Winter North Atlantic Oscillation Hindcast Skill: 1900–2001

CHRISTOPHER G. FLETCHER* AND MARK A. SAUNDERS

Department of Space and Climate Physics, Benfield Hazard Research Centre, University College London, Surrey, United Kingdom

(Manuscript received 7 January 2005, in final form 13 March 2006)

ABSTRACT

Recent proposed seasonal hindcast skill estimates for the winter North Atlantic Oscillation (NAO) are derived from different lagged predictors, NAO indices, skill assessment periods, and skill validation methodologies. This creates confusion concerning what is the best-lagged predictor of the winter NAO. To rectify this situation, a standardized comparison of NAO cross-validated hindcast skill is performed against three NAO indices over three extended periods (1900–2001, 1950–2001, and 1972–2001). The lagged predictors comprise four previously published predictors involving anomalies in North Atlantic sea surface temperature (SST), Northern Hemisphere (NH) snow cover, and an additional predictor, an index of NH subpolar summer air temperature (T_{SP}). Significant ($p < 0.05$) NAO hindcast skill is found with May SST 1900–2001, summer/autumn SST 1950–2001, and warm season snow cover 1972–2001. However, the highest and most significant hindcast skill for all periods and all NAO indices is achieved with T_{SP}. Hindcast skill is nonstationary using all predictors and is highest during 1972–2001 with a T_{SP} correlation skill of 0.59 and a mean-squared skill score of 35%. Observational evidence is presented to support a dynamical link between summer T_{SP} and the winter NAO. Summer T_{SP} is associated with a contemporaneous midlatitude zonal wind anomaly. This leads a pattern of North Atlantic SST that persists through autumn. Autumn SSTs may force a direct thermal NAO response or initiate a response via a third variable. These findings suggest that the NH subpolar regions may provide additional winter NAO lagged predictability alongside the midlatitudes and the Tropics.

1. Introduction

The North Atlantic Oscillation (NAO) is the dominant mode of boreal winter climate variability over the North Atlantic (Walker and Bliss 1932; Barnston and Livezey 1987). The NAO is strongly linked to patterns of winter temperature, precipitation, and storminess over the whole North Atlantic sector (Hurrell 1995; Trigo et al. 2002). Accurate and timely forecasts of the winter NAO are therefore an important challenge for seasonal forecasters and increasing research is focusing on improving the quality and lead time of NAO predictions.

Empirical winter NAO forecast models are formulated using observations of slowly varying boundary variables such as sea surface temperature (SST) and snow cover. The persistence of these boundary conditions and their forcing of the atmosphere can make seasonal atmospheric conditions predictable at a later time (Goddard et al. 2001). Winter NAO forecasts have also recently been produced using general circulation models (GCMs; e.g., Palmer et al. 2004). However, only one or two GCMs produce statistically significant forecast skill at levels comparable to empirical models, and only when assessed over a relatively short time period (1980–2001).

Recent empirical studies have uncovered links between North Atlantic SST anomalies in the preceding late spring/summer/autumn and the upcoming winter NAO. Rodwell and Folland (2002) highlighted a pattern of May North Atlantic SST that is related significantly to the upcoming winter mean 500-hPa geopotential height field, achieving a hindcast correlation skill (r_s) of $r_s = 0.45$ for the period 1948–97. Saunders and Qian (2002) found that two modes of late summer/early autumn North Atlantic SST variability 1950–2001 were skillful in predicting a range of upcoming winter NAO

* Current affiliation: Atmospheric Physics Group, Department of Physics, University of Toronto, Toronto, Ontario, Canada.

Corresponding author address: Christopher Fletcher, Atmospheric Physics Group, Department of Physics, University of Toronto, 60 St. George St., Toronto, ON M5S 1A7, Canada.
E-mail: cgf@atmosp.physics.utoronto.ca

indices with r_s between $r_s = 0.47$ and $r_s = 0.63$. A horseshoelike North Atlantic summer SST pattern has also been shown to explain up to 16% of the early winter atmospheric variance 1958–97 (Drévillon et al. 2001; Czaja and Frankignoul 2002; Cassou et al. 2004).

Another lagged NAO predictor is the monthly mean areal extent of snow cover over different regions of the Northern Hemisphere (NH). Reliable satellite observations of snow cover are available only since 1972 (Robinson et al. 1993). Thus, the role of anomalous snow cover in influencing atmospheric circulation and climate has received less attention than that of SST anomalies. Saunders et al. (2003) used NH snow cover to predict a range of NAO indices and found that the June–July mean produced the most significant hindcast skill with $r_s = 0.61$ for 1972–2001. A significant correlation of $r = -0.56$ is highlighted by Bojariu and Gimeno (2003) between April and October mean Eurasian (EU) snow cover and the upcoming winter NAO index 1973–98. These studies suggest a recent significant link between NH climate in the warm season and that in the upcoming winter. Cohen and Entekhabi (1999) highlighted the significant relationship between EU snow cover during fall and the upcoming winter Arctic Oscillation (AO) index. However, these links were not significant when assessed against the NAO index.

The above discussion highlights the current confusing situation in the literature where several authors claim to have uncovered significant predictive links and skill for the winter NAO using different prior climatic predictors. Furthermore, these authors all reference these links and skill to different NAO indices and/or employ different predictand time periods and/or use different skill assessment methodologies and measures. In this study we compare the most skillful recent empirical hindcast models using the same hindcast procedure for all predictors. This provides a standardized platform on which to evaluate and compare the skill arising from the chosen NAO predictors. Hindcasts are made for three different winter NAO indices over three extended time periods out to 100 yr and skill is assessed using two skill measures. Where possible, hindcast skill sensitivity to choice of predictor observational dataset is also examined by using two datasets for each analysis. These methods will allow us to quantify which of the lagged predictors drawn from the recent literature provides the best hindcast skill for the upcoming winter NAO.

The paper is structured as follows. Section 2 outlines the chosen assessment periods, predictand NAO indices, predictors and the different observational datasets used to compute them. Here we also describe our hindcast procedure and the method for evaluating hindcast skill and its statistical significance. Results from the hindcast comparison are presented in section 3, where stationarity of the predictive relationships is also discussed. Section 4 provides an interpretation of the key findings and examines the physical mechanisms linking the most skillful lagged predictor to the winter NAO. Section 5 discusses the implications of our work for seasonal forecasting in general and a summary and conclusions are presented in section 6.

2. Data and methods

Henceforth, "winter" denotes the three-month December–February (DJF) seasonal mean and "winter NAO" will be referred to as NAO_{DJF}. Previous studies of NAO_{DJF} hindcast skill have examined data since 1950. Here, we assess hindcast skill, where possible, over the entire twentieth century. Specifically, NAO_{DJF} hindcast skill is assessed for the three periods 1900–2001, 1950–2001, and 1972–2001, with the 1972–2001 period corresponding to the interval of reliable satellite snow cover data (Robinson et al. 1993). These three assessment periods also allow the stationarity in NAO_{DJF} hindcast skill to be examined.

Figure 1 shows the time series of the three NAO_{DJF} indices employed as predictands in this study. The first index is the standardized difference in mean sea level pressure (MSLP) between Reykjavik, Iceland, and Gibraltar maintained by the Climatic Research Unit (CRU) at the University of East Anglia (henceforth the CRU NAO_{DJF} index; Jones et al. 1997). The second is defined as the standardized MSLP difference between Stykkisholmur, Iceland, and Ponta Delgada, Azores, provided by the National Center for Atmospheric Research (NCAR) Climate Analysis Section (henceforth the Hurrell NAO_{DJF} index). The third index is the time series of 500-hPa North Atlantic geopotential height coefficients from a singular value decomposition (SVD) analysis with May North Atlantic SST (henceforth the $Z500_{DJF}$ index; Rodwell and Folland 2002).

a. *Previously published NAO_{DJF} predictors*

The four previously published NAO_{DJF} predictors examined in this study are shown in Table 1 and are taken from results published since 2001. They were selected for comparison because they are believed to encompass the highest currently observed seasonal predictive skill for the winter NAO.

