Spectral characteristics and predictability of the NAO assessed through Singular Spectral Analysis

S. R. Gámiz-Fortis,[1] D. Pozo-Vázquez,[2] M. J. Esteban-Parra,[1] and Y. Castro-Díez[1]

[1] For the period 1826–2000, we analyze the spectral characteristics of the winter North Atlantic Oscillation (NAO) index and its predictability based on Singular Spectral Analysis (SSA) and Autoregressive Moving Average (ARMA) models. In the first part, SSA is applied to the winter NAO index to isolate its main spectral characteristics. Based on the SSA, a reconstruction (filtering) of the winter NAO index series was carried out. Results of the SSA indicate that the winter NAO index can be broken down into some modulated amplitude oscillations with periods around 7.7 and 4.8 years, some oscillations associated with a broadband peak of period around 2.4 years along with nonlinear trends. The sum of these components, the SSA-filtered series, explains 56% of the variance of the raw winter NAO index. The SSA-filtered series is particularly reliable, reproducing the NAO phase during extreme events (winter NAO index ≥ 1 or ≤ -1); for this subset of events, the phase of the actual and SSA-filtered series shows to be the same in 91% of the cases. The high positive values observed in the winter NAO index in the last eighties and nineties appear to be associated with the simultaneous presence of a positive trend, starting in the early eighties and of unprecedented steepness, and an oscillation with period around 7.7 years, having very high amplitude. In the second part, an ARMA model has been fitted to the filtered winter NAO index and a forecasting experiment was conducted; results are tested against the raw winter NAO index. Results show that the ARMA modeling has useful 1-year-ahead forecasting abilities. Particularly, over the period 1986–2000, not used to fit the model, the model skill is 27.8% better than climatology and 43.3% better than persistence (38.5% and 47.6%, respectively, when taking into account only extreme NAO events). Additionally, percentage of cases in which the NAO phase was accurately predicted proved to be 80% (88% for extreme NAO events). For 2001/2002 and 2002/2003 winters, persistence in the negative phase of the NAO is predicted, having an index value close to -1.

1. Introduction

[2] Among the several modes of low-frequency variability in geopotential heights in the Northern Hemisphere, one of the most important is known as the North Atlantic Oscillation (NAO) [*Wallace and Gutzler*, 1981; *Barnston and Livezey*, 1987]. The NAO mode of variability is particularly important during winter and is characterized by a dipolar pattern of NS sea level pressure (SLP), with one of the centers located over Iceland and the other approximately over the Azores Islands. This dipolar pattern reflects the strong contrast in meridional pressure over the North Atlantic region. The NAO presents remarkable interannual variability that can be summarized as the existence of two phases. The positive phase of the NAO reflects below-normal heights and pressure across the high latitudes of the North Atlantic, with above-normal heights and pressure over the central North Atlantic. The negative phase is characterized by opposed anomalies to those that are observed during the positive phase. Both phases of the NAO are associated with basin-wide changes in the intensity and location of the North Atlantic jet stream, storm track and changes in the patterns of zonal and meridional heat and moisture transport from the Atlantic Ocean to the continental areas of Europe. Particularly, a northward shift of the axis of maximum moisture transport occurs when the NAO is in the positive phase, causing an intensified westerly flow that brings warm maritime air to Europe during winter, reducing the polar outbreaks over Europe and leading to a warming of central and southern Europe and a cooling of the northwestern Atlantic area [*Hurrell*, 1995]. These effects make the NAO the most important cause of climate variability in the North Atlantic area on interannual timescales [*Hurrell*, 1995; *Hurrell and van Loon*, 1997;

---

[1]Department of Applied Physics, University of Granada, Granada, Spain.

[2]Department of Physics, University of Jaén, Jaén, Spain.

Osborn et al., 1999; *Rodrigo et al.*, 2001; *Pozo-Vázquez et al.*, 2001a].

[3] A proxy way to monitor the NAO and to study its temporal variability back in time is by means of an index constructed using the pressure differences between a station located near the southern center of the dipole and the other one near the northern center. In recent years, several NAO proxy indices have been proposed using different stations and different time-averaged intervals of the year [*Hurrell and van Loon*, 1997; *Jones et al.*, 1997].

[4] The observed temporal variability of the NAO during the recent decades shows some persistent episodes, which have yielded important climate-related socioeconomic impact over the NAO influence area [*Fromentin and Planque*, 1996]. For instance, from the 1979/1980 winter to the 1994/1995 winter, the positive phase of the oscillation dominated the circulation. This persistence of the positive phase produced anomalous climatic conditions over Europe. During this period southern Europe and Africa were abnormally dry, resulting in a severe drought, while northern Europe and Scandinavia were abnormally wet and warm [*Halpert and Bell*, 1997]. The situation reversed in the 1995/1996 winter, in which the NAO changed to the negative phase.

[5] Despite the intense research during the last decades, the physical mechanism underlying the temporal variability of the NAO still remains unclear. The spectral analysis of the NAO indices, both using classical and wavelet analysis, provides a near-white noise like behavior, but also some significant power associated with periods in the 2-year band, between 5 and 6 years and at quasi-decadal scales [*Rogers*, 1984; *Pozo-Vázquez et al.*, 2000, 2001a; *Stephenson et al.*, 2000]. These low-frequency variability modes have been associated with processes involving the North Atlantic Ocean basin [*Delworth*, 1996; *Taylor and Stephens*, 1998; *Rajagopalan et al.*, 1998; *Timmermann et al.*, 1998; *Rodwell et al.*, 1999], and also with internally generated atmospheric processes [*Perlwitz and Graf*, 1995]. Based on the association between the sea surface temperature (SST) in the North Atlantic basin and the NAO, some attempts have been made to predict the NAO state on seasonal to interannual basis [*Taylor and Stephens*, 1998; *Rodwell et al.*, 1999]. However, the importance of the relationship between the NAO and the Atlantic Ocean SST is still largely unknown and, unlike the El Niño–Southern Oscillation (ENSO), forecasting of the NAO state on seasonal to interannual scale upon the North Atlantic SST state is still far from reliable. Recent works suggest that the ENSO influences the NAO state [*Dong et al.*, 2000; *Pozo-Vázquez et al.*, 2001b] to a degree that could lead to an ENSO-related NAO variability.

[6] The aim of the present study is to increase current knowledge of the temporal modes of variability of the NAO and to evaluate its potential predictability based on its own history. The study is divided in two parts. In the first part, we use Singular Spectral Analysis (SSA) to determine and isolate the significant temporal modes of variability of the winter NAO index. SSA acts as a data-adaptive filter, removing the background noise and retaining the leading statistically significant signals [*Ghil and Vautard*, 1991; *Vautard et al.*, 1992]. The filtered signal is composed by modulated oscillatory signals and trends. Particularly, we

are interested in the presence and onset time of trends in the winter NAO index series. In the second part, an interannual linear prediction of the SSA-filtered winter NAO index series is carried out. To this end, Autoregressive Moving Average (ARMA) models [*Box and Jenkins*, 1976] are fitted to the series. ARMA models can be regarded as a special case of general linear stochastic processes and provide a linear representative structure of the temporal evolution of the data. We assume that the SSA-filtered winter NAO index contains, basically, the linearly predictable signal contained in the raw winter NAO index. Our purpose is not to supply an operational methodology to forecast the NAO state, but to study the extend to which the NAO state can be linearly predicted based in is own history. Particularly, since this methodology only uses for the prediction of the NAO the own history of the series, our results may be useful to evaluate the importance of external variables in predicting the NAO state.

[7] The work is organized as follows: section 2 describes the data used, section 3 shows the results of the analysis and a discussion and some conclusions are provided in section 4. In the appendices, a description of the methodologies used in this work is given. Particularly, SSA (Appendix A) and ARMA modeling (Appendix B.1) are fully discussed, as well as the methodology used to cross-validate the ARMA model (Appendix B.2), the setup procedure of the forecasting experiments (Appendix B.3) and the accuracy and skill scores employed (Appendix C).

2. Data

[8] The pressure data used correspond to Gibraltar (36.1°N, 5.4°W) and SW Iceland, the latter comprised mainly of data from Reykjavik and Stykkisholmur (65.0°N, 22.8°W) [*Jones et al.*, 1997]. The records, taken on a monthly basis, extend from 1826 to 2000. The final purpose is to establish a winter NAO index comparing the pressure data of two stations. Due to the different statistical characteristics of the data from the northern and southern stations, and due to the change of both mean and standard deviations over the year, a normalization process for each pressure time series is necessary in order to avoid creating a biased and misleading index. Thus, a monthly normalized series is constructed for each station, which consists of calculating the difference between each raw monthly value and a time-averaged monthly mean value, and then dividing by a time-averaged monthly standard deviation. Normalization relative to the period 1961–1990 has been used. Notably, this process removes the annual deterministic cycle. We construct a monthly NAO index by subtracting both series, and a seasonal index for winter is constructed using averages of three consecutive monthly index values (December–February). Although other normalization procedures have been proposed [*Ropelewski and Jones*, 1987], this has been argued to be the optimal method [*Trenberth*, 1984]. Due to the selected normalization period, the index does not have mean zero or standard deviation unity; the variance of the series is 1.23. Although the existence of several formulation for the winter NAO index, the index constructed using Gibraltar data for the southern center has been found to be the most representative of the NAO variability during winter [*Pozo-Vázquez et al.*, 2000]. An annual index of the

NAO, using the Gibraltar as the southern stations, has been also analyzed. The annual index was computed by averaging the seasonal indices (March–May for spring, June–August for summer, and September–November for the autumn index).

[9] To test the sensitivity of the results against the way in which the index is formulated, we have also analyzed other formulations of the winter NAO index, which used SLP data from Azores and Lisbon as representatives of the southern NAO center of action. These indices were computed using the same procedure used for the Gibraltar case, and cover the period 1865–2000. The Principal Component (PC) series of the Northern Hemisphere SLP during winter, as defined by *Barnston and Livezey* [1987], is also used. This series covers the period 1950–2000.

[10] Since the main application in predicting the NAO involves predicting its impact on the climate, and given that this impact is strongly dependent on the NAO strength and phase, we have investigated the performance of the proposed SSA and ARMA models in reproducing and predicting the extreme phases of the NAO, defining those as the cases in which the raw winter NAO index is ≥ 1 or ≤ -1. Results using this subset of data are also of interest given that the differences between the different winter NAO index formulations diminishes when considering only the extreme events. Thus, the results presented for these particular cases can be regarded somehow index independent.

