



# AMERICAN METEOROLOGICAL SOCIETY

*Journal of Climate*

## **EARLY ONLINE RELEASE**

This is a preliminary PDF of the author-produced manuscript that has been peer-reviewed and accepted for publication. Since it is being posted so soon after acceptance, it has not yet been copyedited, formatted, or processed by AMS Publications. This preliminary version of the manuscript may be downloaded, distributed, and cited, but please be aware that there will be visual differences and possibly some content differences between this version and the final published version.

The DOI for this manuscript is doi: 10.1175/JCLI-D-11-00539.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Zanna, L., 2012: Forecast Skill and Predictability of Observed Atlantic Sea Surface Temperatures. *J. Climate*. doi:10.1175/JCLI-D-11-00539.1, in press.

© 2012 American Meteorological Society



1 **Forecast Skill and Predictability of observed Atlantic sea surface**  
2 **temperatures**

3 **LAURE ZANNA \***

*Atmospheric, Oceanic and Planetary Physics, University of Oxford, UK*

PRELIMINARY ACCEPTED VERSION

---

\* *Corresponding author address:* Laure Zanna, Atmospheric, Oceanic and Planetary Physics, Dept of Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK.  
email:zanna@atm.ox.ac.uk

## ABSTRACT

4  
5 An empirical statistical model is constructed to assess the forecast skill and the linear pre-  
6 dictability of Atlantic sea surface temperatures (SST) variability. Linear inverse modeling  
7 (LIM) is used to build the statistical model based on observed Atlantic SST anomalies be-  
8 tween latitudes 20S and 66N from 1870 to 2009. LIM allows to fit and test a multivariate  
9 red noise model to the observed annually averaged SST anomalies. Forecast skill is assessed  
10 and shown to be on the order of 3 to 5 years. After a few years, the skill is greatly reduced  
11 especially in the subpolar region. In the stable dynamical system determined by LIM, skill  
12 of annual average SST anomalies arises from four damped eigenmodes. The four eigen-  
13 modes are shown to be relevant in particular for the optimal growth events of SST variance  
14 with a pattern reminiscent of the low-frequency mode of variability, and in general for the  
15 predictability and variability of Atlantic SSTs on interannual timescales. LIM might be a  
16 useful benchmark for interannual and decadal forecasts of SST anomalies based on numerical  
17 models.

# 1. Introduction

In the Atlantic, 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 timescales. The possible 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). Despite the large number of studies, it is still difficult to exclude any of the proposed mechanisms due in part to the relatively short historical record. In numerical model simulations, the timescales 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 has obviously many economical and societal consequences as it is associated for example with rainfall in Sahel (Palmer 1986; Battisti et al. 2008) and hurricane activity (Zhao et al. 2009) and 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 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 Models(GCM)-based tools being developed for regional climate predictions are rather complex, challenging and imperfect, the observed record has been mostly ignored when it comes to determining forecast skill

45 and predictability on interannual timescales. The main goal of this study is to extract the  
46 statistical properties of the observed record in order to investigate the dynamics of Atlantic  
47 SSTa and their forecast skill.

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

## 53 **2. Linear Inverse Modeling**

### 54 *a. Data*

55 While numerical climate models are extremely useful, they often do not reproduce the  
56 observed variability properly and introduce structural errors difficult to quantify. To inves-  
57 tigate the variability and forecast skill of Atlantic SST anomalies (SSTa), the Met Office  
58 Hadley Centre’s SST raw dataset HadSST2 on a  $5^\circ$  by  $5^\circ$  latitude-longitude grid from 1870  
59 to 2009 (Rayner et al. 2006) is used.

60 The SST anomalies are constructed by removing the climatology over the entire period  
61 considered and the trend. Separating between the natural and anthropogenically forced  
62 variability is the subject of active research. Methods attempting to separate between the  
63 oscillatory-like component and the upward trend are still largely debated (e.g., Ting et al.  
64 2009; DelSole et al. 2011), we therefore opt for a simple solution. The data at each grid  
65 was detrended using two methods: removing the linear trend as commonly done in various  
66 studies (Enfield et al. 2001; Sutton and Hodson 2005), or using a cubic spline with two  
67 breakpoints similarly to Hawkins et al. (2011). The results presented in this study are using  
68 the latter (cubic spline) however the main conclusions remain unchanged if the linear trend  
69 is removed. Yet, perhaps a part of the external forced component of SST variability is still

70 present and could potentially affect the predictions.

71 The domain is confined to the Atlantic basin between latitudes 20S and 66N for several  
72 reasons: the frequency of SST measurements during 1870-2009 is relatively high; the im-  
73 pact of sea-ice variability on SSTa is minimized compared to all other regions; the possible  
74 modulations of mid-latitudes SSTa from the tropics due to the large-scale ocean circulation  
75 are taken into account. The timeseries at each grid point are averaged from January to  
76 December of each year to obtain annual averages, only years having 10 or more months are  
77 retained. Grid-cells with less than 75% of the years are omitted. The high-latitudes de-  
78 spite being important in Atlantic variability are left out due to their poor sampling over the  
79 period considered and as mentioned the spurious effects of sea-ice. It would have been pos-  
80 sible to use interpolated datasets such as HadISST (Rayner et al. 2003) or RSA MOHSST5  
81 (Kaplan et al. 1998) instead of the raw data to include the high-latitudes (e.g., Labrador  
82 and Greenland Seas), however these datasets are constructed using empirical orthogonal  
83 functions optimal interpolation methods which could potentially comprise the analysis.

84 *b. Estimation of the linear inverse model*

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

93 The main assumption underlying LIM is that the time-dependent annual average SSTa  
94 can be separated into a linear deterministic part and a nonlinear part represented by a  
95 linearized term dependent on the base state and white noise. To reduce the number of degrees

96 of freedom, we construct a reduced space for SSTa using empirical orthogonal functions  
 97 (EOFs) as the spatial fields and their respective principal components (PCs) as the time  
 98 evolution of the EOFs. The data is weighted by the surface area of the grid-box and their  
 99 standard deviation (Moore and Kleeman 2001). Figure 1 shows the spatial patterns of the  
 100 first three EOFs explaining 32%, 10% and 7.7% of the variability and are reminiscent of  
 101 many studies (e.g., Deser and Blackmon 1993).