The first NAO_{DJF} predictor [henceforth the May SST (SVD)] is the time series of May North Atlantic (10°–80°N) SST derived from an SVD analysis with DJF

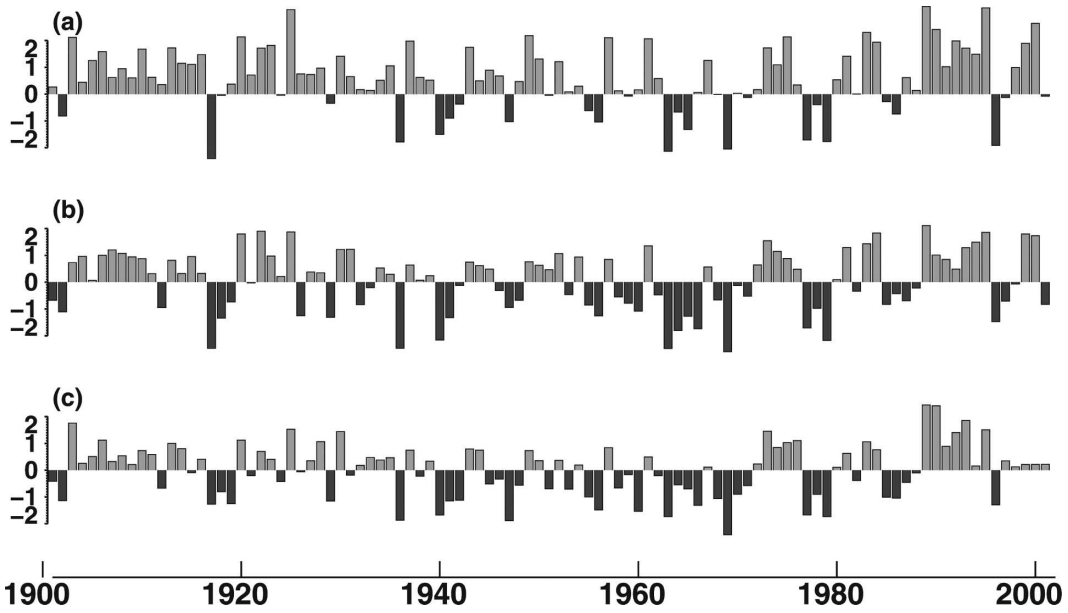

FIG. 1. Temporal evolution of the (a) CRU, (b) Hurrell, and (c) $Z500_{DJF}$ indices 1900–2001 in standardized units. The $Z500_{DJF}$ index is computed using HadSLP data 1900–98. Pearson correlation coefficients between the indices (1900–2001): (a) and (b) = 0.88, (a) and (c) = 0.88, (b) and (c) = 0.85.

mean 500-hPa geopotential height over the North Atlantic sector (Rodwell and Folland 2002). The SVD procedure was performed in "cross-validated" mode, with the resulting SST expansion coefficients called the "predicted" NAO series and the corresponding geopotential height coefficients the "observed" NAO series. The correlation skill was taken as the Pearson correlation between these two time series.

Our implementation of the SVD analysis differs from that employed by Rodwell and Folland (2002). First, in accordance with other analyses in this study, we employ a 5-yr block elimination to the SVD procedure. This means that we remove each year in turn and additionally 2 yr on either side. Second, we remove sea ice from the SST fields prior to performing the SVD (sea ice was retained in the original analysis). Third, we employ different data periods. Fourth, we compute skill from a detrended data analysis. These differences in implementation mean that our results are expected to differ from those of Rodwell and Folland (2002). However, as a preliminary check, we first reproduced the cross-validated hindcast correlation skill (r_s = 0.45) using their original implementation. This predictor maximizes covariance with the $Z500_{DJF}$ index; however, we also assess its hindcast skill against all NAO indices.

The second predictor is the second principal component (PC2) of the June–October (JJASO) mean North Atlantic (0°–65°N) SST. This is slightly the stronger of two lagged JJASO SST predictive modes employed by Saunders and Qian (2002). The PC2 loading pattern (see their Fig. 1a) features a ring of SST anomalies around an opposite-signed center off southeast Newfoundland.

The third and fourth predictors employ the monthly mean areal extent of snow cover, with observations considered reliable since 1972. The third predictor is snow cover over Eurasia for the period April–October (A–O; Bojariu and Gimeno 2003). The fourth predictor

TABLE 1. Published NAO_{DJF} lagged predictors examined in this paper.

Lagged predictor	Assessment period	Lead time (months)	Domain area	NAO_{DJF} index	Reference
May SST (SVD)	1948–98	6	10°–80°N, 90°W–40°E	$Z500_{DJF}$	Rodwell and Folland (2002)
JJASO SST (PC2)	1950–2001	1	0°–65°N, 100°W–0°	CRU/CPC*/PC1	Saunders and Qian (2002)
AMJJASO EU snow cover	1972–2000	1	n/a	Hurrell (1995)	Bojariu and Gimeno (2003)
JJ NH snow cover	1972–2002	4	n/a	CRU/CPC/PC1	Saunders et al. (2003)

* U.S. Climate Prediction Center index.

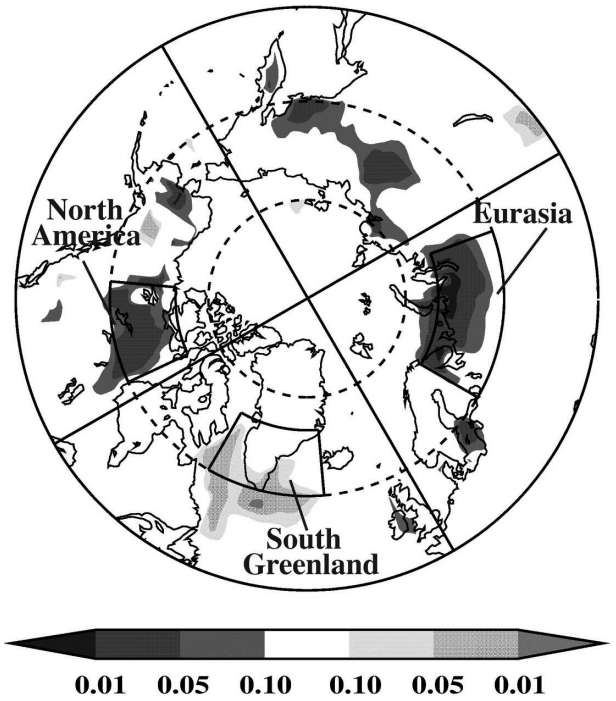

FIG. 2. The correlation pattern significance between detrended time series of JJ NH snow cover extent and gridded JJ 2-m air temperature. Significances are corrected for serial correlation with lags out to 15 yr included. Grayscale denotes where correlation is positive (light) or negative (dark). (After Saunders et al. 2003.)

is the June–July (JJ) mean snow cover for the entire NH (Saunders et al. 2003).

b. An additional NAO_{DJF} predictor

Alongside the four previously published predictors outlined above, we also compare an index of JJ mean NH subpolar (60°–70°N) 2-m air temperature. This index and its link to the upcoming winter NAO was introduced by Saunders et al. (2003). The index is defined as

$$T_{\text{SUBPOLAR}} = \frac{\text{NA} + \text{EU}}{2} - \text{SG}, \qquad (1)$$

where North America (NA), Eurasia (EU), and southern Greenland (SG) refer to subpolar (60°–70°N) 2-m air temperatures over these areas (Fig. 2). These centers correspond to where gridded JJ air temperature is correlated most significantly with contemporaneous NH snow cover ($|r| \sim 0.5$ 1972–2001).