3. Analysis
3.1. SSA of the Winter NAO Index

[11] In this section we carry out a SSA on the raw winter NAO index. Procedures followed in this analysis are fully discussed in Appendix A.
3.1.1. Oscillatory Modes

[12] SSA has been carried out on the lagged-covariance matrix based on the Broomhead–King algorithm, using a M = 40 window length. In the results of this analysis, we have considered 15 eigenvalues, which explain 60% of the total data variance. Eigenvalues 1 and 2, 3 and 5, 6 and 7, 8 and 9, and 11 and 12 conform 5 pairs, each one in quadrature, representing oscillatory modes. On the other hand, eigenvalues 4 and 10 reflect the general trends of the data, and particularly the Empirical Orthogonal Function (EOF) corresponding to the eigenvalue 4 (EOF 4), is associated with an oscillation of frequency 0.008 cycles/yr (period 125 years). Finally, eigenvalues number 13, 14, and 15 do not show oscillatory behavior.

[13] The Maximum Entropy Method (MEM) has been used to evaluate the spectral contents of the PC time series corresponding to the EOFs. Results reveal that EOFs pairs 1–2 and 6–7 contain oscillations associated with periods around 7.8 and 4.8 years, respectively. EOFs 3–5, 8–9, and 11–12 have associated oscillations of periods 2.4, 2.3, and 2.9 years, respectively.

[14] The use of a window length of M = 40 years in the SSA yields to an approximate spectral resolution of 1/40 years = 0.025 cycles/yr. Given that the difference between periods of 2.4 and 2.3 years is only 0.018 cycles/yr, it is more likely that the individual oscillations of periods between 2.4 and 2.9 years found in the SSA analysis represent a broadband peak, rather than distinct periodici-

ties. Regardless of the choice of the window length (M), the oscillations of period 2.3 and 2.4 years are very close to resolvable frequency using yearly sampled data (Nyquist frequency) [*Robertson et al.*, 2000]. We have used monthly and annual data to confirm the robustness of oscillation modes found in the winter-annual SSA. Results both using monthly and annual data show the presence of a single oscillation with period around 2.6 years. This confirms the hypothesis that the oscillations of periodicities between 2.3 and 2.9 years found for the winter NAO index represent in a fact a broadband peak.
3.1.2. Testing the Significance

[15] Results of the previous sections suggest the existence of oscillatory components in the winter NAO index. The extent to which this hypothesis can be considered true is assessed in this section through the use of a Monte Carlo technique. First, results of the previous section will be tested against the hypothesis of the winter NAO index to be the result of a pure AR(1) process with a lag-one autocorrelation value corresponding to that of the winter NAO index (0.05). To this end, we use the data-adaptive basis, projecting each surrogate Monte Carlo realization onto the EOFs of the data and comparing the result with those of the original data. Figure 1a shows the results of this hypothesis test. The extremes of the vertical bars indicate the 2.5th and 97.5th percentiles of the diagonal elements of $\Lambda^{\text{surrogate}}$ corresponding to the EOFs whose eigenvalues they overlie. The vertical bars show the variance expected from the AR(1) process. Results show that EOFs 1, 2, 3, and 5 contain more power than expected from the null hypothesis, indicating that these EOFs are individually significant at the 97.5% level. On the other hand, the rest of the EOFs, explain just the variance expected from the AR(1) process, lying within the surrogate data bars.

[16] An alternative way of displaying the information of Figure 1a is to plot the eigenvalues and surrogate data bars against the dominant frequency associated with their corresponding EOFs, as shown in Figure 1b. Since the EOFs obtained from SSA are not generally pure sinusoidal, identifying a single frequency with an EOF is difficult. For display purposes, a frequency can be associated with an EOF simply by maximizing the squared correlation with a sinusoidal signal. Figure 1b shows EOFs 1 and 2 to be significant, forming a pair centered on 0.13 cycles/yr (period of 7.7 years). EOFs 3 and 5, centered on 0.42 cycles/yr (period 2.4 years), also prove to be significant at the 97.5% confidence level.

[17] Merely observing a pair of data eigenvalues lying above the 97.5th percentiles of the corresponding surrogate distributions is generally not enough to conclude that we have detected an oscillation at that frequency at the 97.5% significance level. Even if we are analyzing a segment of pure noise, the average number of excursions above the 97.5% percentile will be 0.025M by construction. The correct interpretation of Figure 1a, therefore depends on our prior knowledge and expectations. If we know beforehand that it is EOF-k that we are interested in, then the position of the kth eigenvalue of $\mathbf{S}^{\text{data}}$ relative to the corresponding surrogate data distribution translate straightforwardly into the significance level of the test. Often, however, we use the results of such spectra to decide which EOFs to focus on [*Allen et al.*, 1996].

a)

b)
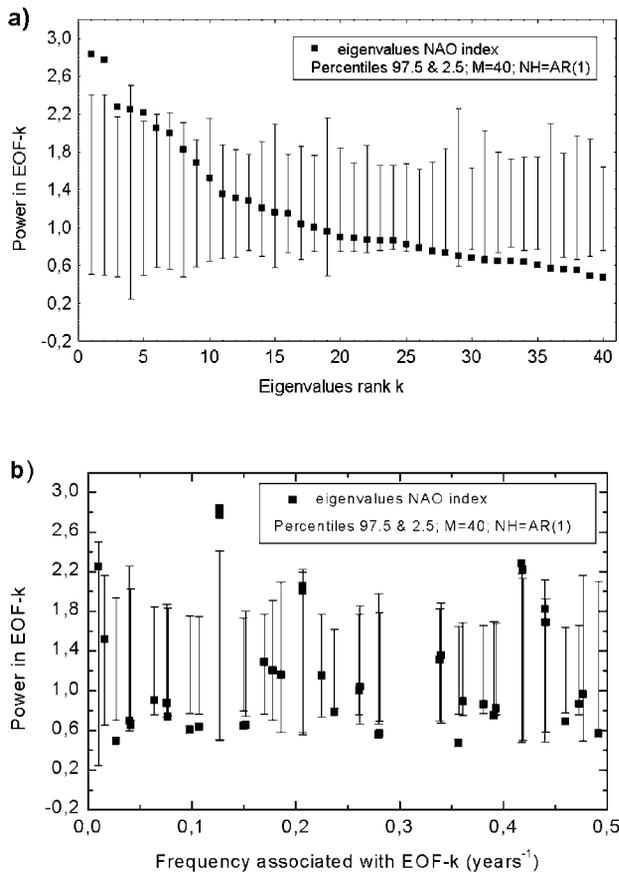
Figure 1. (a) Eigenvalues resulting from the SSA of the winter NAO index ranked by associated power. Vertical bars show the variance we should expect in the directions defined by these EOFs in a segment of AR(1) noise (percentiles 97.5th and 2.5th). (b) Eigenvalues resulting from the SSA of the winter NAO index against the dominant frequency associated with their corresponding EOFs. Vertical bars show the variance we should expect in the directions defined by these EOFs in a segment of AR(1) noise (percentiles 97.5th and 2.5th).

[18] The Monte Carlo procedure can also estimate the probability of n excursions above the mth percentile of the test presented in Figure 1a. If we look for any excursions above the 97.5th percentiles of the surrogate distributions with a window width of 40, then we are, in fact, performing 40 "minitests." The probability of at least one excursion above the 97.5th percentiles of the surrogate distributions is clearly greater than 2.5%. Since the 40 "minitests" are not mutually independent, the probability of a given number of excursions does not, in general, conform to an analytically calculable distribution. For the AR(1) null hypothesis, this is well approximated by the binomial distribution that we would expect if the excursions are independent, so it can be parameterized [*Allen et al.*, 1996]. For the AR(1) process, the probability of having at least four excursions above the 97.5th percentiles, as observed in Figure 1a, for a given member of the surrogate ensemble is 8%. Therefore, the true confidence level at which the AR(1) null hypothesis can be rejected is around 92%. However, taking into account the existence of oscillations having periods 7.7

and 2.4 years, obtained from the spectral analysis, and given our a priori interest in EOFs 1, 2, 3, and 5, we reject the null hypothesis of the winter NAO index to be a red noise AR(1) process.

[19] The test procedure must continue until a final null hypothesis is not rejected. The following step is to test the composite null hypothesis consisting of a red noise whose parameters are chosen in order to exclude the signal EOFs identified (1−2 and 3−5 EOFs pairs). For this hypothesis the lag one autocorrelation value is 0.024. This value corresponds to the residual time series resulting from the reconstruction of the winter NAO index just using 1−2 and 3−5 EOFs pairs. Figures 2a and 2b show the results of this new hypothesis test. As expected, again, EOFs 1, 2, 3, and 5, previously identified signal, are trivially significant at 97.5% confidence level. Additionally, EOFs 4, 6, and 7, this latter being just at the limit, were significant at this significance level. The EOF 4 has an associated oscillation of a period of 100 years, thus representing the present trend in the data series. EOFs 6 and 7 form a pair with a similar associated frequency of 0.21 cycles/yr (period 4.8 years). As in the previous case, the true confidence level to which this hypothesis can be rejected, taking into account the probability of having at least three excursions above the 97.5th percentiles, is the 79% level. On the base of these results and the a priori information, the composite null hypothesis can be rejected.

a)

b)
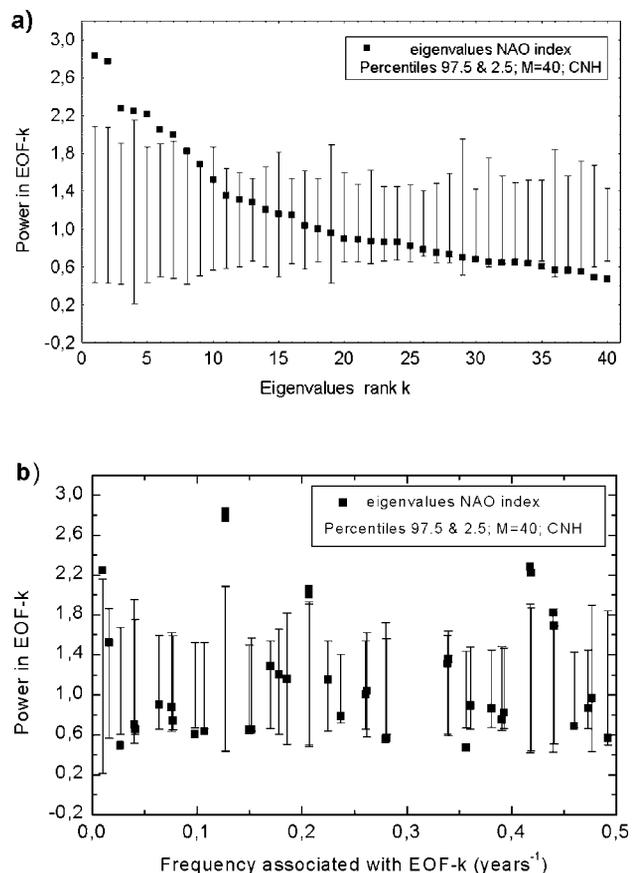
Figure 2. Test series against the null hypothesis of winter NAO index consists of two oscillations of 7.7 and 2.4 years plus AR(1) noise. (a) As in Figure 1a. (b) As in Figure 1b.