102 The evolution of the PCs can therefore be written as a linear dynamical system forced  
 103 with white stochastic forcing ( $\eta$ ) such that

$$104 \quad \frac{d\mathbf{P}}{dt} = \mathbf{A}\mathbf{P} + \eta. \quad (1)$$

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

$$108 \quad \mathbf{A} = \frac{\ln[\mathbf{C}(\tau_0)\mathbf{C}(0)^{-1}]}{\tau_0}, \quad (2)$$

109 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  $\langle \cdot \rangle$   
 110 denotes an ensemble average or an average over all  $t$  if the statistics of  $\eta$  are stationary. Only  
 111 13 EOFs/PCs, explaining about 79% of the variability, are retained to evaluate  $\mathbf{A}$ , ensuring  
 112 that  $\mathbf{A}$  remains well-defined with negative eigenvalues and a relatively low condition number.

### 113 *c. Eigen-analysis of $\mathbf{A}$*

114 To verify that the LIM adequately describes the Atlantic SSTa variability, several tests  
 115 can be performed. To explore the properties of the LIM and its spectrum, the eigenvalues  
 116  $\gamma_j = \alpha_j + i\beta_j$  and corresponding eigenvectors  $\mathbf{v}_j$  (or empirical normal modes) of  $\mathbf{A}$  are com-  
 117 puted (Penland and Sardeshmukh 1995, hereafter PS95). The eigenmodes are often referred  
 118 to as Principal Oscillation Patterns (POPs) (Hasselmann 1988).  $\mathbf{A}$  must be dissipative with  
 119 negative eigenvalues for the LIM to have stationary statistics and to be a valid model for  
 120 the SSTa dynamics, which is confirmed by the tests performed on the eigen-analysis of  $\mathbf{A}$ .

121 Table 1 shows the decay and oscillatory timescales of the different eigenmodes. Some of  
 122 the decaying eigenmodes oscillates with periods ranging from 8.5 to 137 years. Note that  
 123 perhaps some of the dynamical properties of the system are poorly represented due to the  
 124 short timeseries or to sampling issues. For example, similar results for the eigenanalysis of  
 125 the matrix  $\mathbf{A}$  were obtained when evaluated for lead times up to  $\tau_0 = 4 \text{ yr}$ , however this  
 126 is not the case for larger lead times. In this event, we run into the Nyquist problem for  
 127 which the lead time  $\tau_0$  is close to half the period of one the eigenmodes (eigenmode 2/3 in  
 128 Table 1) and the eigenmodes cannot be reliably estimated. Different EOF truncations lead  
 129 to qualitatively similar results with some quantitative changes in the eigenvalues.

130 *d. Multivariate model of SSTa*

131 To further validate that Eq. 1 is an adequate model of the observed Atlantic SSTa,  
 132 the LIM should in part reproduce the SSTa power spectrum. The noise forcing  $\eta$  and  
 133 its statistics can be evaluated from the fluctuation-dissipation relationship (PS95), that is  
 134  $\mathbf{A}\mathbf{C}(0) + \mathbf{C}(0)\mathbf{A}^T + \langle \eta\eta^T \rangle dt = 0$ . The covariance matrix of the forcing,  $\langle \eta\eta^T \rangle dt = \mathbf{Q}$ ,  
 135 has to be positive definite. All the eigenvalues of  $\mathbf{Q}$  are found to be positive except for  
 136 one negative with a small amplitude. The higher order EOFs are possibly not reliably  
 137 estimated from the (noisy) data leading to the small negative eigenvalue of  $\mathbf{Q}$  since using  
 138 less EOFs leads to a purely positive definite matrix  $\mathbf{Q}$ . Eq. 1 is integrated forward for  
 139 49000 years (Penland and Matrosova 1994) with a timestep of 1 hour with  $\eta$  generated using  
 140 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  
 141 independent Gaussian white noises with unit variance.

142 The LIM reproduces the main features of the observed power spectrum of the domain  
 143 averaged Atlantic SSTa as shown in Fig. 2a. The LIM spectrum is as expected much smoother  
 144 than the observed spectrum, due to the relatively few degrees of freedom in the truncated  
 145 EOF-space. Some of the irregularities could also be due to the poor sampling and errors  
 146 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  $\mathbf{B}(\tau) = e^{\mathbf{A}\tau} = e^{\frac{\tau}{\tau_0} \ln \frac{\mathbf{C}(\tau_0)}{\mathbf{C}(0)}}$ . 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 independent of that used to determine  $\mathbf{A}$  which is done as follows: the data is 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; this procedure was repeated every year.

Figure 2b shows the reconstructed domain-averaged Atlantic SSTa timeseries in the EOF-space (black curve) and associated forecasts (grey curves) using the LIM for lead times of up to 10 years, plotted at 1-year 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 to an autoregressive (AR1) process (equivalent to damped persistence), shown in Figure 2b. 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)

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

The estimates of the forecast error covariance provide an estimate of the statistics of the noise forcing  $\eta$  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 of their theoretical estimates at a lead time of 5 years. This does not necessarily invalidate the linearity assumption. While the differences between the actual and theoretical errors can be

173 seen as a measure of the nonlinearities of the system, they could also be due to the severe  
174 EOF truncation or the non-stationarity of the data.

175 Maps of mean square errors relative to climatology (using the climatology from HadSST2  
176 and simply assuming  $\hat{\mathbf{P}}(t+\tau) = 0$ ), shown in Fig 3, indicate the regions for which the forecast  
177 skill is greatly reduced. For years 1-2, the forecast skill in the Atlantic basin is largely  
178 superior to climatology. For years 3-5, the forecast skill decreases south of Greenland, in  
179 the subpolar region. For longer lead times, the model fails in the same subpolar region and  
180 skill is lost in the tropics, especially near the African coast. In the subpolar region, the  
181 variance of SSTa is large and ocean-atmosphere coupling in addition to ocean mixed-layer  
182 processes are strong leading to the possible reduction of the forecast skill of the LIM. If the  
183 anomalies were constructed using averages running from July to June (instead of January to  
184 December), the main conclusions would still hold however the forecast skill would be reduced  
185 by roughly 10%. Including the full winter seems to have a negative effect on the forecast skill,  
186 again potentially due to the large air-sea interaction or sporadic ocean variability during this  
187 season.