Gridded 2-m air temperature data back to 1900 are available from the CRU Air Temperature Amomalies version 2 (CRUTEM2) dataset (Jones and Moberg 2003) for NA, EU, and SG. We therefore calculate the T_{SUBPOLAR} index (T_{SP}) for the period 1900–2001 and employ it as an NAO_{DJF} predictor. The summer period is selected because during the summer months the relationship peaks between the T_{SP} index and the upcoming winter NAO (Saunders et al. 2003). Figure 3 shows that the link is strongest in JJ using both the CRUTEM2 and the (National Centers for Environmental Prediction) NCEP–NCAR reanalysis data (Kalnay et al. 1996) for the 50- and 30-yr assessment periods. However, for the 100-yr period the link is significant over an extended summer period May–September (MJJAS). Therefore, we also evaluate the NAO_{DJF} predictive skill of the longer MJJAS mean T_{SP}.

c. Predictor datasets

The May SST (SVD) assessment back to 1900 is performed using MSLP data because pre-1948 geopotential height data are not available. This analysis is performed over the shorter period 1900–98. This is the period of extended MSLP data available from the United Kingdom Met Office (UKMO) Hadley Centre Sea Level Pressure (HadSLP) dataset, which is an update of the Met Office's Global Mean Sea Level Pressure (GMSLP2.1f; Basnett and Parker 1997). Geopotential heights at 500 hPa (1950–2001) are taken from the NCEP–NCAR reanalysis (Kalnay et al. 1996). The NH gridded surface air temperatures 1900–2001 are from the University of East Anglia CRUTEM2 (5° × 5° latitude–longitude grid) dataset (Jones and Moberg 2003) and for the periods since 1950 are also from the NCEP–NCAR reanalysis.

Two global gridded SST datasets are employed in this study. For the period 1900–2001, we use the Met Office Hadley Centre Coupled Sea Ice and SST (HadISST) dataset (Rayner et al. 2003) and for the period 1950–2001 SSTs are also taken from NCEP–NCAR reanalysis. All SST data are interpolated onto a 2.5° × 2.5° latitude–longitude grid for computational efficiency. Snow cover data for EU and NH are satellite-derived estimates of monthly mean areal extent (Robinson et al. 1993).

Throughout our analysis we employ linear detrended time series corrected for autocorrelation to compute regressions and hindcast skill. This approach minimizes the influence of time series trends and multiyear-to-decadal signal variability on the strength and significance of the deduced hindcast skill. The use of raw (not detrended) time series corrected for autocorrelation gives in most cases hindcast skills that are of similar magnitude and significance to those achieved with the detrended data analysis. For reference, Table 2 shows the linear trends in NAO_{DJF} and the three predictors for each assessment period. Over the 100-yr period the trends are negative or close to zero. However, since

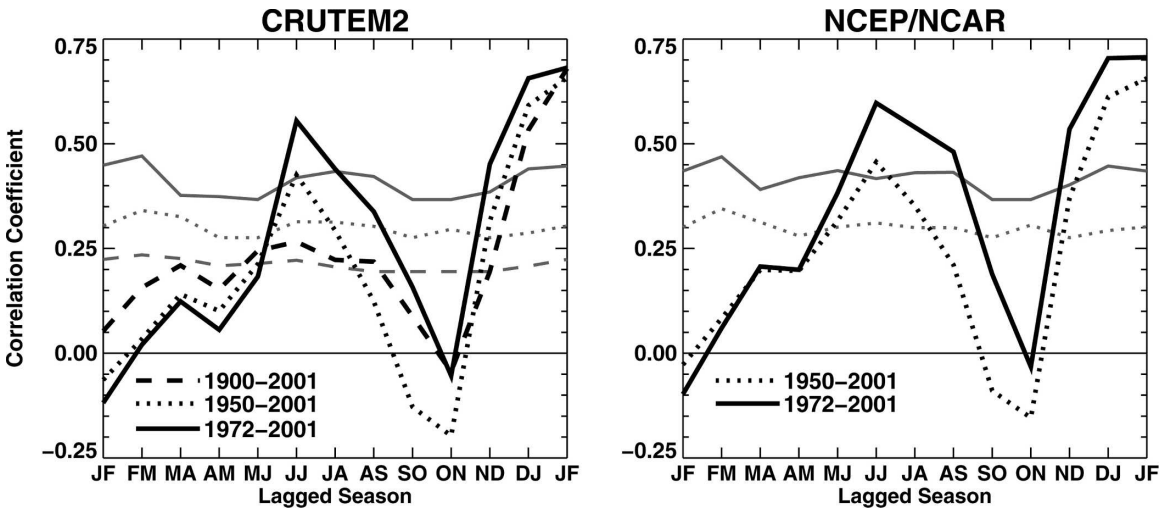

FIG. 3. Magnitude, sign, and significance of linear correlation between the lagged bimonthly T_{SP} temperature index and the upcoming CRU NAO_{DJF} index. (left) Correlations for the CRUTEM2 T_{SP} index and (right) NCEP–NCAR data. Dark lines denote the correlation magnitude and sign. Faint lines denote the correlation 5% significance level corrected for serial correlation.

1950 all three NAO_{DJF} indices and JJ T_{SP} exhibit positive trends of ~0.2–0.3 standard deviations per decade. Also shown is the percentage of NAO_{DJF} hindcast correlation skill that is attributable to linear trend, which is calculated as

$$\Delta_{\text{SKILL}} = \left(1 - \frac{r_s}{r_s'}\right) \times 100\%, \tag{2}$$

where r_s is the correlation skill from the detrended data analysis and r_s' is the correlation skill obtained using raw (not detrended) time series. For the 1900–2001 period the inclusion of trends decreases the hindcast skill by 4%–200%; for the 1950–2001 period it increases skill by 14%–46% (for May SST there is a 30%–50% decrease); for 1972–2001 it decreases skill by 0%–40% (for May SST there is a 40%–60% increase). These differences reflect the decadal and longer time-scale variability in the NAO_{DJF} and predictor indices (see also section 3b).

d. Hindcast methodology

The predictive skill of the selected NAO_{DJF} predictors is computed for each assessment period using cross-validated (Michaelsen 1987) hindcasts with block elimination. Except for the May SST (SVD) predictor, linear empirical models for the NAO_{DJF} index $\hat{y}$ in a given year t using predictor x take the following form:

$$\hat{y}_t = \beta_0 + \beta_1 x_t + e_t, \tag{3}$$

where the coefficients β_0 and β_1 are determined by an ordinary least squares regression and e_t is the residual. A 5-yr block elimination is employed, which means that for a given predictand year t, data for the years $t - 2$, $t - 1$, t, $t + 1$, and $t + 2$ are removed from the regression. Furthermore, predictors derived from principal

TABLE 2. Linear trend in NAO_{DJF} indices and predictors in standardized units per decade (columns 3, 4, 6, and 8). The percentage of NAO_{DJF} hindcast correlation skill attributable to linear trend (ΔSkill, columns 5, 7, and 9). See text for calculation of ΔSkill.

Assessment period	NAO_{DJF} index	NAO_{DJF} trend	May SST (SVD)		JJASO SST (PC2)		JJ T_{SP}	
			Trend	ΔSkill	Trend	ΔSkill	Trend	ΔSkill
1900–2001	CRU NAO_{DJF}	−0.07	0.03	−4	0.07	−111	0.12	−200
	Hurrell NAO_{DJF}	−0.02	—	−4	—	−24	—	−26
	$Z500_{DJF}$	−0.05	—	−4	—	−70	—	−100
1950–2001	CRU NAO_{DJF}	0.21	0.33	−40	−0.15	19	0.31	14
	Hurrell NAO_{DJF}	0.21	—	−29	—	23	—	21
	$Z500_{DJF}$	0.27	—	−50	—	46	—	19
1972–2001	CRU NAO_{DJF}	0.23	−0.52	41	0.18	−40	0.29	−2
	Hurrell NAO_{DJF}	0.13	—	40	—	−17	—	−4
	$Z500_{DJF}$	0.23	—	60	—	−7	—	0

components (PCs) are also subjected to block elimination in their formulation. For each predictor, a time series of hindcasted NAO_{DJF} values is calculated whose skill is verified against the corresponding observed NAO_{DJF} series.

e. Skill assessment

In validating the hindcast skill from each predictor, we employ the Pearson product-moment correlation skill (r_s) and the mean-squared skill score (MSSS) against a simple prediction of climatology. Climatology in this study is the long-term mean for each assessment period because the NAO_{DJF} is known to be nonstationary during the twentieth century (Hurrell and van Loon 1997). This nonstationarity means that no single 30-yr climatology period is representative of the twentieth century as a whole. However, our conclusions are not sensitive to the choice of climatology.