a)

b)
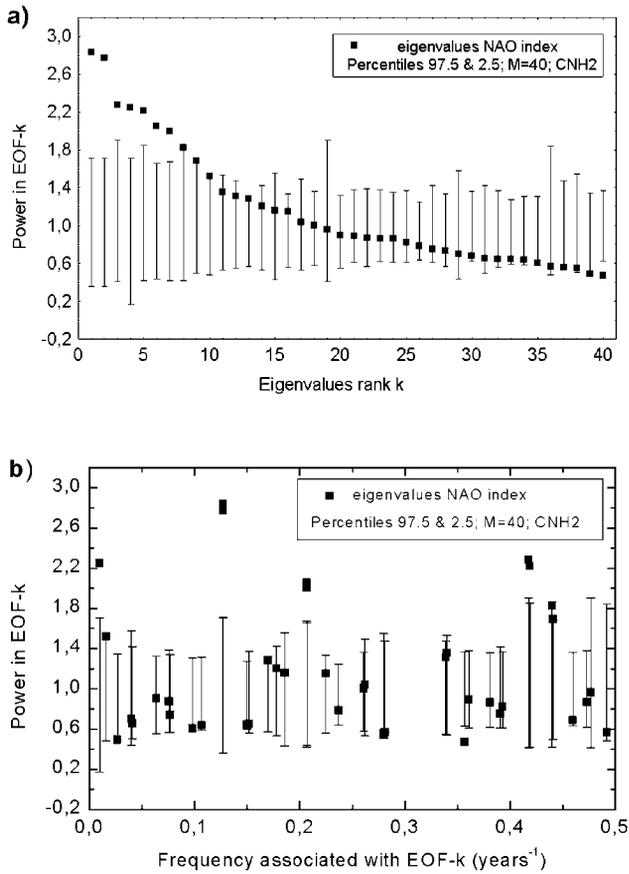
Figure 3. Test series against the null hypothesis of winter NAO index consists of low-frequency variability (period > 100 years), three oscillations of 7.7, 2.4, and 4.8 years plus AR(1) noise. (a) As in Figure 1a. (b) As in Figure 1b.

[20] The null hypothesis currently consists of a red noise, being the EOF 4 (representing the trend), EOFs 1−2 (oscillation of period 7.7 years), EOFs 3−5 (oscillation of period 2.4 years), and EOFs 6−7 (oscillation of period 4.8 years), chosen as signal. Under this conditions, the AR(1) process has a lag-one autocorrelation value of −0.039. This value corresponds to the residual time series derived from the reconstruction of the winter NAO index using 1−2, 3−5, and 6−7 EOFs pairs and the EOF 4. Figures 3a and 3b show the results of this new hypothesis test. Results show that, in addition to the EOFs 1−2, 3−5, and 6−7, also EOFs 9 and 10 are now significant at the 97.5% level. EOF 10 has an associated frequency of 0.016 cycles/yr (period 63 years). EOF 9 and EOF 8 form a pair with an associated period of 2.3 years, being EOF 8 almost significant (significance level is 97%). The true confidence level at which this hypothesis can be rejected is 79%. Based on these results, the second composite null hypothesis is rejected.

[21] Finally, the null hypothesis for the winter NAO index to be composed of low-frequency variability (periods greater than 63 years, EOFs 4 and 10), an oscillation of period 7.7 years (EOFs 1−2), some oscillations associated with a broadband peak or period around 2.4 years (contained in EOFs 3−5 and 8−9), an oscillation of period 4.8

years (EOFs 6−7) and a AR(1) process with a lag-one autocorrelation value of −0.0039 will be tested. Figure 4 shows the results. An additional EOF (EOF 12) proves to be significant, but the true confidence level is very low, only 68%. On the basis of the former results, this null hypothesis cannot be rejected.

[22] We thus conclude that the winter NAO index can be represented by the following model, a nonlinear trend that contains variability at periods of 63 (EOF 10) and 100 years (EOF 4), an oscillation with associated period of 7.7 (EOFs 1−2), some oscillations associated with a broadband peak of period around 2.4 years (contained in EOFs 3−5 and 8−9), an oscillation of period 4.8 years (EOFs 6−7), and a red noise process with lag-one autocorrelation −0.0039 and variance 0.84. This model was checked using the Noise Null Hypothesis test [*Allen et al.*, 1996], obtaining similar results.

[23] We have explored the sensitivity of the results to the way in which the winter NAO index is defined. For this purpose, we have carried out a SSA on the winter NAO indices defined using the Azores and Lisbon as the southern stations, and also on PC SLP series defined by *Barnston and Livezey* [1987], although its limited temporal extension. Overall, results show that the most important modes of variability found in the Gibraltar formulation of the winter NAO index are also found when the other formulations are analyzed. Particularly, the long-term trend, the oscillations of period around 7.7 and the broadband peak of period around 2.4 years are present in the four analyzed NAO indices. Main differences are the oscillation of period around 4.8 years, found in the Gibraltar formulation but not in the rest, and one oscillation of period between 3.3 and 3.8 just found for the Azores formulation.

[24] To test the results against the danger of aliasing subannual frequencies into annual modes, we have carried out an additional SSA using an annual Gibraltar NAO index. Oscillations found for the winter index are also found in the annual index, while some oscillations are found in the annual index but not in the winter case (5, 5.9, 13, and 4 years). Thus, the modes of variability found
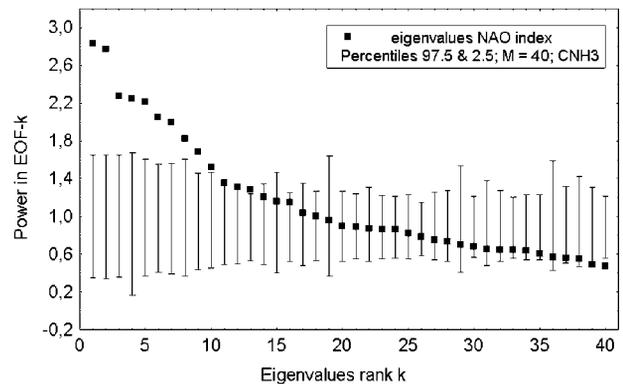
Figure 4. Test series against the null hypothesis of winter NAO index consists of low-frequency variability (period > 63 years), four oscillations of 7.7, 2.4, 4.8, and 2.3 years plus AR(1) noise. (a) As in Figure 1a.

Table 1. Statistical Results for the NAO SSA Modeling and the ARMA Forecasting Experiment

	Raw winter NAO versus SSA-filtered winter NAO (period 1826–2000)	Raw winter NAO versus one-step-ahead forecast (period 1826–1985)	Raw winter NAO versus one-step-ahead forecast (period 1986–2000)	Raw winter NAO versus several-steps-ahead forecast (period 1986–2000)
MSE	0.72	0.75	1.49	1.8
MAE	0.69	0.71	0.97	1.1
Correlation coeff.	0.75[a]	0.70[a]	0.48	0.40
MSE_{cli}		1.15	1.98	1.98
MSE_{per}		2.23	2.64	2.64
% $SMSE_{cli}$		34.8	24.7	9
% $SMSE_{per}$		66.3	43.5	31.8
% Phase accordance	74	69	80	60

[a] Correlation coefficients that are statistically significant at the 95%.

for Gibraltar formulation of the winter NAO index can be considered interannual modes of variability.

3.1.3. Reconstruction of the Winter NAO Index

[25] In the previous section, we concluded that the winter NAO index can be represented by a nonlinear trend that contains variability at periods of 63 and 100 years and a set of oscillations with associated periods around 7.7, between 2.3–2.4 and 4.8 years, plus a red noise process. Based on these results, we have carried out a reconstruction of the winter NAO index (see Appendix A), called SSA-filtered series. The raw index has a variance value of 1.23, while the SSA-filtered NAO has a variance of 0.51. Over the period 1826–2000, the correlation between the original and the SSA-filtered series is 0.75 (statistically significant at 95%). Thus, the variance explained by the model is 56%. Additionally, Maximum Square Error (MSE) of the model is 0.72, Maximum Absolute Error (MAE) is 0.69 and the percentage of cases in which the phase of the SSA-filtered series is the same as that of the raw winter NAO index is 74% (see first column in Table 1). Figure 5a shows the raw and SSA-filtered indices along with the trend component. The model is particularly reliable capturing the winter NAO index behavior in the period 1826–1900, during which extreme values are in many cases correctly modeled, and from 1970 onwards, during which the tendency in the winter NAO index is clearly reflected. During the period 1826–1890 no trend is discerned. From 1890 to 1920 a slightly positive trend can be observed while from 1920 to 1970 a negative trend appears in the data. From 1980 onwards a strong positive trend is evident. Note the unprecedented steepness of this trend. Figures 5b and 5c show, respectively, the oscillations associated with periods 2.4 and 2.3 years. There is a clear modulation of the amplitude of these components, with higher amplitude in the periods 1826–1850 and 1890–1930, and the lowest amplitude at the end of the record. The same amplitude-modulated behavior is found associated with periods 7.7 and 4.8 years (Figure 5d). The high amplitude associated with period 7.7 years at the final of the record is notable. Many of the high positive values observed in the winter NAO index in the eighties and nineties (particularly during 1983–1984, 1989–1992, and 1998–2000) appears to be associated with the simultaneous presence of a positive trend, starting in the early eighties and of unprecedented steepness, and an oscillation with period around 7.7 years with very high amplitude.

[26] For the extreme phases of the NAO, the MSE is 1.20, MAE is 0.95, the correlation between the original and the

SSA-filtered series is 0.83 (statistically significant at 95%), and the percentage of cases in which the phase of SSA-filtered winter NAO index is the same as that of the raw index is 91% (see first column in Table 4).

3.2. Stochastic Modeling and Forecasting of the Winter NAO Index

[27] In this section we carry out an ARMA modeling of the SSA-filtered winter NAO index. The ARMA model fitted to the SSA-filtered winter NAO index is then used in a forecasting experiment.

[28] Data from 1826 to 1985 were used to fit the model while data from 1986 to 2000 were used for a true comparison in the forecasting experiment. See the appendices for a throughout explanation of the ARMA modeling procedure (Appendix B.1), the methodology used for cross-validation the ARMA model (Appendix B.2), the setup procedure of the forecasting experiments (Appendix B.3) and the accuracy and skill scores employed (Appendix C).

3.2.1. ARMA Modeling

[29] Figures 6a and 6b show, respectively, the sample Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the raw winter NAO index series, while Figures 6c and 6d show the sample ACF and PACF of the SSA-filtered winter NAO index. The analysis of these figures is the starting point of the ARMA modeling process.