188 *b. Transient growth, optimal initial conditions and predictability*

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

197 In several GCM experiments, significant transient amplification of annual mean Atlantic  
198 SSTa is found on timescales of 4 to 20 years (Zanna et al. 2011; Tziperman et al. 2008;

199 Hawkins and Sutton 2009). For the current LIM based on observations, the largest possible  
 200 transient growth of the domain-integrated Atlantic SST variance without forcing is given by  
 201  $\lambda(\tau) = \mathbf{P}(\tau)^T \mathbf{P}(\tau) / \mathbf{P}(0)^T \mathbf{P}(0)$  and shown in Fig. 4a as function of the lead time  $\tau$ .  $\lambda(\tau)$   
 202 is the leading singular value of  $\mathbf{B}(\tau)$  such that the maximum amplification of SST variance  
 203 is found at about  $\tau = 3$  (to 4) yrs with an amplitude of roughly 8. Further exploration of  
 204 the singular values of  $\mathbf{B}(\tau = 3 \text{ yrs})$  reveals that three additional singular values are larger  
 205 than unity therefore exhibiting possible linear growth; those three additional singular values  
 206 are ranging between 1.5 and 4 therefore leading to significantly less amplification than the  
 207 leading singular value. The maximum amplification curve in Fig. 4a implies that the actual  
 208 growth of variance from the interference of non-orthogonal eigenmodes can be sustained for  
 209 more than 10 years, actually up to 18 years (time at which  $\lambda$  decreases to unity). This  
 210 timescale can be viewed as an optimistic estimate or an upper bound on the predictability  
 211 time of the linear events without forcing. However, given that the system is not a perfect  
 212 linear system and that forcing (noise) is present, predictability is lost due to the white noise  
 213 forcing. In fact, for the domain-averaged SSTa to be predictable,  $\lambda(\tau)$  has to be larger than  
 214 the error covariance  $1 + \langle \mathbf{e}(\tau) \mathbf{e}(\tau)^T \rangle$  and the limit for predictability using this criteria is  
 215 slightly under 10 years. The other way to think about the growth as mentioned earlier is  
 216 in term of error variance growth such that the rapid increase in SST variance over 3 to 4  
 217 years could lead to a limited predictability due to errors in initial conditions and potentially  
 218 explain the forecast skill of 3 to 5 yrs, especially in the subpolar regions.

219 While the system seems to be able to sustain SST variance growth, it is interesting  
 220 to explore if these potentially predictable linear growth events exist in the record and to  
 221 determine their spatial structure. The initial conditions leading to the maximum growth  
 222 at a given time  $\tau$  are given by the leading singular vector of  $\mathbf{B}(\tau)$ , often termed optimal  
 223 initial conditions (Farrell and Ioannou 1996). Optimal perturbations can be relevant to the  
 224 variability of the observed system; for example, PS95 showed that the leading singular vector  
 225 amplifying the tropical Pacific SST corresponds to the most efficient initial condition for the

226 development of ENSO.

227 The spatial pattern of the optimal initial perturbations corresponding to the largest  
228 growth (obtained for  $B(\tau = 3 \text{ yrs})$ ) is shown in Fig. 5a. The initial state of the optimal  
229 perturbations (Fig. 5a) exhibits several signals in Northern Hemisphere: in the vicinity  
230 of the Gulf Stream and its extension, the southern tip of Greenland, the Gulf of Mexico  
231 and the Canary Islands; and an additional anomaly in the Southern Hemisphere around  
232 20S. The shape of the leading singular vector at different  $\tau$  does not change very much  
233 with always a relatively pronounced anomaly in the middle of the basin, in vicinity of the  
234 Gulf Stream's extension. Therefore the LIM has only one main leading optimal initial  
235 structure. The evolution of the perturbations (Fig. 5) over the optimal growth timescale  
236 hints at a non-local growth of the perturbations (at least over the first year) with possible  
237 mechanisms involving ocean-atmosphere interaction, mixed-layer processes, re-emergence of  
238 SST, and advection by the large-scale ocean circulation. The optimally evolved pattern  
239 (Fig. 5d), reminiscent of the first EOF and the low-frequency mode of variability in the  
240 Atlantic (Schlesinger and Ramankutty 1994), has a single sign over the Northern Hemisphere  
241 with the strongest anomalies in the subpolar gyre region and just north of the tropics. In  
242 the Southern Hemisphere, a small anomaly of opposite sign has developed and is centered  
243 around 27W and 20S. The optimal perturbation found can be seen as a precursor event to  
244 the large warming (or cooling) over the basin as shown in Fig. 5d.

245 When analysing optimal structures, it is important to consider whether the growth is  
246 actually observed to occur as predicted by the LIM. Figure. 4b shows the projection of  
247 observed anomalies onto the optimal initial conditions vs. their projection onto the predicted  
248 evolved structure 3 years later. The remarkable agreement between the slopes of least-square  
249 fit and the linear amplification of SST,  $\sqrt{\lambda}(\tau = 3\text{yrs})$ , in addition to correlation of 0.6 are  
250 indications that optimal SST growth is relatively well captured by the LIM and relevant for  
251 observed Atlantic variability. As expected, some scatter is observed about the least-square  
252 fit due to the noise  $\eta$  reducing the correlation.

253 *c. Dynamical eigenmodes of the Atlantic variability*

254 A vast number of studies have been looking at the temporal and spatial variations of  
255 the Atlantic Multidecadal Variability or Atlantic Multidecadal Oscillation (AMO) and their  
256 impact on climatic phenomena such as rainfall or hurricane activity (Enfield et al. 2001;  
257 Sutton and Hodson 2005). The Atlantic multidecadal variability is characterized by basin-  
258 wide fluctuations in North Atlantic SST anomalies, this SSTa pattern of variability exhibits  
259 some similarities with the optimally excited pattern shown in Fig. 5d and is found in several  
260 numerical models.

261 In the present section, we consider how the unsmoothed AMO index timeseries and  
262 the optimal patterns project onto the eigenmodes of the linearized matrix  $\mathbf{A}$  to gain some  
263 insight about the source of predictability of Atlantic SSTa. The AMO index timeseries are  
264 calculated from the monthly Kaplan SST dataset. The index is an area weighted average  
265 of linearly detrended SST anomalies over the North Atlantic between 0 to 70N. The index  
266 was obtained from <http://www.cdc.noaa.gov/data/timeseries/AMO/>. The AMO index is  
267 simply a measure of the domain-averaged annually averaged SST variability (and not defined  
268 as the leading EOF/PC explaining most of the variance unlike the Pacific Decadal Oscillation  
269 in the North Pacific). The 5-yr running mean of the temporal evolution of the AMO index  
270 is often used and seems to suggest shifts from one sign to the other in Atlantic SSTa on  
271 multidecadal timescales (70 to 80 years as in Kushnir (1994)). Due to the presence of the  
272 external anthropogenic forcing and our difficulty to adequately remove the forced component  
273 of SST fluctuations to isolate the internal variability, different studies define different periods  
274 as a shift in the sign of the AMO index. We will not worry about such issues in the present  
275 work and only examine year to year variations.