The statistical significance of each skill value is estimated using a randomized resampling method (Manly 1997). Our null hypothesis (H_0) is that there is no link between a given predictor and the NAO_{DJF} and therefore any observed skill is achieved by random chance. The observed and hindcast NAO_{DJF} time series are both randomly shuffled (with replacement) to create two new synthetic time series drawn from the same populations as the original time series. The length of the two synthetic time series is reduced to equal the "effective" number of degrees of freedom between the original time series, taking into account serial correlation (Davis 1976). Skill values are calculated for the two randomized time series and this shuffling/resampling process is repeated in Monte Carlo fashion 25 000 times. The statistical significance is determined by the number of cases where the skill value from the random data exceeds the original observed skill value. The a priori threshold of statistical significance is set at 0.05, which represents the probability that H_0 is incorrectly rejected. Our test is one tailed as only positive values of r_s and MSSS will lead to a rejection of H_0 (Wilks 1995).

3. NAO hindcast skill

a. Comparison by assessment period

Table 3 shows the NAO_{DJF} cross-validated hindcast skill (and significance) for each predictor against the three NAO_{DJF} indices over all three assessment time periods. For the 100-yr period, significant hindcast skill is found using May SST (SVD) and both T_{SP} index predictors. The long summer [May–September (MJJAS)] mean produces the highest skill (a 6%–9% improvement over a prediction of climatology) and the JJ mean results in a 2%–5% improvement. Values of r_s range from 0.26 to 0.31 for MJJAS, which represent 7%–10% of NAO_{DJF} variance over 100 yr. The JJASO SST (PC2) predictors shows positive, but not significant, skill over this period. There is little variability in predictive skill by predictor with NAO index.

For the 50-yr assessment period, statistically significant skill is achieved using the two SST predictors and both T_{SP} indices. However, the May SST (SVD) predictor is significant only when calculated using NCEP–NCAR SST and not using HadISST. In contrast to the 1900–2001 period, it is the JJ T_{SP} rather than the MJJAS T_{SP} that is most skillful for all three NAO indices. The former has maximum $r_s = 0.48$ and MSSS = 22%, whereas for the MJJAS index the maximum skill values are $r_s = 0.41$ and MSSS = 16%. The CRUTEM2 T_{SP} indices show skill 20%–30% lower in this assessment period than the NCEP–NCAR T_{SP} indices. An ensemble mean T_{SP} index is calculated using the mean of the CRUTEM2 and NCEP–NCAR data. This produces skill levels close to those achieved using NCEP–NCAR alone. In terms of skill variation with NAO_{DJF} index, all predictors except May SST (SVD) perform best against $Z500_{DJF}$ for this assessment period.

Highest hindcast skill in the 1972–2001 period is achieved using the JJ T_{SP} index and the two snow cover predictors. The JJ T_{SP} index shows the highest skill values (maximum $r_s = 0.59$ and $MSSS = 35\%$) this skill showing little sensitivity to NAO_{DJF} index or choice of temperature dataset. The JJ NH index is slightly the more skillful of the two snow cover predictors (maximum $r_s = 0.53$ and MSSS = 28%) against all three NAO_{DJF} indices. As in the 50-yr assessment, the link between T_{SP} and the NAO_{DJF} indices is strongest for the JJ period rather than for the long summer (MJJAS) period. The combined CRUTEM2–NCEP–NCAR mean JJ T_{SP} index shows skill slightly higher than that achieved with each single dataset JJ T_{SP} index. For MJJAS T_{SP}, skill is similar to that achieved using single dataset indices for CRU and Hurrell NAO_{DJF} but higher against $Z500_{DJF}$. The A–O mean Eurasian snow cover produces significant skill against all three NAO indices, although it displays sensitivity to NAO_{DJF} index. Neither of the SST-based predictors produce significant skill for the 30-yr assessment period (not shown).

b. Stationarity

The differences observed in NAO_{DJF} predictive skill from summer T_{SP} between the 1900–2001 assessment period and the 1972–2001 period suggest that the relationship between T_{SP} and the NAO_{DJF} is nonstationary.

TABLE 3. Cross-validated hindcast skill for three assessment periods and three NAO_{DJF} indices. Dataset abbreviations: Had = HadISST, CRUT = CRUTEM2, NCEP = NCEP–NCAR reanalysis, NCEP–CRUT = combined CRUTEM2–NCEP–NCAR mean, Rut = Rutgers. Here r is the Pearson correlation skill value and *MSSS* is the percentage improvement in mean-squared error over climatology. Here p is the probability (values in brackets) that the observed skill was obtained by random chance. Boldface type denotes significant ($p < 0.05$) skill values as determined by a Monte Carlo resampling test.

			CRU NAO_{DJF}				Hurrell NAO_{DJF}				$Z500_{DJF}$			
Period	Lagged predictor	Dataset	r	(p)	MSSS	(p)	r	(p)	MSSS	(p)	r	(p)	MSSS	(p)
	May SST (SVD)*	Had	**0.23**	**(0.03)**	**4**	**(0.03)**	**0.27**	**(0.01)**	**6**	**(<0.01)**	**0.22**	**(0.04)**	**5**	**(0.03)**
	JJASO SST (PC2)	Had	0.13	(0.11)	1	(0.08)	0.15	(0.09)	1	(0.07)	0.11	(0.14)	0	(—)
1900–2001	JJ T_{SP}	CRUT	**0.18**	**(0.05)**	**2**	**(0.04)**	**0.24**	**(0.02)**	**5**	**(0.02)**	**0.20**	**(0.04)**	**4**	**(0.05)**
	MJJAS T_{SP}	CRUT	**0.26**	**(0.01)**	**6**	**(0.01)**	**0.31**	**(<0.01)**	**9**	**(<0.01)**	**0.26**	**(0.01)**	**6**	**(0.01)**
	May SST (SVD)	Had	0.15	(0.15)	0	(—)	0.22	(0.07)	3	(0.08)	0.17	(0.11)	1	(0.16)
	May SST (SVD)	NCEP	**0.24**	**(0.05)**	4	(0.07)	**0.31**	**(0.02)**	**8**	**(0.03)**	**0.29**	**(0.04)**	**8**	**(0.04)**
	JJASO SST (PC2)	Had	**0.24**	**(0.05)**	**5**	**(0.03)**	**0.25**	**(0.05)**	**6**	**(0.03)**	**0.38**	**(0.01)**	**14**	**(0.01)**
	JJASO SST (PC2)	NCEP	**0.24**	**(0.05)**	**5**	**(0.03)**	**0.24**	**(0.05)**	**5**	**(0.03)**	**0.26**	**(0.03)**	**5**	**(0.04)**
	JJ T_{SP}	CRUT	**0.33**	**(0.02)**	**10**	**(0.01)**	**0.27**	**(0.05)**	**6**	**(0.03)**	**0.36**	**(0.02)**	**13**	**(0.02)**
1950–2001	JJ T_{SP}	NCEP	**0.46**	**(<0.01)**	**21**	**(<0.01)**	**0.36**	**(0.02)**	**12**	**(0.01)**	**0.48**	**(<0.01)**	**22**	**(<0.01)**
	JJ T_{SP}	NCEP–CRUT	**0.44**	**(<0.01)**	**19**	**(<0.01)**	**0.36**	**(0.02)**	**12**	**(0.01)**	**0.45**	**(<0.01)**	**20**	**(<0.01)**
	MJJAS T_{SP}	CRUT	0.21	(0.10)	4	(0.07)	0.27	(0.06)	**7**	**(0.04)**	**0.32**	**(0.03)**	**10**	**(0.03)**
	MJJAS T_{SP}	NCEP	**0.33**	**(0.02)**	**10**	**(0.01)**	**0.32**	**(0.04)**	**9**	**(0.02)**	**0.41**	**(0.01)**	**16**	**(0.01)**
	MJJAS T_{SP}	NCEP–CRUT	**0.29**	**(0.04)**	**8**	**(0.03)**	**0.33**	**(0.03)**	**10**	**(0.02)**	**0.42**	**(0.01)**	**17**	**(0.01)**
	JJ T_{SP}	CRUT	**0.54**	**(<0.01)**	**29**	**(<0.01)**	**0.57**	**(<0.01)**	**32**	**(<0.01)**	**0.34**	**(0.04)**	10	0.06
	JJ T_{SP}	NCEP	**0.58**	**(<0.01)**	**34**	**(<0.01)**	**0.56**	**(<0.01)**	**31**	**(<0.01)**	**0.55**	**(<0.01)**	**30**	**(<0.01)**
	JJ T_{SP}	NCEP–CRUT	**0.59**	**(<0.01)**	**35**	**(<0.01)**	**0.57**	**(<0.01)**	**33**	**(<0.01)**	**0.55**	**(0.01)**	**30**	**(<0.01)**
	MJJAS T_{SP}	CRUT	**0.45**	**(0.03)**	**20**	**(0.02)**	**0.49**	**(0.02)**	**23**	**(0.01)**	0.11	(0.24)	**0**	(—)
1972–2001	MJJAS T_{SP}	NCEP	**0.51**	**(0.01)**	**26**	**(<0.01)**	**0.47**	**(0.01)**	**22**	**(0.01)**	**0.40**	**(0.01)**	**16**	**(0.03)**
	MJJAS T_{SP}	NCEP–CRUT	**0.48**	**(0.02)**	**22**	**(0.01)**	**0.47**	**(0.02)**	**22**	**(0.01)**	**0.47**	**(0.02)**	**22**	**(0.01)**
	A–0 EU snow cover	Rut	**0.47**	**(0.01)**	**21**	**(0.01)**	**0.36**	**(0.02)**	**11**	**(0.03)**	**0.43**	**(0.02)**	**17**	**(0.03)**
	JJ NH snow cover	Rut	**0.53**	**(0.01)**	**28**	**(<0.01)**	**0.51**	**(0.01)**	**24**	**(<0.01)**	**0.48**	**(0.01)**	**22**	**(0.01)**