[30] As a preliminary step, we studied the stationarity in mean and variance of the raw and the SSA-filtered NAO indices. The study was carried out by analyzing the ACF and following *Pankratz* [1991, chapter 2]. As a result of this study, the stationarity in the mean and variance for the two analyzed series can be assumed. A study of the histogram, skewness and kurtosis coefficients suggests that also the normality can reasonably be assumed for both series.

[31] Note in Figure 6 that, while the raw data behaves like a white noise process, the SSA-filtered series shows a strong autocorrelation pattern. This makes the SSA-filtered winter NAO index more predictable compared to unfiltered one, given the weak autocorrelation pattern of this latter. It appears that the autocorrelation pattern of the winter NAO index was hidden, in some sense, by other random phenomena that were present in the original pressure data, and that the filtering was successful in removing these other phenomena.

[32] Following the ARMA modeling process outlined in Appendix B.1, we have found for the raw winter index an ARMA(9,9) model. The estimated innovations variance is $\hat{\sigma}_e^2 = 0.97$, around 21% of reduction in the variance from

a)

b)

c)

d)
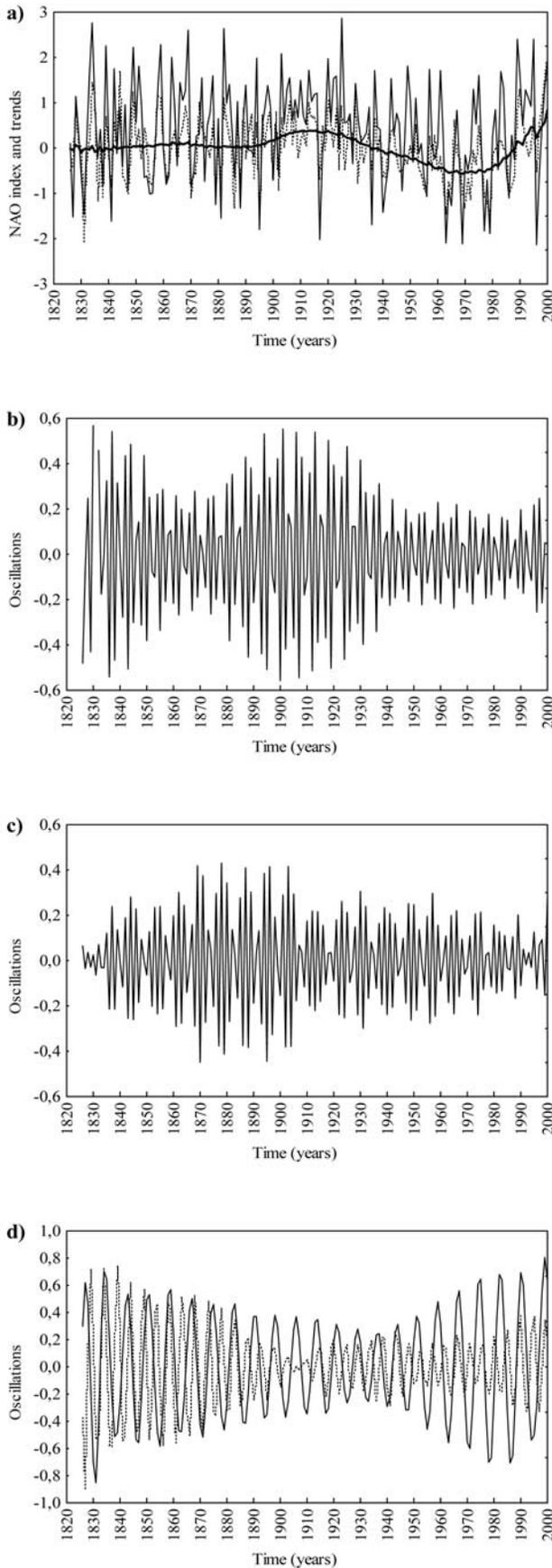
that of an uncorrelated process (variance of raw winter NAO index is 1.23).

[33] For the SSA-filtered index, we have finally chosen an ARMA(8,9) model with parameters:

$$AR = (\phi_1 = 0.37^*, \ \phi_2 = 0.43^*, \ \phi_3 = -0.97^*, \ \phi_4 = 0^*,$$
$$\phi_5 = 0.95^*, \ \phi_6 = -0.66^*, \ \phi_7 = 0, \ \phi_8 = 0.66^*),$$

$$MA(\theta_1 = 0.26^*, \ \theta_2 = -0.22, \ \theta_3 = 0, \ \theta_4 = 0, \ \theta_5 = -0.68^*,$$
$$\theta_6 = 0, \ \theta_7 = -0.19, \ \theta_8 = 0, \ \theta_9 = -0.56^*).$$

[34] The estimated innovations variance is $\hat{\sigma}_e^2 = 0.03$, around 94% reduction in the variance from that of an uncorrelated process (SSA-filtered NAO variance is 0.51). Since the SSA-filtered winter NAO index accounts for the 56% of the raw winter NAO index variance, the ARMA model explains a similar amount of this raw winter NAO index. Significance of the parameters was computed using approximate t-values, derived from the parameter standard errors (those with "*" are significant at 95% level). If not significant, a parameter can in most cases be dropped from the model without substantially affecting the overall fit, but improving its parsimony (monitored through the Akaike Information Criterion (AIC)). Thus, some parameters have been set to zero. Some parameters were not significant, but a comprehensive study showed their importance (for instance, the AIC was not improved when setting these parameters to zero). The high-order model finally selected, through the use of an independent information criterion, indicates a certain degree of persistence. Given the values of some of the parameters of the model and the ACF and PACF shown in Figures 6c and 6d, the current winter NAO state is considerably dependent on its state up to 5, 6, and even 8 years earlier. Also notable is the importance of the MA part of the model, that is, the departure of the expected past values (up to 9 years before) substantially influences the current value.

3.2.2. Forecasting Experiment

[35] We have conducted a forecasting experiments for the SSA-filtered winter NAO index based on its ARMA(8,9) model. Results are tested against the raw winter NAO index. Procedures followed in this forecasting experiment are fully discussed in the appendices.

[36] Figure 7 presents the result of this forecasting experiment and Table 1 shows a summary of the performance of the model. Particularly, Figure 7a shows the one-step-ahead forecast, that is, the prediction made for 1 year into the future, during the period 1965–1985, while Figure 7b shows the forecasting for the period 1986–2000. The period 1826–1985 was used to construct the model (calibration period) and therefore, only the predictions from 1986 to 2000 (Figure 7b) can be regarded as genuine. For the sake of comparison, the raw winter NAO index is also shown. Over the calibration period, Figure 7a, the quasi-decadal oscillation, as the main characteristic of the series, is clearly

Figure 5. (opposite) (a) Raw (thin continuous line) and SSA-filtered (dashed line) series of the winter NAO index. The trend components are also shown (thick continuous line). (b) Oscillation of period 2.4 years. (c) Oscillation of period 2.3 years. (d) Oscillations of period 7.7 years (continuous line) and 4.8 years (dashed line).
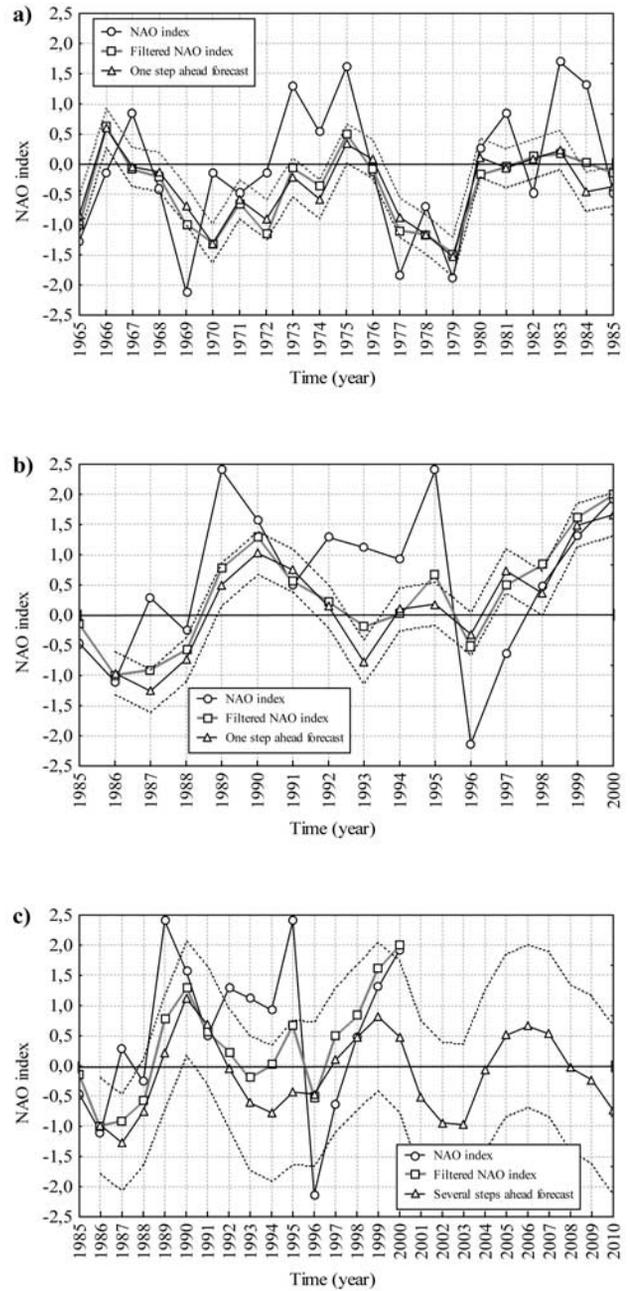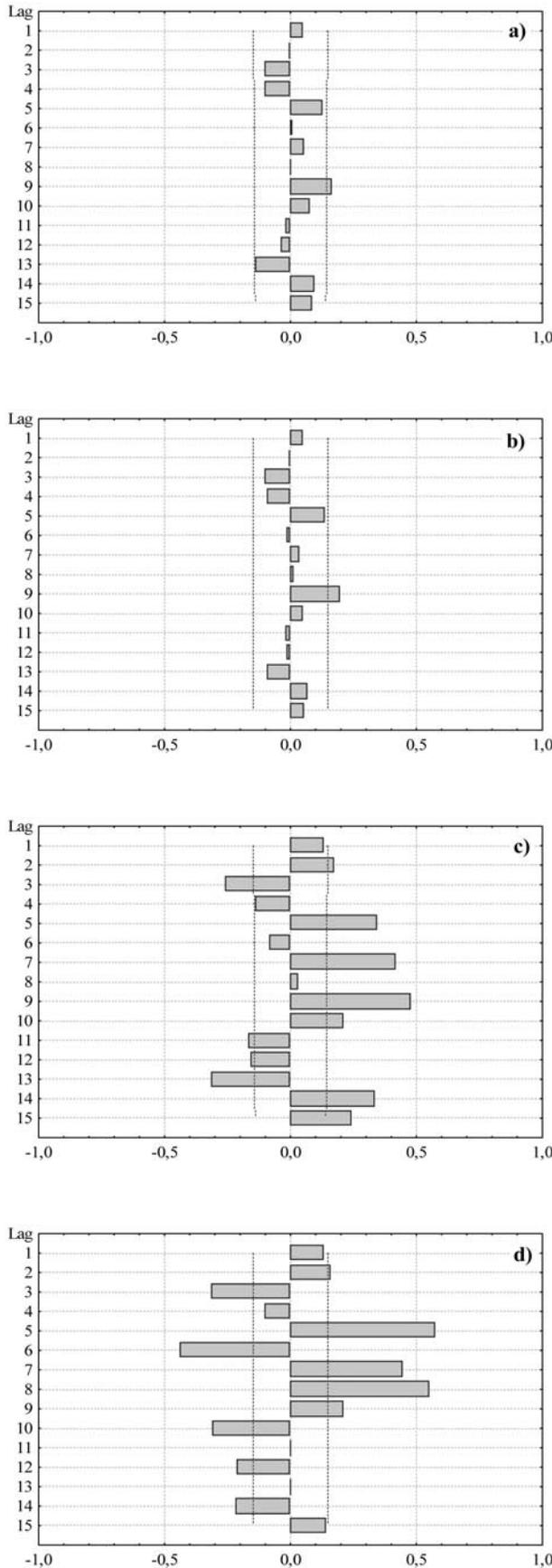
Figure 7. (a) Results of the forecasting experiment of winter NAO index. The one-step-ahead forecast for the SSA-filtered winter NAO index is shown for the period 1965–1985. Raw (unfiltered) NAO data, filtered NAO data, and one-step-ahead forecasts along with the 95% confidence intervals are displayed. (b) As in (a), but during the period 1986–2000 (used for true comparison). (c) Results from the several-steps-ahead forecasting experiment. Using data until 1985, NAO is forecasted starting in 1986 and continuing until 2010.