276 The regression coefficients  $\mathbf{R}_{AMO}$  between the PCs and the AMO index are found sim-  
277 ilarly to the procedure described in (Tziperman et al. 2008). Therefore the AMO index  
278 reconstructed from the PCs is given by  $\tilde{AMO} = \mathbf{R}_{AMO}^T \mathbf{P}$ . The correlation between the  
279 original AMO index timeseries and  $\tilde{AMO}$  is roughly 0.91. The time dependent PCs,  $\mathbf{P}(t)$ ,

280 can be written as the sum of the contributions of the different eigenmodes of  $\mathbf{A}$  such that

$$281 \quad \mathbf{P}(t) = \sum_{j=1}^{j=13} \phi_j(t) \mathbf{v}_j. \quad (4)$$

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

$$283 \quad \phi_j(t) = \frac{\mathbf{g}_j^* \mathbf{P}(t)}{\mathbf{g}_j^* \mathbf{v}_j} \quad (5)$$

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

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

297 Looking at Fig. 6, one could assume that the oscillation of eigenmode 5/6 would lead  
298 the optimal growth of the anomalies and possibly define the evolution of the AMO index.  
299 However, the variance explained by eigenmode 5/6 is only half the observed AMO variance  
300 and the timeseries of the eigenmode 5/6 correlates only at 0.43 with the AMO index. Similar  
301 results are obtained for the evolution of the optimal pattern.

302 The projection of the AMO index timeseries onto the eigenmodes reveals that the same  
303 four eigenmodes are necessary to reproduce the evolution of the AMO index. This is perhaps  
304 not surprising as the AMO index is closely related to the SST variance norm used to maximize  
305 the optimal growth. The reconstructed AMO with the four eigenmodes and  $\tilde{AMO}$  timeseries

306 are shown in Fig. 7 and found to have a correlation of approximately 0.89. It therefore  
307 appears that the skill and predictability of Atlantic SSTa arises from the decaying timescales  
308 of the participating eigenmodes and not their oscillatory timescales similarly to the PDO  
309 variability (Newman 2007).

## 310 4. Discussion and Conclusions

311 Using the Linear Inverse Modeling (LIM) approach, we constructed a statistical model of  
312 annually averaged Atlantic SST anomalies. The main results can be summarized as follows:

- 313 • SST variability in the Atlantic domain can be well captured by linearized dynamics  
314 forced with white noise;
- 315 • the LIM captures the main features of Atlantic SST anomalies including their power  
316 spectrum and optimal growth events;
- 317 • linear forecasts of SSTa are successful for up to 5 years after which the skill deteriorates;  
318 using a 16 year-long timeseries of surface temperature in the vicinity of the Grand  
319 Banks, Wunsch (2012) shows that some linear predictive skill is possible up to about  
320 5 years, similarly to our results.
- 321 • the forecast skill is mainly lost in the subpolar region south of Greenland, possibly due  
322 to the nonlinearities arising from air-sea interaction and/or subduction processes;
- 323 • optimal patterns develop into large anomalies reminiscent of the 1st EOFs of the system  
324 and low-frequency SST patterns found in models (DelSole et al. 2011).
- 325 • linear growth events and AMO variability can be explained by four eigenmodes of the  
326 linearized operator, the four non-orthogonal eigenmodes are responsible for the skill  
327 and the predictability in the Atlantic sector.

328 Yet, 140 years of data might still be too short to adequately explore the interannual  
329 (and multidecadal) variability and predictability of the Atlantic especially given the large  
330 uncertainties in observations during the first half of the record. The model and its skill  
331 could perhaps be improved by adding additional datasets such as sea level pressure, salinity  
332 or sub-surface temperature or other temperatures from other ocean basins. However careful  
333 analysis will be necessary as these datasets are short and extremely noisy (even more noisy  
334 that the one used in the present study).

335 While statistical models are useful, they have several limitations. It is extremely difficult  
336 to investigate the physical mechanisms and to confirm that the results obtained are real  
337 and not an artifact of the statistics or the short timeseries. For example, it is difficult to  
338 compare the eigenmodes or singular vectors obtained to numerical studies as most of them  
339 concentrate on the EOFs and not the dynamical modes. We hope to address these issues  
340 in future work with a designed set of experiments using idealized and complex numerical  
341 models in addition to observations. We expect to identify the dynamical eigenmodes and  
342 explore their relationship to the skill and variability in the Atlantic and compare them with  
343 the observations.

344 The results suggests that statistical models such as LIM could provide valuable informa-  
345 tion regarding regional predictability and variability before using complex nonlinear tools,  
346 including GCM-based tools. LIM as a global or regional tool could serve as a benchmark for  
347 interannual and decadal climate forecasts.

348 *Acknowledgments.*

349 Many thanks to Ed Hawkins and David Munday for useful discussions and to Cecile Pen-  
350 land and two additional reviewers for their comments which helped improve the manuscript.  
351 LZ was supported by a Fellowship from the Oxford Martin School and Balliol College, Uni-  
352 versity of Oxford.

## REFERENCES

- 355 Battisti, D. S., U. S. Bhatt, and M. A. Alexander, 1995: A modeling study of the interannual  
356 variability in the wintertime North Atlantic Ocean. *J. Climate*, **8**(12), 3067–3083.
- 357 Biasutti, M., I. M. Held, A. H. Sobel, and A. Giannini, 2008: SST forcings and Sahel rainfall  
358 variability in simulations of the twentieth and twenty-first centuries. *J. Climate*, **21**(14),  
359 3471–3486.
- 360 Bjerknes, J., 1964: Atlantic air–sea interaction. *Adv. Geophys.*, **10**, 1–82.
- 361 Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global  
362 circulation. *J. Atmos. Sci.*, **52**(9), 1434–1456.
- 363 Chen, F., and M. Ghil, 1995: Interdecadal variability of the thermohaline circulation and  
364 high–latitude surface fluxes. *J. Phys. Oceanogr.*, **25**, 2547–2568.
- 365 Czaja, A., and J. Marshall, 2001: Observations of atmosphere-ocean coupling in the North  
366 Atlantic. *Q. J. R. Meteorol. Soc.*, **127**(576), 1893–1916.
- 367 DelSole, T., M. K. Tippett, and J. Shukla, 2011: A significant component of unforced  
368 multidecadal variability in the recent acceleration of global warming. *J. Climate*, **24**(3),  
369 909–926.
- 370 Deser, C., and M. L. Blackmon, 1993: Surface climate variations over the North Atlantic  
371 ocean during winter: 1900–1989. *J. Climate*, **6**, 1743–1753.
- 372 Enfield, D. B., A. M. Mestas-nunez, and P. J. Trimble, 2001: The Atlantic multidecadal  
373 oscillation and its relation to rainfall and river flows in the continental us. *Geophys. Res.  
374 Lett.*, **28**(10), 2077–2080.

