series of surface temperature in the vicinity of the Grand Banks, Wunsch (2012) shows that some linear predictive skill is possible up to ~ 5 yr, similar to our results.
- The forecast skill is mainly lost in the subpolar region south of Greenland, possibly because of the nonlinearities

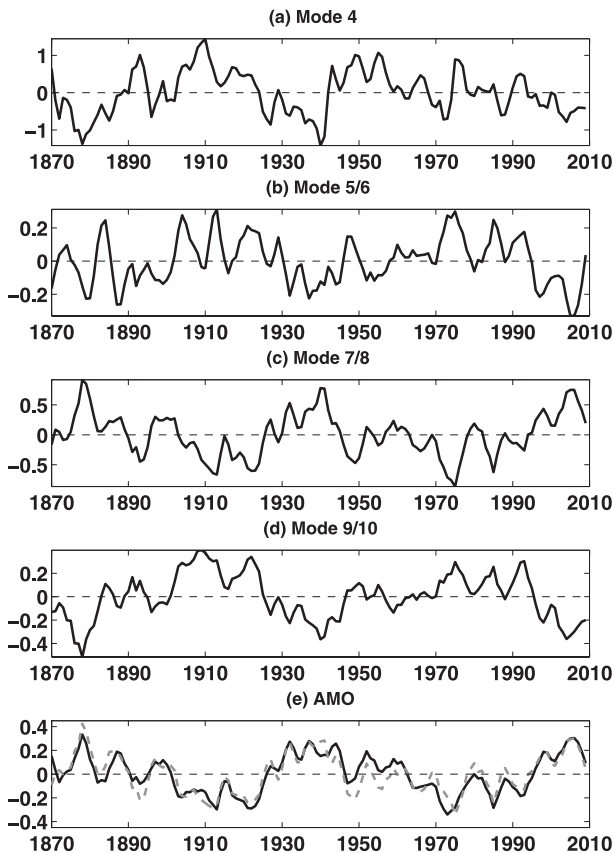


FIG. 7. Projection time series $\phi_i(t)$ of eigenmodes (a) 4, (b) 5/6, (c) 7/8, and (d) 9/10, and (e) time series of the AMO index (black line) and the reconstructed AMO index (dashed gray line) from the four eigenmodes in (a)–(d).

arising from air–sea interaction and/or subduction processes.

- Optimal patterns develop into large anomalies reminiscent of the first EOFs of the system and low-frequency SST patterns found in models (DelSole et al. 2011).
- Linear growth events and AMO variability can be explained by four eigenmodes of the linearized operator; the four nonorthogonal eigenmodes are responsible for the skill and the predictability in the Atlantic sector.

Yet, 140 yr of data might still be too short to adequately explore the interannual (and multidecadal) variability and predictability of the Atlantic, especially given the large uncertainties in observations during the first half of the record. The model and its skill could perhaps be improved by adding additional datasets such as sea level pressure, salinity, subsurface temperature, or temperatures from other ocean basins. However, careful analysis will be necessary because these datasets are short and extremely noisy (even more noisy than the one used in this study).

While statistical models are useful, they have several limitations. It is extremely difficult to investigate the physical mechanisms and to confirm that the results obtained are real and are not an artifact of the statistics or of the short time series. For example, it is difficult to compare the eigenmodes or singular vectors obtained with numerical studies because most of them concentrate on the EOFs and not on the dynamical modes. Moreover, the physical mechanisms (e.g., different oceanic and atmospheric processes) leading to the growth of the optimal structure and therefore participating in exciting Atlantic variability remain unexplained. We hope to address these issues in future work with a designed set of experiments using idealized and complex numerical models in addition to observations. We expect to identify the dynamical eigenmodes, explore their relationship to the skill and variability in the Atlantic, and compare them with the observations.

The results suggest that statistical models such as LIM could provide valuable information about regional predictability and variability before using complex nonlinear tools, including GCM-based tools. LIM as a global or regional tool could serve as a benchmark for interannual and decadal climate forecasts.

Acknowledgments. Many thanks are given to Ed Hawkins and David Munday for useful discussions and to Cecile Penland and two additional reviewers for their comments that helped improve the manuscript. The author was supported by a Fellowship from the Oxford Martin School and Balliol College, University of Oxford.

REFERENCES

- Battisti, D. S., U. S. Bhatt, and M. A. Alexander, 1995: A modeling study of the interannual variability in the wintertime North Atlantic Ocean. *J. Climate*, **8**, 3067–3083.
- Biasutti, M., I. M. Held, A. H. Sobel, and A. Giannini, 2008: SST forcings and Sahel rainfall variability in simulations of the twentieth and twenty-first centuries. *J. Climate*, **21**, 3471–3486.
- Bjerknes, J., 1964: Atlantic air–sea interaction. *Advances in Geophysics*, Vol. 10, Academic Press, 1–82.
- Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52**, 1434–1456.
- Chen, F., and M. Ghil, 1995: Interdecadal variability of the thermohaline circulation and high-latitude surface fluxes. *J. Phys. Oceanogr.*, **25**, 2547–2568.
- Czaja, A., and J. Marshall, 2001: Observations of atmosphere–ocean coupling in the North Atlantic. *Quart. J. Roy. Meteor. Soc.*, **127**, 1893–1916.
- DelSole, T., M. K. Tippett, and J. Shukla, 2011: A significant component of unforced multidecadal variability in the recent acceleration of global warming. *J. Climate*, **24**, 909–926.
- Deser, C., and M. L. Blackmon, 1993: Surface climate variations over the North Atlantic Ocean during winter: 1900–1989. *J. Climate*, **6**, 1743–1753.