Figure 6. (opposite) Estimated ACF (a) and PACF (b) along with 95% confidence intervals of raw winter NAO index. Standard errors assume AR order of k - 1 for PACF and white-noise estimates for ACF. (c and d) As in (a) and (b), but for the SSA-filtered winter NAO index.

visible. The actual filtered winter NAO index values are always within the 95% confidence intervals, indicating that the ARMA model properly represents the main characteristics of the SSA-filtered winter NAO index. The raw winter NAO index shows a considerable variability, greater than those of the filtered index, causing that the raw values to fall outside of the confidence bands in many cases. Nevertheless, the overall behavior of the raw winter NAO index closely recalls that of the filtered one, and the model forecasts with generally little error the state of the NAO. Particularly, MSE is 0.75, MAE is 0.71, correlation coefficient is 0.70 while percentage of cases in which the phase of the NAO was accurately predicted is 69%. This leads us to conclude that the autocorrelation structure of the winter NAO index in the SSA-filtered index, which was overwhelmed by the presence of random signals in the raw data, captures to certain degree the behavior of the raw winter NAO index over time. Over the period 1986–2000 (Figure 7b), do not used to fit the model, MSE is 1.49, MAE is 0.97, correlation coefficient is 0.48 while the percentage of cases in which the NAO phase was accurately predicted proved to be 80%. Reliability was particularly strong in the forecast during the period 1998–2000.

[37] To assess the skill of the forecast ARMA model, we first computed the MSE when climatology (MSE_{cl}) and persistence (MSE_{pe}) are used for forecasting; then, we have obtained the percentage improvement in MSE forecast over a climatological forecast ($SMSE_{cl}$) and the percentage improvement in MSE forecast over a persistence forecast $SMSE_{pe}$. Additionally, we computed the Linear Error in Probability Space skill score (LEPS), see Appendix B.3. Results are shown in Table 1. Over the period 1826–1985, the ARMA model skill is 34.8% better than climatology and 66.3% better than persistence. Over the period 1986–2000, percentage improvement is lower but still high, $SMSE_{cl}$ is 24.7% and $SMSE_{pe}$ 43.5%. Note the considerable better results in the prediction using climatology compared to persistence. Furthermore, over the period 1826–1985, the LEPS value is +36.8% while over the period 1986–2000 the LEPS showed a value +23.4%. We can, then, conclude that the ARMA model shows considerable skill, improving notably the results of using climatology or persistence or random forecast.

[38] An additional forecasting experiment was made in order to check the forecasting skill of the model in predictions several steps ahead, for which Figure 7c shows the results. Using information until 1985, the NAO state is forecasted from 1986 to 2010. It is easy to note the quasi-decadal structure of the forecast. Results of the forecast are especially reliable during the periods 1986–1991 and 1998–2000. MSE for the period 1986–2000 is 1.8, MAE is 1.1, correlation coefficient is 0.40 and percentage of accurately predicted phase is 60%. The skill scores show percentage improvements, as expected, considerably lower than in the one-step-ahead forecast, but still the ARMA model shows better results than climatology ($SMSE_{cl}$ is 9%) and persistence (MSE_{pe} is 31.8%). Confidence in the prediction seems to lose reliability when an anomalous persistence in one of the phases (as those of the early nineties) takes place. A problem to be considered is possible "external" factors not accounted for in the history of the series and which could affect the evolution of the phenomenon. For the winter of 2000/2001

the prediction of the model was −0.523 while the actual value was −0.44. For the winters of 2001/2002 and 2002/2003, persistence in the negative phase is predicted, having winter NAO index value close to −1.

[39] As in the case of the SSA-filtered winter NAO index, a forecasting experiment has been carried out for the raw winter NAO index using its ARMA(9,9) model. Results are considerably worse than those using the SSA-filtered index. Particularly, during the period 1986–2000, the one-step-ahead forecast has a MSE of 2.2, MAE is 1.6, correlation between the actual and the forecasted series is 0.36 and the percentage of accurate predicted phases is 62%.

3.2.3. Cross-Validation

[40] Cross-validation of the model was carried out by dividing the development period 1826–1985 in two subperiods, 1826–1899 and 1900–1969, and then fitting ARMA models to these two subseries (see Appendix B.2). For the period 1826–2000, we found an ARMA(8,9) model, having parameters:

$$AR = (\phi_1 = 0.27^*, \ \phi_2 = 0.21^*, \ \phi_3 = -0.95^*, \ \phi_4 = 0,$$
$$\phi_5 = 0.75^*, \ \phi_6 = -0.80^*, \ \phi_7 = 0, \ \phi_8 = 0.40^*),$$

$$MA(\theta_1 = 0.30^*, \ \theta_2 = -0.44^*, \ \theta_3 = 0, \ \theta_4 = 0, \ \theta_5 = -0.70^*,$$
$$\theta_6 = 0, \ \theta_7 = -0.37, \ \theta_8 = 0, \ \theta_9 = -0.14).$$

For the period 1900–1969, the model was an ARMA(8,9) having parameters:

$$AR = (\phi_1 = 0.27^*, \ \phi_2 = 0.34^*, \ \phi_3 = -0.83^*, \ \phi_4 = 0,$$
$$\phi_5 = 0.82^*, \ \phi_6 = -0.50^*, \ \phi_7 = 0.1, \ \phi_8 = 0.68^*),$$

$$MA(\theta_1 = 0.14, \ \theta_2 = -0.27, \ \theta_3 = 0, \ \theta_4 = 0, \ \theta_5 = -0.45^*,$$
$$\theta_6 = 0, \ \theta_7 = -0.35, \ \theta_8 = 0, \ \theta_9 = -0.47^*).$$

[41] Estimated reduction in the variance from that of an uncorrelated process is 95% for the 1826–1899 period model and 94% for the 1900–1969 model. Note the strong similarities between these two models and the model found for the whole period. Only slightly differences in some coefficients can be appreciated, but the order of the model remains the same.

[42] Using the former ARMA models, we have carried out one-step-ahead forecast and several-steps-ahead forecast over the period 1900–1915 for the 1826–1899 model and over the period 1970–1985 for the 1900–1969 model. Results are shown, respectively, in Tables 2 and 3. The comparison of the results in Tables 1, 2, and 3 do not show important differences between the performances of the complete period ARMA model and those of the two subperiods, nor in MSE and MAE neither in the correlation and percent of phase accordance. The main difference is a greater MSE in the one-step-ahead forecast over the period 1985–2000 (corresponding to the 1826–1985 model) than the MSE in the periods 1900–1915 and 1970–1985 corresponding two subperiod models. In the first case is 1.49 while in the last two are 0.95 and 0.94.

[43] Previous results show than the ARMA model can reasonably be considered independent of the period of fitting, provide this period be long enough.

3.2.4. Extreme NAO Values

[44] As stated before, the main application in predicting the NAO state involves predicting its impact on the climate.

Table 2. Statistical Results for the Validation ARMA Model Developed Over the Period 1826–1899

	Raw winter NAO versus one-step-ahead forecast (period 1826–1899)	Raw winter NAO versus one-step-ahead forecast (period 1900–1915)	Raw winter NAO versus several-steps-ahead forecast (period 1900–1915)
MSE	0.80	0.95	1.06
MAE	0.72	0.89	0.91
Correlation coeff.	0.74[a]	0.74[a]	0.65[a]
% Phase accordance	74.4	73	62

[a] Correlation coefficients that are statistically significant at the 95%.

Given that this impact is strongly dependent on the NAO strength and phase, we have investigated the performance of the ARMA model in forecasting the extreme phases of the NAO. Furthermore, the results for these particular cases can be regarded somehow index independent. Table 4 shows the results. When comparing results in Tables 1 and 4, MSE and MAE show to be greater for the extreme cases, but results of the correlation coefficients and percentage of phase accordance are better. Particularly, the model is highly reliable predicting the phase: 82% during the period 1826–1985 and 88% during the period 1986–2000. During this latter period, only the year 1993 presented an extreme phase of the NAO not predicted 1 year earlier. In the several-steps-ahead forecasting experiment, a 66% of the phases were accurately predicted during the period 1986–2000 and the skill scores of these extreme cases also show an overall improvement compared to the whole cases. For the one-step-ahead forecast and over the period 1985–2000 the skill is 38.5% better than climatology (24.7% in Table 1). A similar improvement is found for the several-steps-ahead forecast, $SMSE_{cl}$ is 17.8% for extreme NAO values and 9% when considering all the winter NAO index values.