- 375 Farrell, B., 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.*, **46**, 1193–1206.
- 376 Farrell, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*,  
377 **39**(8), 1663–1686.
- 378 Farrell, B. F., and P. J. Ioannou, 1996: Generalized stability theory Part I: Autonomous  
379 operators. *J. Atmos. Sci.*, **53**, 2025–2040.
- 380 Griffies, S. M., and K. Bryan, 1997: Predictability of North Atlantic multidecadal climate  
381 variability. *Science*, **275**(5297), 181–184.
- 382 Griffies, S. M., and E. Tziperman, 1995: A linear thermohaline oscillator driven by stochastic  
383 atmospheric forcing. *J. Climate*, **8**(10), 2440–2453.
- 384 Halliwell, G. R., 1998: Simulation of North Atlantic decadal/multidecadal winter SST  
385 anomalies driven by basin-scale atmospheric circulation anomalies. *J. Phys. Oceanogr.*,  
386 **28**(1), 5–21.
- 387 Hasselmann, K., 1988: PIPs and POPs: the reduction of complex dynamical systems using  
388 principal interaction and Oscillation patterns. *J. Geophys. Res.*, **93**, 11,015–11,021.
- 389 Hawkins, E., and R. Sutton, 2009: Decadal predictability of the Atlantic ocean in a coupled  
390 GCM: Forecast skill and optimal perturbations using linear inverse modeling. *J. Climate*,  
391 **22**(14), 3960–3978.
- 392 Hawkins, E., J. Robson, R. Sutton, D. Smith, and N. Keenlyside, 2011: Evaluating the  
393 potential for statistical decadal predictions of SSTs with a perfect model approach. *Clim.*  
394 *Dyn.*. Submitted.
- 395 Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan,  
396 1998: Analyses of global sea surface temperature 1856-1991. *J. Geophys. Res. - Oceans*,  
397 **103**(C9), 18,567–18,589.

- 398 Kushnir, Y., 1994: Interdecadal variations in North Atlantic sea surface temperature and  
399 associated atmospheric conditions. *J. Climate*, **7**, 141–157.
- 400 Moore, A. M., and R. Kleeman, 1996: The dynamics of error growth and predictability in  
401 a coupled model of ENSO. *Q. J. R. Meteor. Soc.*, **122**, 1405–1446.
- 402 Moore, A. M., and R. Kleeman, 2001: The differences between the optimal perturbations of  
403 coupled models of ENSO. *J. Climate*, **14**(2), 138–163.
- 404 Newman, M., 2007: Interannual to decadal predictability of tropical and North Pacific sea  
405 surface temperatures. *J. Climate*, **20**(11), 2333–2356.
- 406 Palmer, T. N., 1986: Influence of Atlantic, Pacific and Indian Oceans on Sahel rainfall.  
407 *Nature*, **322**, 251–253.
- 408 Penland, C., and L. Matrosova, 1994: A balance condition for stochastic numerical-models  
409 with application to the El Niño–Southern Oscillation. *J. Climate*, **7**(9), 1352–1372.
- 410 Penland, C., and P. D. Sardeshmukh, 1995: The optimal-growth of tropical sea-surface  
411 temperature anomalies. *J. Climate*, **8**(8), 1999–2024.
- 412 Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell,  
413 E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice,  
414 and night marine air temperature since the late nineteenth century. *J. Geophys. Res. -  
415 Atmos.*, **108**(D14).
- 416 Rayner, N. A., P. Brohan, D. E. Parker, C. K. Folland, J. J. Kennedy, M. Vanicek, T. J.  
417 Ansell, and S. F. B. Tett, 2006: Improved analyses of changes and uncertainties in sea  
418 surface temperature measured in situ since the mid-nineteenth century: The HadSST2  
419 dataset. *J. Climate*, **19**(3), 446–469.
- 420 Schlesinger, M. E., and N. Ramankutty, 1994: An oscillation in the global climate system  
421 of period 65–70 years. *Nature*, **367**(6465), 723–726.

- 422 Sutton, R. T., and D. L. R. Hodson, 2005: Atlantic Ocean forcing of North American and  
423 European summer climate. *Science*, **309**(5731), 115–118.
- 424 Ting, M. F., Y. Kushnir, R. Seager, and C. H. Li, 2009: Forced and internal Twentieth-  
425 Century SST trends in the North Atlantic. *J. Climate*, **22**(6), 1469–1481.
- 426 Tziperman, E., L. Zanna, and C. Penland, 2008: Nonnormal thermohaline circulation  
427 dynamics in a coupled ocean-atmosphere GCM. *J. Phys. Oceanogr.*, **38**(3), 588–604.
- 428 Weaver, A. J., J. Marotzke, P. F. Cummins, and E. S. Sarachik, 1993: Stability and vari-  
429 ability of the thermohaline circulation. *J. Phys. Oceanogr.*, **23**, 39–60.
- 430 Wunsch, C., 2012: Covariances and Linear Predictability of the Atlantic Ocean. *Deep Sea*  
431 *Research*. In Press.
- 432 Zanna, L., P. Heimbach, A. M. Moore, and E. Tziperman, 2011: Upper ocean singular vectors  
433 of the North Atlantic climate with implications for linear predictability and variability. *Q.*  
434 *J. R. Meteorol. Soc.*, **x**, DOI: 10.1002/qj.937.
- 435 Zhao, M., I. M. Held, S. J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane cli-  
436 matology, interannual variability, and response to global warming using a 50-km resolution  
437 gcm. *J. Climate*, **22**(24), 6653–6678.

438 **List of Tables**

439     1     Eigenvalues of the linearized matrix  $A$  constructed using 13 EOFs.                     20

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

Mode	$\tau_r = \frac{1}{\alpha}$ [yrs]	$\tau_i = \frac{2\pi}{\beta}$ [yrs]
Mode 1	1.2	-
Modes 2/3	1.4	$\pm 8.5$
Modes 4	3.1	-
Mode 5/6	4.0	$\pm 13.4$
Modes 7/8	4.9	$\pm 21.1$
Mode 9/10	6.7	$\pm 36.8$
Mode 11/12	8.0	$\pm 137.1$
Mode 13	15.3	-