- Enfield, D. B., A. M. Mestas-Nunez, and P. J. Trimble, 2001: The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophys. Res. Lett.*, **28**, 2077–2080.
- Farrell, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, **39**, 1663–1686.
- , 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.*, **46**, 1193–1206.
- , and P. J. Ioannou, 1996: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.*, **53**, 2025–2040.
- Griffies, S. M., and E. Tziperman, 1995: A linear thermohaline oscillator driven by stochastic atmospheric forcing. *J. Climate*, **8**, 2440–2453.
- , and K. Bryan, 1997: Predictability of North Atlantic multidecadal climate variability. *Science*, **275**, 181–184.
- Halliwel, G. R., 1998: Simulation of North Atlantic decadal/multidecadal winter SST anomalies driven by basin-scale atmospheric circulation anomalies. *J. Phys. Oceanogr.*, **28**, 5–21.
- Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93**, 11 015–11 021.
- Hawkins, E., and R. Sutton, 2009: Decadal predictability of the Atlantic Ocean in a coupled GCM: Forecast skill and optimal perturbations using linear inverse modeling. *J. Climate*, **22**, 3960–3978.
- , J. Robson, R. Sutton, D. Smith, and N. Keenlyside, 2011: Evaluating the potential for statistical decadal predictions of sea surface temperatures with a perfect model approach. *Climate Dyn.*, **37**, 2495–2509.
- Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res.*, **103**, 18 567–18 589.
- Kushnir, Y., 1994: Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Climate*, **7**, 141–157.
- Marshall, J., H. Johnson, and J. Goodman, 2001: A study of the interaction of the North Atlantic oscillation with ocean circulation. *J. Climate*, **14**, 1399–1421.
- Moore, A. M., and R. Kleeman, 1996: The dynamics of error growth and predictability in a coupled model of ENSO. *Quart. J. Roy. Meteor. Soc.*, **122**, 1405–1446.
- , and —, 2001: The differences between the optimal perturbations of coupled models of ENSO. *J. Climate*, **14**, 138–163.
- Newman, M., 2007: Interannual to decadal predictability of tropical and North Pacific sea surface temperatures. *J. Climate*, **20**, 2333–2356.
- Palmer, T. N., 1986: Influence of Atlantic, Pacific and Indian Oceans on Sahel rainfall. *Nature*, **322**, 251–253.
- Penland, C., and L. Matrosova, 1994: A balance condition for stochastic numerical-models with application to the El Niño–Southern Oscillation. *J. Climate*, **7**, 1352–1372.
- , and P. D. Sardeshmukh, 1995: The optimal-growth of tropical sea-surface temperature anomalies. *J. Climate*, **8**, 1999–2024.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.
- , P. Brohan, D. E. Parker, C. K. Folland, J. J. Kennedy, M. Vanicek, T. J. Ansell, and S. F. B. Tett, 2006: Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The HadSST2 dataset. *J. Climate*, **19**, 446–469.
- Schlesinger, M. E., and N. Ramankutty, 1994: An oscillation in the global climate system of period 65–70 years. *Nature*, **367**, 723–726.
- Sutton, R. T., and D. L. R. Hodson, 2005: Atlantic Ocean forcing of North American and European summer climate. *Science*, **309**, 115–118.
- Ting, M. F., Y. Kushnir, R. Seager, and C. H. Li, 2009: Forced and internal twentieth-century SST trends in the North Atlantic. *J. Climate*, **22**, 1469–1481.
- Tziperman, E., L. Zanna, and C. Penland, 2008: Nonnormal thermohaline circulation dynamics in a coupled ocean–atmosphere GCM. *J. Phys. Oceanogr.*, **38**, 588–604.
- Vimont, D. J., 2012: Analysis of the Atlantic meridional mode using linear inverse modeling: Seasonality and regional influences. *J. Climate*, **25**, 1194–1212.
- Weaver, A. J., J. Marotzke, P. F. Cummins, and E. S. Sarachik, 1993: Stability and variability of the thermohaline circulation. *J. Phys. Oceanogr.*, **23**, 39–60.
- Wunsch, C., 2012: Covariances and linear predictability of the Atlantic Ocean. *Deep-Sea Res.*, in press.
- Zanna, L., P. Heimbach, A. M. Moore, and E. Tziperman, 2012: Upper ocean singular vectors of the North Atlantic climate with implications for linear predictability and variability. *Quart. J. Roy. Meteor. Soc.*, **138**, 500–513.
- Zhao, M., I. M. Held, S. J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50-km resolution GCM. *J. Climate*, **22**, 6653–6678.