[45] Overall, ARMA model for extreme NAO values shows better performance than for the whole NAO values, highly improving the results of using simple climatology or persistence.

4. Discussion and Concluding Remarks

[46] The variability and predictability of the winter NAO index have been studied through the period 1826–2000 using SSA and ARMA models.

[47] Results of the SSA show that the winter NAO index can be represented by the following model: a nonlinear trend which contains variability at periods of 63 and 100 years, amplitude-modulated oscillations with associated periods around 7.7, 4.8, between 2.3 and 2.4 years and a red noise process with lag-one autocorrelation −0.0039 and variance 0.84. Using the former model, a reconstruction of the winter NAO index (called the SSA-filtered NAO index) was carried out, accounting for the 56% of its variance. The percentage of cases in which the actual phase of the winter NAO index was accurately reproduced by the SSA-filtered index is 74% and correlation between the these series is 0.75. For extreme events (NAO index ≥ 1 or ≤ -1) percentage reaches 91% and correlation 0.87. Similar modes of NAO temporal variability to those found in the SSA have been also reported in different studies of the spectral characteristics of the NAO index. *Rogers* [1984] analyzed the Fourier spectrum of the winter NAO index from 1900 to 1983, using pressure data from Iceland and the Azores, finding pronounced peaks at periods 5, 7, and 20 years. *Hurrell and van Loon* [1997] reported the Fourier power spectrum of winter NAO index from 1865 to 1997, identifying peaks at periods 2–3 and 6–10 years and interdecadal bands. They also found a trend of the NAO spectra to become redder with time. Perhaps the most prominent feature of the SSA result is the presence of an upward trend, of unprecedented steepness, which begins around 1980. During the last years, the observed positive trend of the NAO index since 1960 has been speculated to be associated with an anthropogenic effect on climate [*Trenberth*, 1995; *Wallace et al.*, 1995; *Hurrell*, 1996]. This trend arises from a running mean computing of the NAO index series. Our methodology, using SSA, suggest that the trend begins not around 1960 but between 1970 and 1980. Many of the high positive values observed in the winter NAO index in the eighties and nineties appears to be associated with the simultaneous presence of this positive trend and the amplitude-modulated oscillation with period around 7.7 years, which in these periods shows very high amplitude. Additionally, our results suggest the presence of a short upward trend between 1890 and 1920 followed by a downward trend between 1920 and roughly 1970, whose slopes are considerably lower than that of the positive trend of the eighties. This low-frequency variability in the NAO index (interdecadal and longer), has been associated with processes involving the ocean [*Taylor*

Table 3. As in Table 2, but for the Model Developed Over the Period 1900–1969

	Raw winter NAO versus one-step-ahead forecast (period 1900–1969)	Raw winter NAO versus one-step-ahead forecast (period 1970–1985)	Raw winter NAO versus several-steps-ahead forecast (period 1970–1985)
MSE	0.72	0.94	1.12
MAE	0.69	0.77	0.86
Correlation coeff.	0.68[a]	0.71[a]	0.69[a]
% Phase accordance	68.8	71.4	57.0

[a] Correlation coefficients that are statistically significant at the 95%.

Table 4. As in Table 1, but for the Subset of Cases in Which the Raw Winter NAO Index is NAO ≥ 1 or NAO ≤ -1

	Raw winter NAO versus SSA-filtered winter NAO (period 1986–2000)	Raw winter NAO versus one-step-ahead forecast (period 1826–1985)	Raw winter NAO versus one-step-ahead forecast (period 1986–2000)	Raw winter NAO versus several-steps-ahead forecast (period 1986–2000)
MSE	1.20	1.35	1.9	2.54
MAE	0.95	1.03	1.23	1.36
Correlation coeff.	0.83[a]	0.82[a]	0.56	0.53
MSE_{cli}		2.22	3.09	3.09
MSE_{per}		3.28	3.63	3.63
% $SMSE_{cli}$		35.0	38.5	17.8
% $SMSE_{per}$		58.8	47.6	30.0
% Phase accordance	91	82	88	66

[a] Correlation coefficients that are statistically significant at the 95%.

and Stephens, 1998; *Rajagopalan et al.*, 1998; *Timmermann et al.*, 1998; *Rodwell et al.*, 1999].

[48] The ARMA modeling of the SSA-filtered winter NAO index results in an ARMA(8,9) model. A similar model was found by *Stephenson et al.* [2000], modeling the stochastic behavior of the NAO index. The forecasting experiment shows that the ARMA(8,9) model presents useful forecasting skills. For one-step-ahead forecasts and over the period 1986–2000, used for true comparison, the model skill is 27.8% better than climatology (in Means Square Error) and 43.3% better than persistence, while percentage of cases in which the actual NAO phase is accurately predicted is 80%. When taking into account only extreme NAO events, the model skills are 38.5% and 47.6% better than climatology and persistence respectively, while percentage of cases in which the actual NAO phase is accurately predicted is 88%. For several-steps-ahead forecast, and over the same 1826–2000 period, skill scores are lower but still better than climatology (9%) and persistence (31.8%), while percentage of cases in which the actual NAO phase is accurately predicted is 60%. For 2000/2001 winter, the prediction was −0.523 while the actual value was −0.44. For 2001/2002 and 2002/2003 winters, persistence in the negative phases is predicted, having NAO index value close to −1. For the 2001/2002 winter, a negative value is also forecasted by *Rodwell and Folland* [in press] using SST data.

[49] The results of the analysis of an univariate time series representative of a complex and nonlinear dynamical system, as those involved in the NAO, must be taken with care [*Yiou et al.*, 1996; *Wunsch*, 1999]. However, this analysis of the NAO index series may help to the understanding of NAO dynamics origin and to provide some complementary information that may be useful for multivariate statistical or dynamical predictions of the NAO. Particularly, since the methodology presented here only uses for the prediction of the NAO state the own history of the series, our results can be used to evaluate the importance of "external" variables in predicting the NAO state.

Appendix A: Singular Spectral Analysis

[50] SSA is a powerful form of the Principal Component Analysis (PCA) of the lag correlation structure of a time series [*Vautard et al.*, 1992], which is particularly successful in isolating multiple period components with fluctuating amplitudes and trends in short and noisy series. SSA was first introduced into the study of dynamical systems by *Broomhead and King* [1986a, 1986b] as a method of visualizing qualitative dynamics from noisy experimental data. In this appendix, a brief review on SSA, including a test to study the significance of the results, is provided. A comprehensive review, explaining in detail the mathematical foundations of SSA, is given by *Vautard et al.* [1992] and *Plaut et al.* [1995].

[51] SSA consists of the diagonalization of the lagged-autocovariance matrix of a time series. As in the PCA, the eigenvectors or EOFs represent patterns of temporal behavior, and the Principal Component series (PCs) are characteristic time series. The order election of the lagged-covariance matrix M (window length or embedding dimension) represents a trade-off between significant information and statistical confidence. A common recommendation is to choose $M \approx N/4$, N being the length of the data [*Vautard and Ghil*, 1989]. Given M as the dimension of the lagged-covariance matrix, the PCs have length $N - M + 1$. An individual PC contains a very limited number of harmonic components. The detailed reconstruction of a set of significant components, called SSA-filtered components (RCs), of the time series is carried out by an optimal linear square fitting between the corresponding PCs and the original data. An RC represents the contribution of its associated EOF to the variance of the time series; the RCs are additive and their sum provides the original time series. When two eigenvalues of the lagged-covariance matrix are nearly equal and their corresponding eigenvectors are orthogonal, they represent an oscillation. So, SSA extracts and reconstructs periodic components from noisy time series. To determine the corresponding frequencies requires, however, estimations of power spectra. A consistent spectral approach, combining SSA with the spectral Fourier method, leads to high-resolution spectral estimates. The MEM is used to evaluate the spectral contents of the PC time series corresponding to the EOFs. The MEM consists of fitting an autoregressive model to the time series and then obtaining its associated spectrum. This method has the advantage of having a high spectral resolution [*Burg*, 1968], allowing to study and detect possible oscillatory modes in the data.

[52] Results from spectral analysis must be taken with caution when analyzing, as in our case, short and noisy time series [*Yiou et al.*, 1996; *Wunsch*, 1999]. Special care must be taken in the study of the significance of the results. We

study the statistical significance of the SSA results by means of a Monte Carlo method, following the indications of *Allen and Smith* [1994]. Given a time series of length N, the lagged-covariance matrix S^{data} is computed and their eigenvalues (λ_k^{data}) and eigenvectors (E_k^{data}) are estimated by diagonalizing the lagged-covariance matrix:

$$\Lambda^{data} = E^{data^T} S^{data} E^{data} \qquad (1)$$

[53] A test against the null hypothesis that the data have been generated by an AR(1) noise process can then be computed. The expression:

$$u_t = \hat{\alpha} u_{t-1} + \varepsilon_t \qquad (2)$$

where $u_0 = 0$, $\hat{\alpha}$ is the expected lag-one covariance and ε_t is a white noise process, can be used to generate an ensemble of surrogate time series, each containing N_t values. Time series generated in this way are simple autocorrelated noise containing no deterministic components. Thus, for each surrogate record in the ensemble, the lagged-covariance matrices, $S^{surrogate}$, and their eigenvalues, $\lambda_k^{surrogate}$, can be computed according to:

$$\Lambda^{surrogate} = E^{data^T} S^{surrogate} E^{data} \qquad (3)$$

[54] The size of the ensemble (p) determines the accuracy to which a significance level can be assessed. Using p = 1000, significance estimates are accurate to the order of 1%. The set of p eigenvalues for each mode (k = 1, ..., M) can be regarded as a sample distribution from which percentiles can be determined. The appropriate significance level percentiles can be computed from the distributions of $\lambda_k^{surrogate}$. If, for a given k, λ_k^{data} lies above the 95th percentile of the $\lambda_k^{surrogate}$, then the kth eigenvector explains an unlikely large portion of the variance in the data series given the null hypothesis. In this work, we use a variation of this approach, proposed by *Allen* [1992]. This variation consists of using a single set of eigenvectors from the data and project the lagged-covariance matrix of the surrogate onto this basis to obtain $\lambda_k^{surrogate}$.

[55] Once a component of the series has been identified as a signal, the rest of the spectrum can be examined to determine whether or not it is simply noise. In this study, the method described by *Allen and Smith* [1996, p. 3387] is used. The aim of this methodology is to identify the null hypothesis parameters of the red noise process, after filtering to suppress variance in the directions defined by the EOFs (that we have already identified as signal). This red noise process must have the same variance and lag-one autocorrelation of the data series, when the same filtering procedure is applied. An estimate of the lag-covariance matrix of the filtered signal is given by:

$$S^{data'} \approx E^{data}(I - K)E^{data^T}S^{data}, \qquad (4)$$

where K is a M × M diagonal matrix with $K_{kk} = 0$ if the eigenvector k has been identified as corresponding to a signal and $K_{kk} = 1$ otherwise. If the BK algorithm is used, equation (4) is satisfied exactly. $S^{data'}$ can be used as an estimate of the filtered lag-covariance matrix. *Allen and*

Smith [1996] provide a comprehensive review of this technique.