## List of Figures

- 440
- 441 1 First three EOFs of annually averaged SST anomalies in the Atlantic basin  
442 using the raw dataset HadSST2. The contour intervals are the same for each  
443 three plots and arbitrary. The fraction of variance explained by each EOF is  
444 indicated in parenthesis. 23
- 445 2 (a) Power Spectrum of Atlantic annual mean SSTa averaged over the domain  
446 20S-66N: observations (solid line) and LIM (dashed line). Grey shading de-  
447 notes the 95% confidence interval evaluated from a 49000-yr run of the LIM  
448 (the model output is separated into 350 segments of 140-year timeseries and  
449 the ensemble mean is used to estimate the power spectrum of the model and  
450 confidence interval); (b) Timeseries of Atlantic annual mean SSTa averaged  
451 over the domain 20S-66N for original dataset (black) and 10-yr segment pre-  
452 dictions using the LIM (grey); (c) Normalized domain-integrated SST error  
453 forecast for the LIM (grey), a perfect linear model (solid black) and an AR1  
454 model (dashed black). The errors are normalized by the trace of the covariance  
455 matrix  $\mathbf{C}(0)$ . 24
- 456 3 Mean Square Error relative to climatology (the climatology is simply deter-  
457 mined from the HadSST2 dataset with  $\hat{\mathbf{P}}(\tau) = 0$ ) for (a)  $t = 1 - 2$  years; (a)  
458  $t = 3 - 5$  years; and (c)  $t = 6 - 8$  years. 25
- 459 4 (a) Maximum amplification curve: Maximum amplification,  $\lambda$ , as function of  
460 lead time  $\tau$ ; (b) Projection of observations upon the optimal initial condition  
461 for amplification of SST anomalies in the LIM over a 3-year interval, versus the  
462 optimal evolved SST state 3 years later. Grey dashed line: Expected growth  
463  $\lambda^{0.5}(\tau = 3yr) = 2.45$ ; black dashed line: linear fit with linear correlation  $r = 0.6$ . 26

464	5	Time evolution of the optimal perturbations: (a) Optimal initial conditions,	
465		$t = 0$ ; (b) Optimals at $t = 1$ yr; (c) Optimals at $t = 2$ yr; (d) Optimals at	
466		maximum amplification time, $t = 3$ yr. Contour intervals are the same for	
467		all panels and the amplitude is arbitrary (due to the linearity of the optimal	
468		growth problem).	27
469	6	Spatial patterns (real and imaginary parts) associated with the four eigen-	
470		modes (4, 5/6, 7/8 and 9/10) necessary to explain the optimal growth and	
471		the AMO index timeseries. Contour intervals are the same for all panels.	28
472	7	Projection timeseries, $\phi_j(t)$ , of eigenmodes (a) 4, (b) 5/6, (c) 7/8 and (d)	
473		9/10; Timeseries of the AMO index (black) and the reconstructed AMO index	
474		(dashed grey) from the four eigenmodes (a) through (d).	29

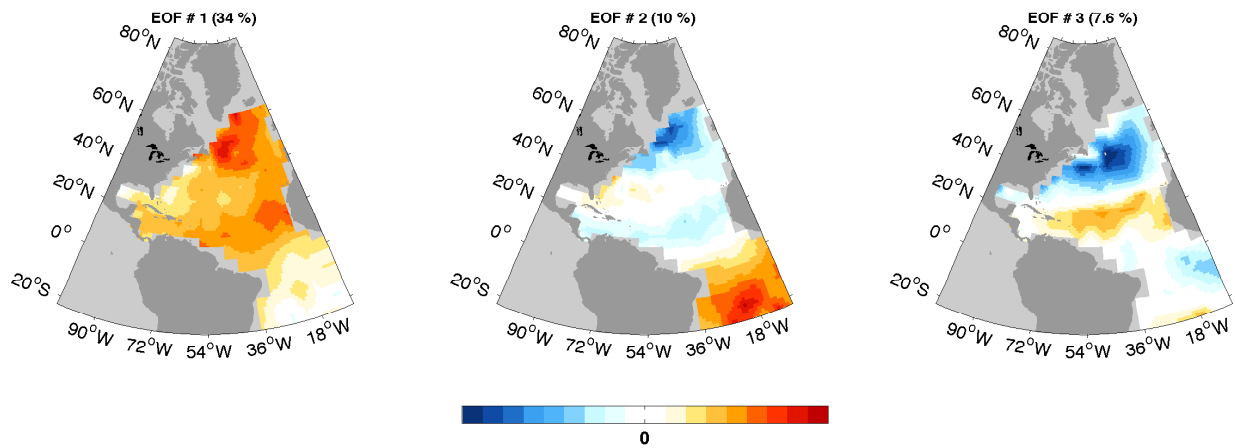


FIG. 1. First three EOFs of annually averaged SST anomalies in the Atlantic basin using the raw dataset HadSST2. The contour intervals are the same for each three plots and arbitrary. The fraction of variance explained by each EOF is indicated in parenthesis.

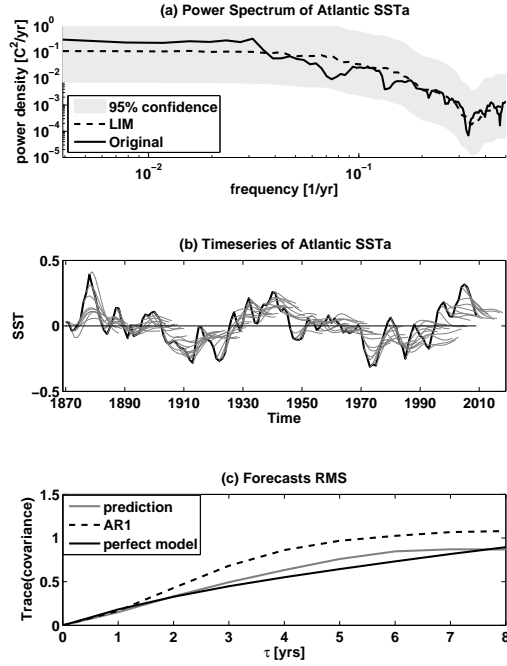


FIG. 2. (a) Power Spectrum of Atlantic annual mean SSTa averaged over the domain 20S-66N: observations (solid line) and LIM (dashed line). Grey shading denotes the 95% confidence interval evaluated from a 49000-yr run of the LIM (the model output is separated into 350 segments of 140-year timeseries and the ensemble mean is used to estimate the power spectrum of the model and confidence interval); (b) Timeseries of Atlantic annual mean SSTa averaged over the domain 20S-66N for original dataset (black) and 10-yr segment predictions using the LIM (grey); (c) Normalized domain-integrated SST error forecast for the LIM (grey), a perfect linear model (solid black) and an AR1 model (dashed black). The errors are normalized by the trace of the covariance matrix  $C(0)$ .