* Data available 1900–98.

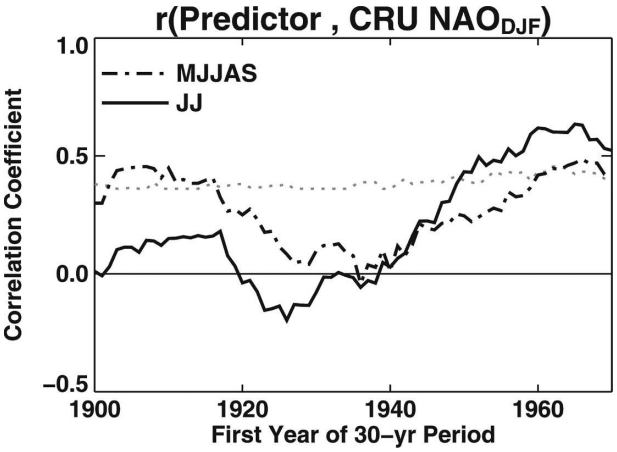

FIG. 4. Correlation coefficients between the JJ and MJJAS T_{SP} indices and the upcoming CRU NAO_{DJF} index for running 30-yr windows commencing 1900–71. Faint dashed line indicates 5% significance level corrected for serial correlation.

Therefore, the modest skill seen over the full 100-yr could be drawn entirely from the 1972–2001 skillful period. To test this we first perform a rolling cross-correlation analysis for both the JJ and MJJAS T_{SP} indices against the CRU NAO_{DJF} index for all possible 30-yr periods 1900–2001 (Fig. 4). A linear trend is removed from both time series prior to each correlation value being computed. The MJJAS T_{SP} index peaks at the beginning and end of the twentieth century. By contrast, JJ T_{SP} is most skillful since 1950. Further investigation reveals that it is the August–September period that contributes to the early 1900s peak in MJJAS (not shown). The strengthening correlations since 1950 may be linked to an increase in decadal NAO_{DJF} variability over the same period (Hurrell and van Loon 1997).

Despite the coincidence between periods of high correlation and periods of strongly positive NAO index in the early and late twentieth century (Fig. 1), the T_{SP} indices perform equally as well predicting either above- or below-median NAO_{DJF} events. Hindcasts using MJJAS T_{SP} are correctly above and below median in 54% of observed above- and below-median CRU NAO_{DJF} events 1900–2001. For the period 1950–2001 the JJ T_{SP} index is correct in 64% of cases, while for 1972–2001 the figure is 71%. We therefore conclude that the hindcast skill from the summer T_{SP} indices is nonstationary and that highest positive skill is observed during the early and late twentieth century.

c. Dataset dependence

Table 3 shows that, with the exception of May SST (SVD), the NAO_{DJF} hindcast skill exhibits little sensitivity to the choice of SST or 2-m air temperature dataset. Specifically, the JJASO SST (PC2) predictor produces similar skill against the three NAO_{DJF} indices whether it is calculated using HadISST or NCEP–NCAR data 1950–2001. The largest skill sensitivity to dataset is observed for the T_{SP} indices calculated using CRUTEM2 and NCEP–NCAR data 1950–2001. Using CRUTEM2 data, the JJ T_{SP} predictor explains 11% of the variance in the CRU NAO_{DJF} index compared with 21% using NCEP–NCAR data. A possible explanation for this difference is that, unlike the reanalysis data, the CRUTEM2 data are missing over the open ocean. The SG T_{SP} region is more than 50% over ocean and has double the weighting of the EU and NA regions. Therefore, SG is likely to contribute the largest errors in T_{SP}. However, only ~50% of JJ T_{SP} skill comes from SG, which explains why CRUTEM2 produces some positive skill.

4. Summer T_{SP} influence on upcoming NAO_{DJF}

a. Relationship between NAO_{DJF} lagged predictors

Quantifying the links between the different NAO_{DJF} lagged predictors will help to explore the physical basis for the observed lagged predictability. Cross correlations are performed for all predictors over the three assessment periods (not shown). For all periods, stable and significant ($p < 0.05$) relationships are found between the summer T_{SP} indices and the summer/autumn (JJASO) SST (PC2) predictor. Since 1972, similar correlations are observed between T_{SP}, JJASO SST (PC2) and the two warm season snow cover predictors. These results imply a potential physical relationship between the predictors both in time (summer–autumn) and in space (North America, Eurasia, to the North Atlantic sector).

The NAO_{DJF} exhibits persistence from one winter to the next (Johansson et al. 1998; Wanner et al. 2001). We therefore ask how much of the observed NAO_{DJF} hindcast skill results from the influence of the previous NAO_{DJF} on the predictors? Table 4 shows that only the Hurrell NAO_{DJF} 1900–2001 and $Z500_{DJF}$ 1950–2001 have a significant (detrended) lag-1 autocorrelation. The table also shows how much hindcast predictability the lag-1 NAO_{DJF} offers for predicting the May SST (SVD), JJASO SST (PC2), and JJ T_{SP} time series. For the period 1900–2001, the skill values from the lag-1 NAO_{DJF} are comparable to those achieved using the predictors themselves to hindcast NAO_{DJF} (Table 3). For 1950–2001 and 1972–2001, with the exception of May SST (SVD), the predictors provide significantly higher NAO_{DJF} skill than does lag-1 NAO_{DJF}, based on all years of data. However, when composited on JJ T_{SP}

TABLE 4. Lag-1 autocorrelation (r[lag-1]) of NAO_{DJF} indices and cross-validated skill for predicting May SST (SVD), JJASO SST (PC2), and JJ T_{SP} from the prior NAO_{DJF}. Notations are the same as in Table 3. SST data are from HadISST and T_{SP} data are from CRUTEM2.