Appendix B: ARMA Modeling and Forecasting
B.1.   Fitting Procedure

[56] In this appendix, a briefly review of the ARMA models [*Box and Jenkins*, 1976], including their definition and modeling guidelines, and comments on the software routines used in the work, are given. Statistical software package S-plus (StatSci 1995) has been used. A comprehensive review, explaining in detail how to fit ARMA models to data sets following the identification, estimation, and diagnostic check stages, is given by *Brockwell and Davis* [1996] and *Hipel and Mcleod* [1994].

[57] A stochastic process $\{X_t\}$, with mean zero, has an ARMA(p,q) (ARIMA(p,0,q), stationary Autoregressive Integrated Moving Average) representation if it can be expressed in the form:

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2^-} \ldots\ldots - \phi_p X_{t-p}$$
$$= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2^-} \ldots\ldots - \theta_q a_{t-q} \qquad (5)$$

where $\{a_t\}$ is a white noise Gaussian process (normality is not necessary in general) with variance σ_a^2 and zero mean; p and q are nonnegative integers, $\{\phi_1, ..., \phi_p\}$ are the autoregressive (AR) coefficients and $\{\theta_1, ..., \theta_q\}$ are the moving average (MA) coefficients.

[58] The order of the model is selected, in a preliminary approach, studying the ACF and PACF. A satisfactory estimate of the *k*th lag ACF value for a time series $\{x_t\}$ can be obtained in the form [*Box and Jenkins*, 1976, chapter 2]:

$$r_k = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\left[\sum_{t=1}^{n}(x_t - \bar{x})\right]^2}, \qquad (6)$$

where $\bar{x}$ is the mean of the series. A estimation of the PACF can be carried out by fitting to the $\{x_t\}$ series, by least squares and successively, autoregressive process of orders 1, 2, 3, etc, and picking our the estimates of the last fitted at each stage [*Box and Jenkins*, 1976, chapter 2]:

$$x_t = a_{11}x_{t-1}; x_t = a_{12}x_{t-1} + a_{22}x_{t-2};$$
$$x_t = a_{12}x_{t-2} + a_{22}x_{t-2} + \ldots + a_{kk}x_{t-k}.$$

The PACF estimate at lag is, then, a_{11}, at lag 2 a_{22} and at lag k a_{kk}.

[59] Gaussian maximum likelihood estimates for the $\{\phi_1, ..., \phi_p\}$ and $\{\theta_1, ..., \theta_q\}$ parameters can be calculated using the S-plus "arima.mle" function. To formulate a physically meaningful model for our time series, we must impose some constraints (concerning the stationarity and invertibility) on the parameters of the ARMA model, and some consideration concerning the parsimony must be taking into account. Besides this, a candidate model must have a white noise process as a residual time series. In our analysis, the residual time series has been checked using S-plus "arima.diag" procedure that includes a Pormanteau test statistic Q for the study of the correlation of a series. This test was developed by *Box and Pierce* [1970] and later refined by *Ljung and Box* [1978], who showed that, for a white noise sequence $\{a_t: t = 1, ..., n\}$, with m ≪ n being

the maximum number of lags and $\hat{\gamma}_k^2$ being the sample autocovariance at lag k, the statistic:

$$Q_m = n(n+2) \sum_{k=1}^{m} (n-k)^{-1} \hat{\gamma}_k^2 \qquad (7)$$

is approximately distributed as a χ^2 random variable with "m" degrees. Besides this test, a further analysis of the residual time series was made studying its ACF and PACF. Only those models that met the requirements were considered further.

[60] In physical terms, the best model has as few parameters as possible. We have used the AIC [*Akaike*, 1974] to select the final model among all the candidates. The AIC is based on information theory and represents a compromise between the goodness of the fit and the number of parameters of the model. The model with the lowest AIC value should be selected. However, the best way to check the adequacy of a model is to study the way in which it forecasts the future values of the modeled series. Given an ARMA(p,q) model, the forecast with the minimum mean squared error for a leading time $\hat{x}_t(L)$ is the conditional expectation $E_t[x_{t+L}]$ of x_{t+L} at origin "t":

$$\hat{x}_t(L) = E_t[x_{t+L}] = \phi_1 E_t[x_{t+L-1}] + \phi_2 E_t[x_{t+L-2}] + \ldots + \phi_p E_t[x_{t+L-p}]$$
$$+ E_t[a_{t+L}] - \theta_1 E_t[a_{t+L-1}] - \theta_2 E_t[a_{t+L-2}] - \ldots - \theta_q E_t[a_{t+L-q}]. \qquad (8)$$

[61] The forecast error $e_t(L)$ can be obtained in the form:

$$e_t(L) = x_{t+L} - \hat{x}_t(L) + a_{t+L} + \psi_1 a_{t+L-1} + \ldots\ldots + \psi_{L-1} a_{t+1}, \qquad (9)$$

"ψ_i" are the coefficients of the ARMA model in the random shock form (that is, expressed as an MA(∞) model). The variance of the forecast error can then be obtained in the form:

$$Var(e_t(L)) = E_t[e_t(L)^2] = [1 + \psi_1^2 + \ldots\ldots + \psi_{L-1}^2] \sigma_\alpha^2. \qquad (10)$$

[62] Particularly, the one-step-ahead forecast error is:

$$e_t(1) = x_{t+1} - \hat{x}_t(1) = a_{t+1}. \qquad (11)$$

[63] The variance of the this {e_t: t = 1, …, n} series is called the innovations variance and gives a measure of the variance of the modeled series not accounted by the ARMA model.

[64] Equation (8) states how an ARMA model forecasts future values of a time series given the current and past values of the series. The coefficients {ϕ_i} and {θ_i} of the ARMA model are obtained in such a way that each value of our sample series {x_t} (that is, a realization of the stochastic process {X_t}) can be calculated from a linear combination of past values and innovations (which are samples of a white noise process) using these coefficients as weights. Assuming that the series is stationary, we might expect future values (not included in our sample realization {x_t}) will follow the same linear relationship between past and current values, and, therefore these future values can be derived using the same weights on past values.

[65] Section 3.2.2 in the work shows some forecasting experiments based on the fitted models; subroutines "arima.filt" and "arima.forecast" of S-plus software were used, respectively, to make one-step-ahead and several-steps-ahead forecasts. We should be aware of the fact that forecasts projected with ARMA models are influenced not only by the goodness of the fit but also by the assumptions that the underlying physical process related to the series does not change during the forecasting time. However, this latter assumption is hardly ever true in dynamic systems like climate.

B.2. Separate Training and Forecast Intervals and Cross-Validation

[66] The firm separation of training and forecast periods is fundamental for true skill assessment. We employ data from 1826 to 1985 to fit the model while data from 1986 to 2000 is used for a true comparison in the forecasting experiment. Additionally, a cross-validation of the model is conducted to validate the model. Commonly, the Cross-validation of a regression model is carried out using development data sets of size n − 1 and verification data sets containing the remainder single observation of the predictand, this leads to n partitions of the data set. The model is then calculated for each of these partitions, resulting in n similar forecast equations, each computed without one of the observations of the predictand.

[67] This procedure cannot be applied in our case for several reasons. Usually, in the regression models, some variable are use to predict one other variable; in our case, we must obtain the potentially predictable signal from the own history of the series. When fitting ARMA models, the temporal location of each data cares: the "history" of the series is very important. When removing one single data in the middle of the data set, the remaining data are not useful to fit the model, because the temporal structure of the data is then broken. Furthermore, in a regression model we know a priori the temporal dependence between the predictand and the predictors. This allows to properly removing some samples from the data set and fitting the model using the rest of the sample. In an ARMA model, we do not know a priori the temporal dependence of the model.

[68] To cross-validate the ARMA model, and taking into account this singularities of the ARMA models, we have divided the development period 1826–1985 in two subperiods, 1826–1899 and 1900–1969. We have then fitted ARMA models in these two subperiods and have carried out one-step-ahead forecast and several-steps-ahead forecast over 1900–1915 for the 1826–1899 subperiod model and over 1970–1985 for the 1900–1985 subperiod model. Results are compared with those of the whole period 1826–1985.

B.3. Setup of the ARMA Forecasting Procedure

[69] The period 1826–1985 of the filtered NAO series is used to develop de ARMA model and the period 1986–2000 for "true forecasting." Two kinds of forecasting experiments are carried out, the one-step-ahead forecast and the several-steps-ahead forecast. In the first case, the values of the winter NAO index are forecasted for the following winter, over the period 1985–2000. Note that the initial condition to forecast the 1986 winter NAO index

were set up based on the state of the NAO until the previous winter (1985), including this late; for the forecast of the 1987 winter NAO index we use the information up to the previous winter (1986) and so on. In the second case, the several-steps-ahead forecast, we forecast the NAO index value for the period 1986–2010 using only the information up to the year 1985 (including this late). Thus, the forecast for the year 1986 is a one-step-ahead forecast, the forecast for the year 1987 is a 2-years-ahead forecast and so on.

Appendix C: Accuracy and Skill Scores

[70]   To asses the extend to which the SSA model is able to reproduce the NAO index and the performance of the ARMA model forecasting experiments, a set of commonly used scores are used.

[71]   First at all, the Pearson correlation coefficient is used. As accuracy measures, the Mean Absolute Error (MAE) and the Mean Square Error (MSE) are employed. To assess the skill of the forecast models, the following scores are used:

1. The percentage improvement in mean square error forecast over a climatological forecast ($SMSE_{cl}$) and the percentage improvement in mean square error forecast over a persistence forecast ($SMSE_{pe}$). Climatology is taken as the 30 values average prior each value being forecast and persistence value is taken from the previous value to those being forecast.

2. Additionally, the LEPS skill [*Potts et al.*, 1996] is computed. LEPS measures an error of a forecast as the "distance" in a chosen climatological cumulative probability distribution between a forecast and the corresponding observation, referred to the chance distance created by random forecast [*Potts et al.*, 1996]. Thus, if the error in forecast was exactly Equal to chance error, the LEPS score would be zero. Three equiprobable categories are used based on Table 1 [*Potts et al.*, 1996]. LEPS are then converted in percentage skill scores, ranging from −100 to +100% as in section 7 of that paper.

References

Akaike, H., A new look at the statistical model identification, *IEEE Trans. Autom. Control*, *19*, 716–723, 1974.