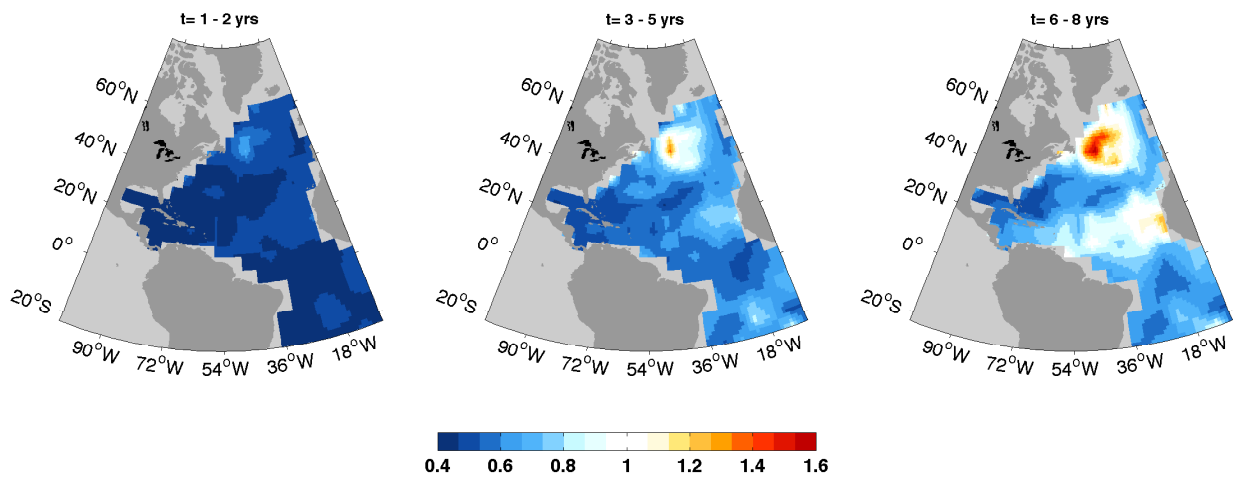


FIG. 3. Mean Square Error relative to climatology (the climatology is simply determined from the HadSST2 dataset with  $\hat{\mathbf{P}}(\tau) = 0$ ) for (a)  $t = 1 - 2$  years; (a)  $t = 3 - 5$  years; and (c)  $t = 6 - 8$  years.

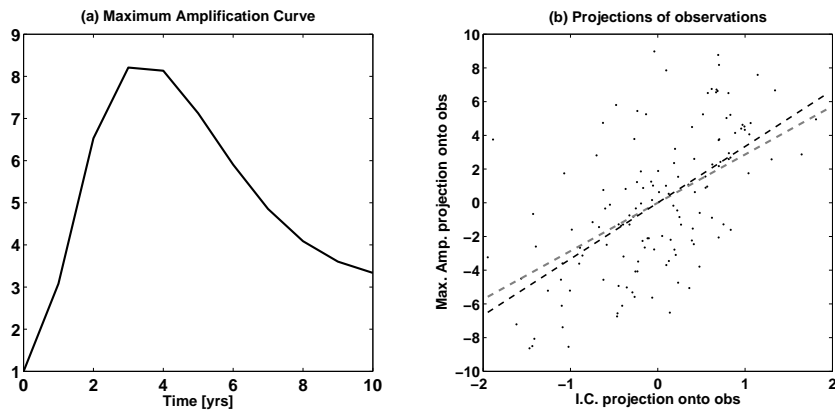


FIG. 4. (a) Maximum amplification curve: Maximum amplification,  $\lambda$ , as function of lead time  $\tau$ ; (b) Projection of observations upon the optimal initial condition for amplification of SST anomalies in the LIM over a 3-year interval, versus the optimal evolved SST state 3 years later. Grey dashed line: Expected growth  $\lambda^{0.5}(\tau = 3yr) = 2.45$ ; black dashed line: linear fit with linear correlation  $r = 0.6$ .

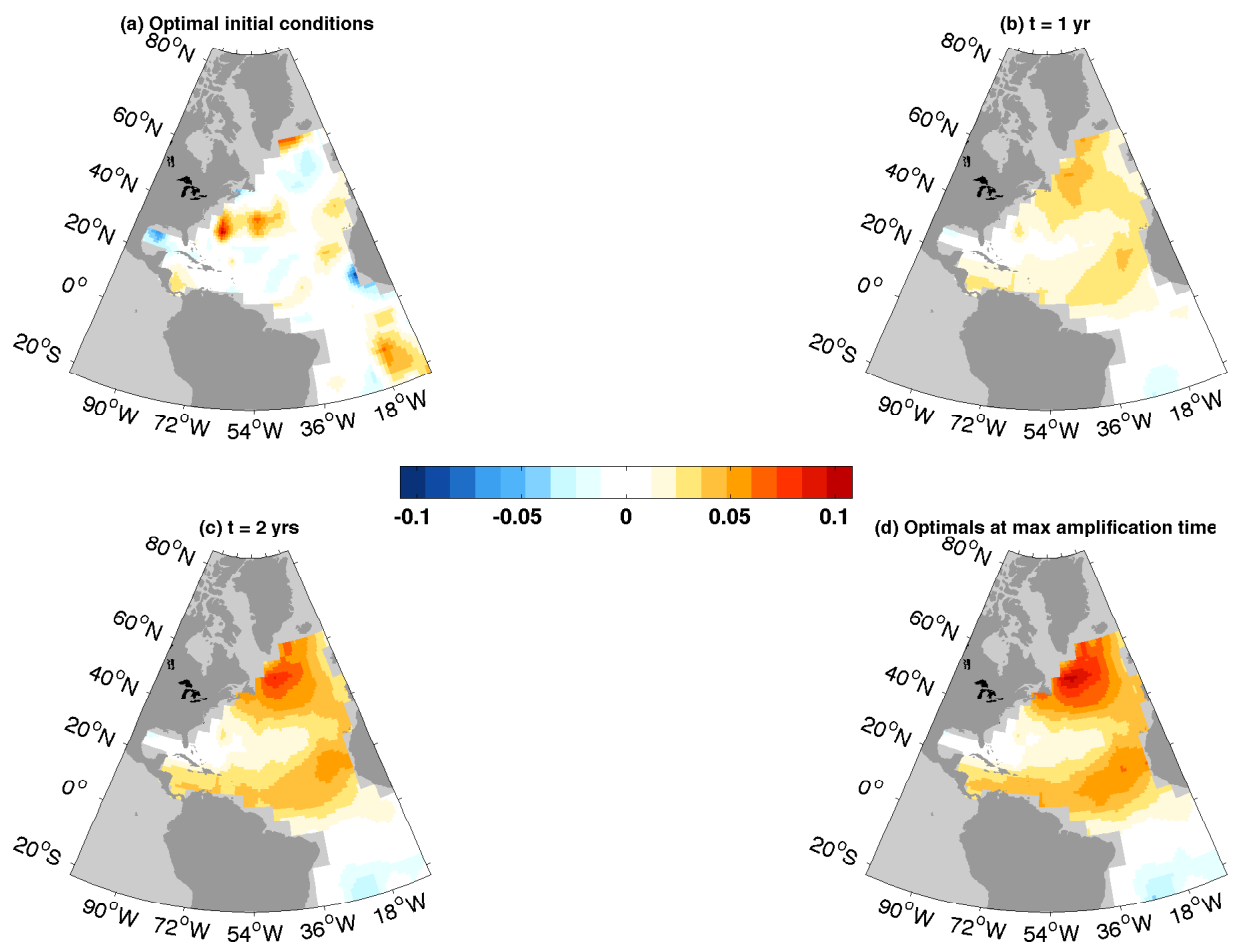


FIG. 5. Time evolution of the optimal perturbations: (a) Optimal initial conditions,  $t = 0$ ; (b) Optimals at  $t = 1$  yr; (c) Optimals at  $t = 2$  yr; (d) Optimals at maximum amplification time,  $t = 3$  yr. Contour intervals are the same for all panels and the amplitude is arbitrary (due to the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