Period	Lagged predictor	NAO		May SST (SVD)			JJASO SST (PC2)			JJ T_{SP}			
		r[lag-1]	(p)	r	(p)	MSSS	r	(p)	MSSS	r	(p)	MSSS	(p)
1900–2001	CRU NAO_{DJF}	0.13	(0.20)	**0.36**	**(<0.01)**	**12**	0.00	**(—)**	0	**0.27**	**(0.01)**	**6**	**(<0.01)**
	Hurrell NAO_{DJF}	**0.24**	**(0.02)**	**0.38**	**(<0.01)**	**14**	0.15	(0.06)	0	**0.31**	**(<0.01)**	**9**	**(<0.01)**
	$Z500_{DJF}$*	0.19	(0.06)	**0.47**	**(<0.01)**	**17**	0.15	(0.54)	0	**0.35**	**(0.03)**	**8**	**(<0.01)**
1950–2001	CRU NAO_{DJF}	0.14	(0.32)	**0.43**	**(0.01)**	**8**	0.00	(—)	0	0.00	(—)	0	(—)
	Hurrell NAO_{DJF}	0.24	(0.10)	**0.50**	**(<0.01)**	**16**	(<0.01)	**(0.03)**	**5**	**(0.05)**	0.14	(0.14)	0	(—)
	$Z500_{DJF}$	**0.29**	**(0.05)**	**0.45**	**(0.02)**	**9**	**(<0.01)**	**(0.05)**	**3**	**(0.07)**	**0.29**	**(0.02)**	**7**	**(0.03)**
1972–2001	CRU NAO_{DJF}	0.18	(0.36)	**0.60**	**(0.01)**	**18**	**(<0.01)**	(—)	0	(—)	0.00	(—)	0	(—)
	Hurrell NAO_{DJF}	0.27	(0.17)	0.46	0.10)	0	(—)	0.01	0	(—)	0.26	(0.06)	1	(0.08)
	$Z500_{DJF}$	0.18	(0.38)	0.47	(0.19)	0	(—)	0.00	0	(—)	0.14	(0.22)	0	(—)

* UKMO HadSLP data 1900–98.

terciles, the lag-1 NAO_{DJF} skill for JJ T_{SP} 1950–2001 increases to $r_s \sim 0.5$. This suggests that when JJ T_{SP} is large, it responds more to the NAO_{DJF} signal from the previous winter. However, JJ T_{SP} is a better NAO_{DJF} predictor than lag-1 NAO_{DJF} alone (Tables 3 and 4). This implies that the signal from the previous winter intensifies during the summer.

The results in Table 4 are similar using data with trends included (not shown). However, some sensitivity to the choice of dataset was found for the May SST (SVD) predictor 1950–2001. Using NCEP–NCAR reanalysis SST data, the lag-1 NAO_{DJF} skill drops to $r_s \sim 0.3$ ($p > 0.10$) for both 1950–2001 and 1972–2001. The JJASO SST (PC2) predictor does not display such sensitivity, which suggests that greater differences exist for the May SST data than for the JJASO SST fields when comparing the NCEP–NCAR and HadISST datasets.

b. *Physical basis for summer T_{SP} influence on upcoming winter NAO_{DJF}*

Seasonal atmospheric predictability in the extratropics is normally assumed to be lower than for the Tropics because of the presence of baroclinic instability. This appears to contradict the links we observe between summer T_{SP}, snow cover, and the upcoming winter climate. There is little persistence intrinsic to the extratropical atmosphere, which has a decorrelation time of ~10 days. Therefore, to enable a summer–winter link the memory of the summer atmospheric state must persist in a slowly varying boundary variable that can feed back onto the atmosphere at a later time.

Following Saunders et al. (2003), we propose the following mechanism for the influence of summer climate on that of the upcoming winter: summer (JJ) NH snow cover anomalies are negatively correlated with JJ subpolar near-surface air temperature (T2m) over northern Siberia and northwest Canada and positively correlated with JJ T2m over Southern Greenland. The resulting subpolar temperature pattern induces a contemporaneous atmospheric response and lagged SST response centered on the North Atlantic. The atmospheric response is characterized by anomalies in MSLP and midlatitude zonal wind. The SST response is characterized by basin-scale anomalies and meridional gradients south and east of Newfoundland. Persistence of these SST gradients into autumn or early winter could force either a direct thermal NAO_{DJF} response or initiate a response via a third variable.

There is good observational evidence to support the link between snow cover, T_{SP} and North Atlantic SSTs. First, there is a significant correlation ($p < 0.01$) between changes in JJ NH snow cover and JJ T_{SP} for the

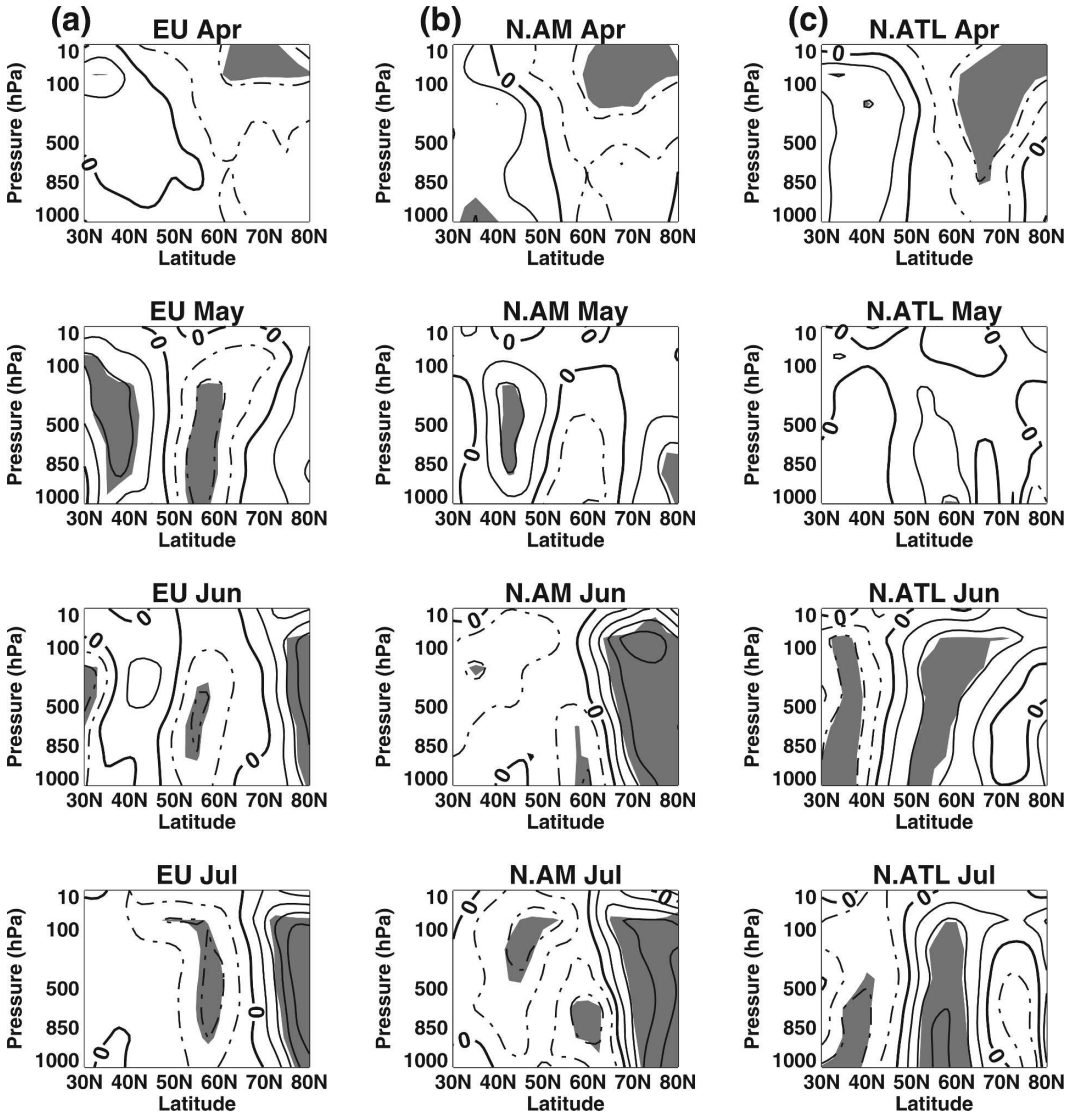

FIG. 5. Composite vertical cross section of zonally averaged zonal wind anomalies based on high minus low terciles of JJ T_{SP} index 1950–2001. Zonal averages are calculated over (a) Eurasia, 25°–70°E, (b) North America, 120°–90°W, and (c) the North Atlantic, 50°–20°W. Data are dimensionless standardized anomalies. Contour interval is 0.3 and dashed contours denote negative anomalies. Shaded areas denote significance ($p < 0.05$) as determined by a Monte Carlo resampling test.