Allen, M. R., Interactions between the atmosphere and oceans on time-scales of weeks to years, Ph.D. thesis, 202 pp., Univ. of Oxford, Oxford, 1992.

Allen, M. R., and L. A. Smith, Investigating the origins and significance of low-frequency modes of climate variability, *Geophys. Res. Lett.*, *21*, 883–886, 1994.

Allen, M. R., and L. A. Smith, Monte Carlo SSA: Detecting irregular oscillations in the presence of coloured noise, *J. Clim.*, *9*, 3373–3404, 1996.

Appenzeller, C., T. F. Socker, and M. Anklin, North Atlantic Oscillation dynamics recorded in Greenland ice cores, *Science*, *282*, 446–450, 1998.

Barnston, A. G., and R. E. Livezey, Classification, seasonality and persistence of low-frequency atmospheric circulation patterns, *Mon. Weather Rev.*, *115*, 1083–1126, 1987.

Box, G. E. P., and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, 543 pp., Holden-Day, Boca Raton, Fla., 1976.

Box, G. E. P., and D. A. Pierce, Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, *J. Am. Stat. Assoc.*, *15*, 1509–1526, 1970.

Brockwell, P. J., and R. A. Davis, *Introduction to Time Series and Forecasting*, 420 pp., Springer-Verlag, New York, 1996.

Broomhead, D. S., and G. P. King, Extracting qualitative dynamics from experimental data, *Physica D*, *20*, 217–236, 1986a.

Broomhead, D. S., and G. P. King, On the qualitative analysis of experimental dynamical systems, in *Nonlinear Phenomena and Chaos*, edited by S. Sarker, pp. 113–144, Adam Hilger, Bristol, 1986b.

Burg, J. P., A new analysis technique for time series data, in *NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics*, 1968.

Currie, R. G., Luni-solar 18.6 and 10–11 year solar cycle signal in USA air temperature records, *Int. J. Climatol.*, *13*, 31–50, 1993.

Delworth, T. L., North Atlantic interannual variability in a coupled ocean–atmosphere model, *J. Clim.*, *9*, 2356–2375, 1996.

Dettinger, M. D., M. Ghil, C. M. Strong, W. Weibel, and P. Yiou, Software for singular spectrum analysis of noisy time series, *Eos Trans. AGU*, *76*(2), 12, 1995.

Dong, B. W., R. T. Sutton, S. P. Jewson, A. O'Neill, and J. M. Slingo, Predictable winter climate in the North Atlantic sector during the 1997–1999 ENSO cycle, *Geophys. Res. Lett.*, *27*, 985–988, 2000.

Elsner, J. B., and A. A. Tsonis, *Singular Spectrum Analysis: A New Tool in Time Series Analysis*, pp. 39–83, Plenum, New York, 1996.

Folland, C. K., D. E. Parker, and F. E. Kates, Worldwide marine temperature fluctuations 1856–1981, *Nature*, *310*, 670–673, 1984.

Fromentin, J. M., and B. Planque, Calanus and environmental in the eastern North Atlantic, 2, Influence of the North Atlantic Oscillation on c-Finmarchicus and c-Helgolandicus, *Mar. Ecol. Prog. Ser.*, *134*(1–3), 111–118, 1996.

Ghil, M., and R. Vautard, Interdecadal oscillations and the warming trend in global temperature series, *Nature*, *310*, 324–327, 1991.

Grötzner, A., M. Lafif, and D. Dommenget, Atmospheric response to sea-surface temperature anomalies during El-Niño 1997/98 as simulated by Echam 4, *Q. J. R. Meteorol. Soc.*, *126*, 2175–2198, 2000.

Halpert, M. S., and G. D. Bell, Climate assessment for, *Bull. Am. Meteorol. Soc.*, *78*, 50, 1997.

Hipel, K. W., and A. I. Mcleod, *Time Series Modelling of Water Resources and Environmental Systems*, 1013 pp., Elsevier Sci., New York, 1994.

Hurrell, J. W., Decadal trends in North Atlantic Oscillation and relationship to regional temperature and precipitation, *Science*, *269*, 676–679, 1995.

Hurrell, J. M., Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperatures, *Geophys. Res. Lett.*, *23*, 665–668, 1996.

Hurrell, J. W., and H. van Loon, Decadal variations in climate associated with the North Atlantic Oscillation, *Clim. Change*, *36*, 301–326, 1997.

Jones, P. D., The early twentieth century Arctic High: Fact or fiction?, *Clim. Dyn.*, *1*, 63–75, 1987.

Jones, P. D., T. Jonsson, and D. Wheeler, Extension to the North Atlantic Oscillation index using early instrumental pressure observations from Gibraltar and southwest Iceland, *Int. J. Climatol.*, *17*, 1–18, 1997.

Kiladis, N., and H. F. Díaz, Global climatic anomalies associated with extremes of the Southern Oscillation, *J. Clim.*, *2*, 1069–1090, 1989.

Latif, M., D. Anderson, T. Barnett, M. Cane, R. Kleeman, A. Leetman, J. O'Brien, A. Rosati, and E. Scheneider, A review of the predictability and prediction of ENSO, *J. Geophys. Res.*, *103*, 14,375–14,393, 1998.

Ljung, G. M., and G. E. Box, On a measure of lack of fit in time series models, *Biometrika*, *65*, 297–303, 1978.

Moses, T., G. Kiladis, H. F. Díaz, and R. Barry, Characteristic and frequency of reversal in mean sea level pressure in the North Atlantic sector and their relationship to long-term temperature trends, *Int. J. Climatol.*, *7*, 13–30, 1987.

Osborn, J., K. R. Briffa, S. F. B. Tett, P. D. Jones, and R. M. Trigo, Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model, *Clim. Dyn.*, *15*, 685–702, 1999.

Panckratz, A., *Forecasting with Dynamic Regression Models*, John Wiley, New York, 1991.

Perlwitz, J., and H. F. Graf, The statistical connection between tropospheric and stratospheric circulation of the Northern Hemisphere in winter, *J. Clim.*, *8*, 2281–2295, 1995.

Plaut, G., M. Ghil, and R. Vautard, Interannual and interdecadal variability in 335 years of central England temperatures, *Science*, *268*, 710–713, 1995.

Pozo-Vázquez, D., M. J. Esteban-Parra, F. S. Rodrigo, and Y. Castro-Díez, An analysis of the variability of the North Atlantic Oscillation in the time and the frequency domains, *Int. J. Climatol.*, *20*, 1675–1692, 2000.

Pozo-Vázquez, D., M. J. Esteban-Parra, F. S. Rodrigo, and Y. Castro-Díez, A study on NAO variability and its possible non-linear influences on European surface temperatures, *Clim. Dyn.*, *17*, 701–715, 2001a.

Pozo-Vázquez, D., M. J. Esteban-Parra, F. S. Rodrigo, and Y. Castro-Díez, The association between ENSO and winter atmospheric circulation and temperature in the North Atlantic region, *J. Clim.*, 2001b.

Rajagopalan, B., Y. Kushnir, and Y. Tourre, Observed decadal midlatitude and tropical Atlantic climate variability, *Geophys. Res. Lett.*, *25*, 3967–3970, 1998.

Robertson, A. W., On the influence of ocean–atmosphere interaction on the

Arctic Oscillation in two general circulation models, *J. Clim.*, *14*, 3240–3254, 2001.

Robertson, A. W., C. R. Mechoso, and Y. J. Kim, The influence of the Atlantic Sea surface temperature anomalies on the North Atlantic Oscillation, *J. Clim.*, *13*, 122–138, 2000.

Rodrigo, F. S., D. Pozo-Vázquez, M. J. Esteban-Parra, and Y. Castro-Díez, A reconstruction of the winter North Atlantic Oscillation index back to A.D. 1501 using documentary data in southern Spain, *J. Geophys. Res.*, *14*, 805–818, 2001.

Rodwell, M. J., and C. K. Folland, Atlantic air–sea interaction and seasonal predictability, *Q. J. R. Meteorol. Soc.*, in press.

Rodwell, M. J., D. P. Rowell, and C. F. Folland, Oceanic forcing of the wintertime North Atlantic Oscillation and European climate, *Nature*, *398*, 320–323, 1999.

Rogers, J. C., The association between the NORTH Atlantic Oscillation and the Southern Oscillation in the Northern Hemisphere, *Mon. Weather Rev.*, *112*, 1999–2015, 1984.

Ropelewski, C. F., and P. D. Jones, An extension of the Tahiti–Darwin Southern Oscillation index, *Mon. Weather Rev.*, *115*, 2161–2165, 1987.

Stephenson, D. B., V. Pavan, and R. Bojariu, Is the North Atlantic Oscillation a random walk?, *Int. J. Climatol.*, *20*, 1–18, 2000.

Taylor, A. H., and J. A. Stephens, The North Atlantic Oscillation and the latitude of the gulf stream, *Tellus*, *A50*, 134–142, 1998.

Timmermann, A., M. Latif, R. Voss, and A. Grötzner, Northern Hemispheric interdecadal variability: A coupled air–sea mode, *J. Clim.*, *8*, 1906–1931, 1998.

Trenberth, K. E., Signal versus noise in the Southern Oscillation, *Mon. Weather Rev.*, *112*, 327–333, 1984.

Trenberth, K. E., Atmospheric circulation climate changes, *Clim. Change*, *31*, 427–453, 1995.

van Loon, H., and R. A. Madden, The Southern Oscillation, part I, Global associations with pressure and temperature in the northern winter, *Mon. Weather Rev.*, *104*, 1354–1361, 1981.

Vautard, R., and M. Ghil, Singular spectrum analysis in non-linear dynamics with applications to paleoclimatic time series, *Physica D*, *35*, 395–424, 1989.

Vautard, R., P. Yiou, and M. Ghil, Singular spectrum analysis: A toolkit for short, noisy chaotic signal, *Physica D*, *58*, 95–126, 1992.

Wallace, J. M., and D. S. Gutzler, Teleconnections in the geopotential height field during the Northern Hemisphere winter, *Mon. Weather Rev.*, *109*, 784–812, 1981.

Wallace, J. M., Y. Zhang, and J. A. Renwick, Dynamical contribution to hemispheric temperature trends, *Science*, *270*, 780–783, 1995.

Wunsch, C., The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations, *Bull. Am. Meteorol. Soc.*, *80*, 245–255, 1999.

Yiou, P., E. Baert, and M. F. Loutre, Spectral analysis of climate data, *Surv. Geophys.*, *17*, 619–663, 1996.

———————

Y. Castro-Díez, M. J. Esteban-Parra, and S. R. Gámiz-Frotis, Department of Applied Physics, University of Granada, E-18071, Granada, Spain.

D. Pozo-Vázquez, Department of Physics, University of Jaén, E-23071, Jaén, Spain.