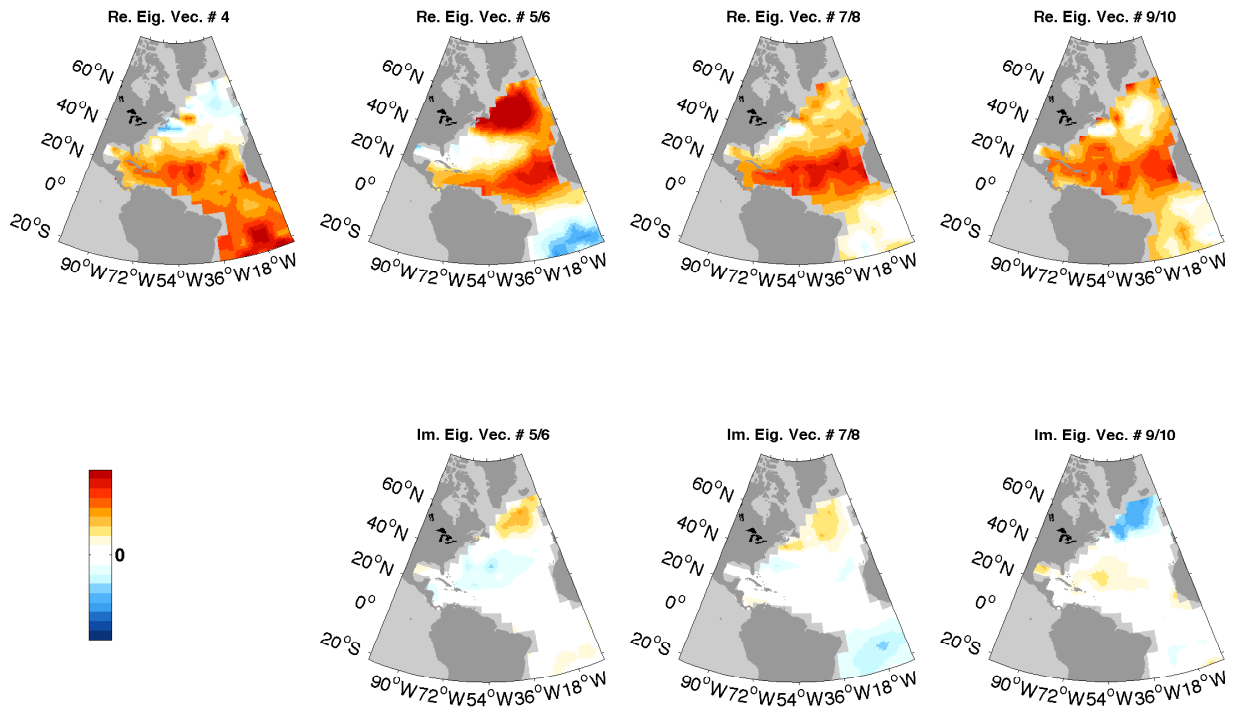


FIG. 6. Spatial patterns (real and imaginary parts) associated with the four eigenmodes (4, 5/6, 7/8 and 9/10) necessary to explain the optimal growth and the AMO index timeseries. Contour intervals are the same for all panels.

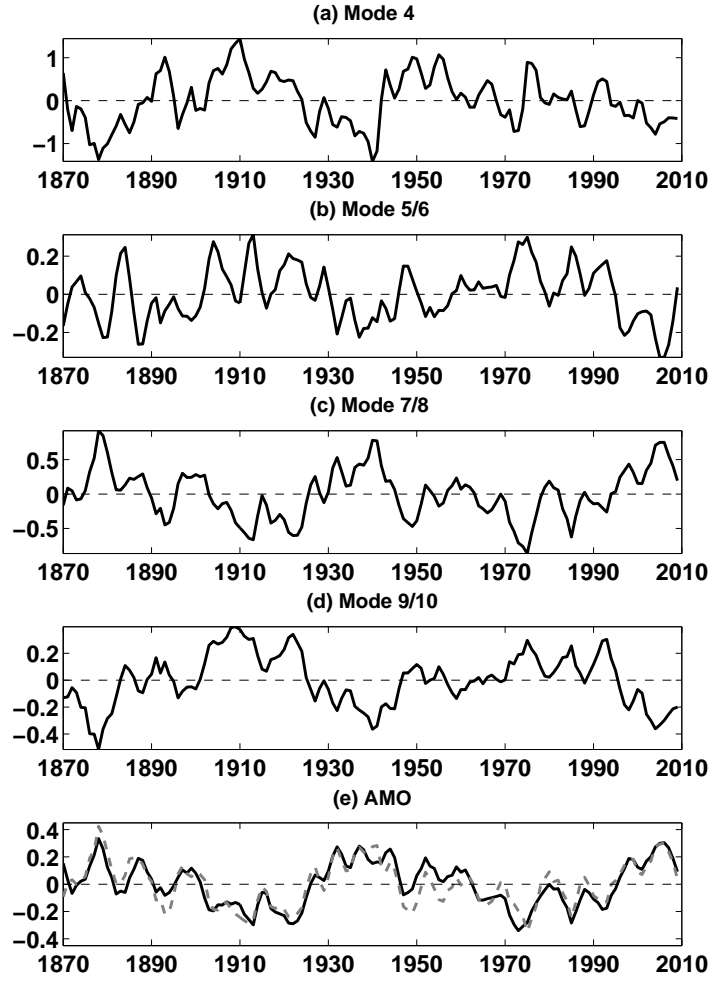


FIG. 7. Projection timeseries,  $\phi_j(t)$ , of eigenmodes (a) 4, (b) 5/6, (c) 7/8 and (d) 9/10; Timeseries of the AMO index (black) and the reconstructed AMO index (dashed grey) from the four eigenmodes (a) through (d).