period 1972–2001 (Saunders et al. 2003). Second, the atmosphere adjacent to the T_{SP} centers is significantly perturbed during June and July. Figure 5 shows zonally averaged zonal wind anomalies as a function of height above the T_{SP} regions and the North Atlantic in months before and during high minus low JJ T_{SP} tercile years 1950–2001. Prior to June and July the only consistent signal for all three regions is located in the stratosphere at high latitudes during April. The signal in June and July over Eurasia and North America is characteristic of a strengthened polar vortex and a weakened midlatitude jet. A corresponding teleconnected signal is seen over the North Atlantic in June and July, which extends from the surface to the lower stratosphere. A band of significant anomalous westerlies is centered between 50° and 60°N, with anomalous easterlies to the north and south. This highlights that during extreme years of T_{SP}, contemporaneous subpolar atmospheric teleconnections are formed between northern Eurasia, northern Canada, and the North Atlantic during June and July.

Our proposed mechanism is further supported by Fig. 6, which shows the hemispheric-scale anomaly patterns in height-averaged (925–200 hPa) zonal wind,

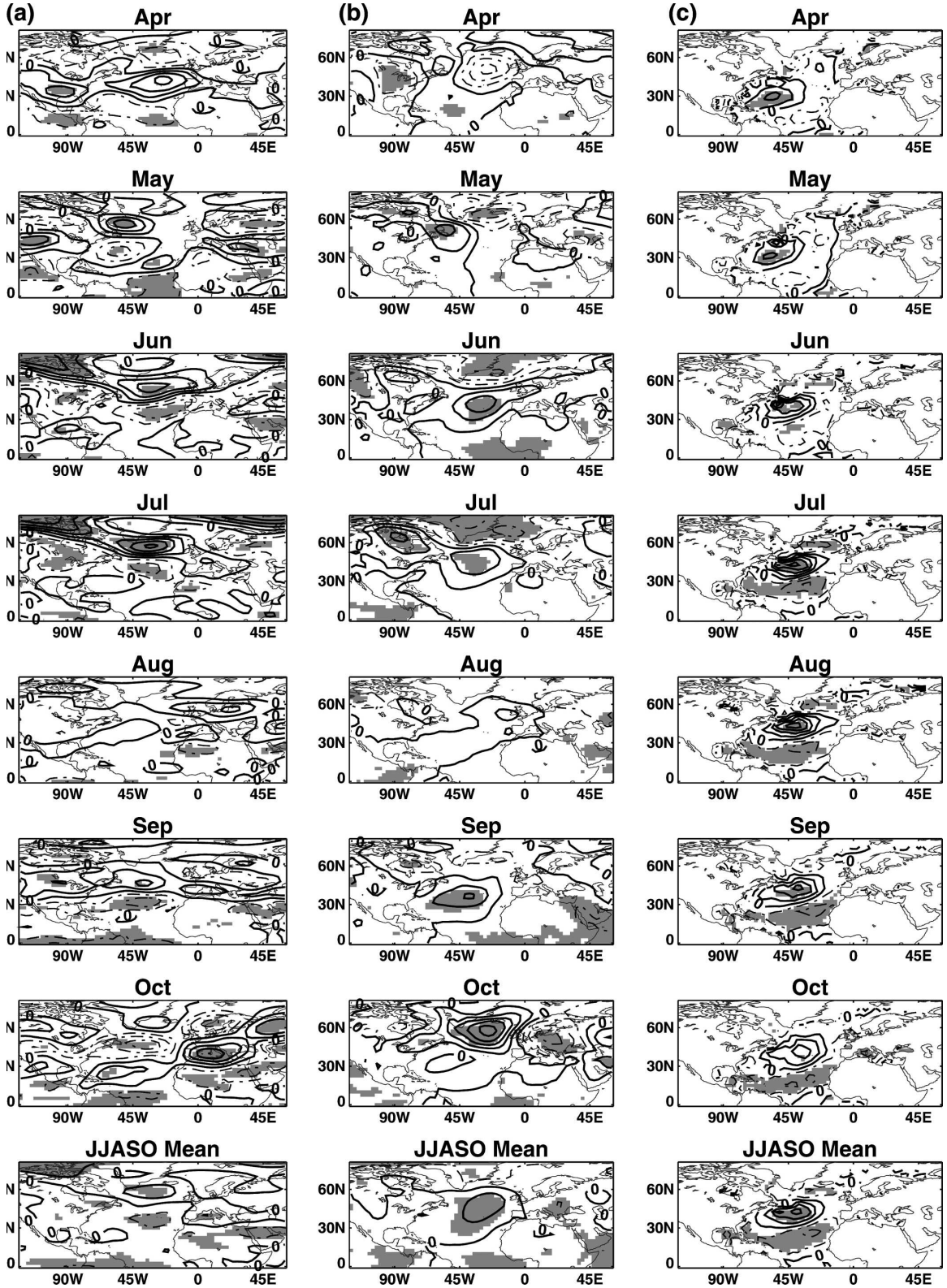

FIG. 6. Composite of (a) vertically averaged (925–200 hPa) North Atlantic sector zonal wind anomalies, (b) North Atlantic sector MSLP anomalies and (c) North Atlantic SST anomalies based on high minus low terciles of JJ T_{SP} index. Contour intervals are (a) 1 m s^{-1}, (b) 1 hPa, and (c) 0.2°C. Dashed contours denote negative anomalies and the zero contour is labeled. Shaded areas denote significance ($p < 0.05$) as determined by a Monte Carlo resampling test.

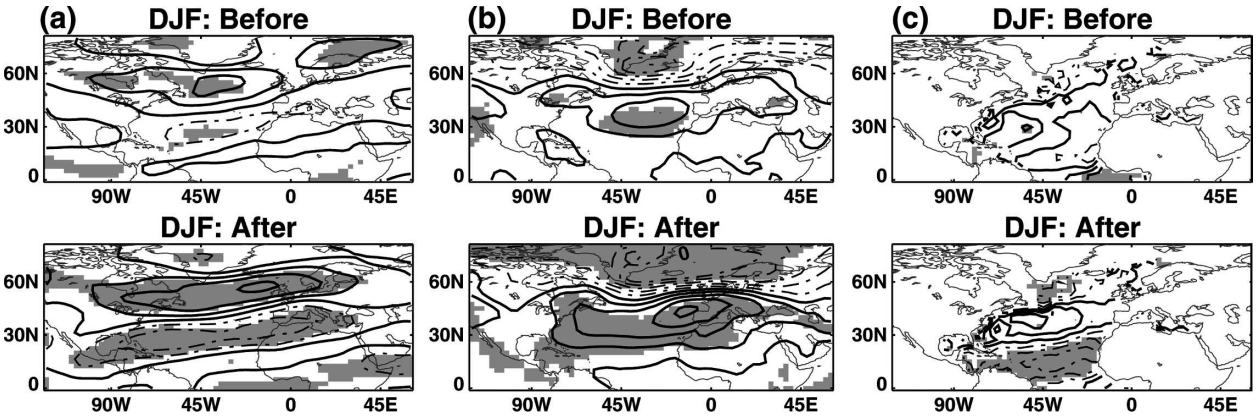

FIG. 7. Same as in Fig. 6 except for winter (DJF) seasonal means (top) preceding and (bottom) following high minus low JJ T_{SP} terciles.

MSLP, and SST in months before, during, and following high-minus-low JJ T_{SP} tercile years 1950–2001. Over the North Atlantic, the atmospheric signal strengthens in June and the SST strengthens after July. The zonal wind, MSLP, and SST signals appear causally connected, with the zonal wind anomalies collocated with gradients in MSLP anomalies and with SST anomalies. The atmospheric signal is prominent in June and July but dissipates thereafter, although it is apparent in the JJASO mean. In contrast, the SST anomaly pattern strengthens in July and persists until October. The timing, lag, and spatial pattern of the SST signal is consistent with its forcing by the anomalous atmospheric circulation associated with JJ T_{SP}. This conclusion is further supported by the significant ($p < 0.05$ for all time periods) link between JJ T_{SP} and JJASO SST (PC2).

Figure 7 shows the mean anomalies in vertically averaged zonal wind, MSLP and SST for the winter seasons before and following high-minus-low JJ T_{SP} terciles. The situation preceding JJ T_{SP} shows a moderate positive NAO pattern in the zonal wind anomalies, which is further evidence that a proportion of the NAO_{DJF} predictability from JJ T_{SP} arises through interannual NAO persistence (see also section 4a). However, SSTs are close to climatology except near the equator, showing that the involvement of SST does not begin until after summer. The SST pattern following extreme JJ T_{SP} events exhibits a strong meridional gradient off Newfoundland, which has persisted from summer. This pattern resembles the NAO-driven North Atlantic SST tripole anomaly pattern (e.g., Czaja and Frankignoul 2002).

Next, we attempt to show whether the evolving anomalies seen in Fig. 6 indicate a real response to JJ T_{SP} or merely represent increased climatological variance in those months. This is quantified by plotting the percentage of grid cells in each month that are locally significant ($p < 0.05$). Figure 8 shows the percentages for zonal wind over the region (30°–80°N, 120°W–40°E) and for SST over the region (0°–65°N, 100°W–0°). The horizontal lines denote approximate levels of field significance ($p < 0.05$) following the method of Livezey and Chen (1983). For each month, we assume 30 spatial degrees of freedom (DOF) in the zonal wind field and 15 DOF in the SST field. These values are appropriate because Livezey and Chen (1983) state that hemispheric atmospheric fields contain ~30–60 DOF. The actual DOF in our fields may oscillate around these estimates depending on the month. However, even with a conservative estimate for DOF of 15, the zonal wind and SST maxima are both field significant.

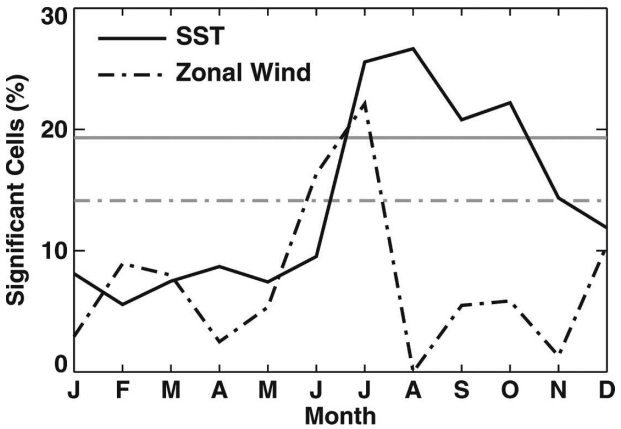

FIG. 8. Percentage of locally significant grid cells in Fig. 6 for SST (solid line) and vertically averaged (925–200 hPa) zonal wind (dashed line) associated with high minus low terciles of JJ T_{SP} index. Domains used are 0°–65°N, 90°W–0° for SST and 30°–80°N, 120°W–40°E for zonal wind. Horizontal lines denote approximate field significance assuming 15 spatial degrees of freedom for SST and 30 for zonal wind.

The zonal wind signal peaks in June and July, consistent with Fig. 6. The SST response develops 1 month later, peaking in July and August, and persisting until October. The timing of the lagged SST response is consistent with other observational studies of atmosphere–ocean interaction (e.g., Deser and Timlin 1997). By October, the majority of the SST signal comes from the subtropics. The SST response decreases to December, which is associated with a more variable atmospheric pattern August–October. We therefore conclude that the anomalous atmospheric circulation associated with JJ T_{SP} is leading North Atlantic SST variability with a time lag of ~1 month.

5. Discussion

NAO_{DJF} hindcast skill is nonstationary during the twentieth century, implying that the predictive relationships vary in time. Data quality may also contribute to the variations in NAO_{DJF} skill. During the early part of the record, observations are more sparse than during the 1972–2001 period of highest skill. This reduction in data quality may degrade NAO_{DJF} hindcast skill and contribute to nonstationarity.

Nonstationarity in predictive relationships may arise from low-frequency multiannual or decadal oscillations in the predictor and/or predictand indices. An analysis of spectral coherency (Bloomfield 2000) was performed on the predictors and NAO indices to determine the dominant time scales of interaction. The results (not displayed) show that the preferred oscillatory period for each predictor with the NAO_{DJF} differs depending on the assessment period. For example, the MJJAS T_{SP} index and CRU NAO_{DJF} index are significantly coherent at a period of 7 yr (1900–2001), 4 yr (1950–2001), and 8 yr (1972–2001). A further cause of the nonstationarity may be competing influences from SST forcings outside the North Atlantic, in particular, variations in the strength of ENSO (Sutton and Hodson 2003).

The May SST (SVD) predictor shows significant skill with HadISST data 1900–2001 and with NCEP–NCAR SST data 1950–2001. However, no significant skill is seen with HadISST over the latter period. Furthermore, our reported skill levels are lower than those found by Rodwell and Folland (2002), who employed an earlier version of HadISST 1948–98. Their original skill ($r_s = 0.45$) is achieved using data with trends included, which suggests that 30%–60% of that skill comes from linear trend 1950–2001 (cf. Table 1). Furthermore, this original skill is affected by serial correlation in the $Z500_{DJF}$ index. Using data with trends included, the lag-1 autocorrelation of the $Z500_{DJF}$ index rises from 0.29 to 0.42. We also note that despite optimizing covariance against the 500-hPa DJF North Atlantic geopotential height field, May SST (SVD) produces highest skill against the Hurrell NAO_{DJF} index. This results because the $Z500_{DJF}$, CRU, and Hurrell NAO_{DJF} indices cross correlate at $r > 0.85$ (1950–2001).

Significant SST persistence associated with JJ T_{SP} is shown to occur from summer into October. The persistent SST signal located southeast of Newfoundland in Fig. 6 lies adjacent to the main region of North Atlantic cyclogenesis, which marks the beginning of the North Atlantic storm track. Anomalous meridional SST gradients and associated turbulent heat, moisture, and momentum fluxes in this region could therefore initiate a direct thermal response in NAO_{DJF} (Rodwell et al. 1999; Peng and Whittaker 1999; Peng et al. 2003; Cassou et al. 2004). However, the weakened SST signal from October to December (Fig. 8) suggests that the T_{SP}-induced SST anomalies alone may not be sufficient to force subsequent changes in NAO_{DJF}. One possibility is that autumn SST anomalies could affect the upward propagation of anomalous wave activity into the subpolar stratosphere. This would cause a downward-propagating response during early winter, which could influence the surface (Polvani and Waugh 2004). However, typically an AO, rather than an NAO_{DJF}, signal is observed in the winter stratosphere. Another theory is that autumn SST could influence either local sea ice outflow (Hilmer and Jung 2000; Kvamsto et al. 2004) and/or remote snow cover extent (Kitaev et al. 2002; Robock et al. 2003) through contemporaneous atmospheric circulation anomalies. These variables have sufficiently long time scales to feed back onto the atmosphere during the subsequent winter. An alternative possibility is that anomalous atmospheric flow could be the lone cause of summer NH snow cover anomalies, T_{SP}, and subsequent North Atlantic SST anomalies.

6. Summary and conclusions

In this study, a detailed assessment of the current levels of empirical seasonal hindcast skill for the North Atlantic Oscillation (NAO) is presented. A standardized hindcast procedure is employed to validate four previously published lagged predictors of the upcoming winter NAO. An additional predictor based on summer Northern Hemisphere subpolar 2-m air temperature (T_{SP}) is examined. Over three extended assessment periods out to 100-yr summer T_{SP} is most skillful in predicting the upcoming winter NAO. For the period 1900–2001, May–September mean T_{SP} offers the highest skill (a ~6%–9% improvement over climatology). Since 1950, June–July mean T_{SP} produces the highest skill (up to 22% improvement). Significant skill is also

observed 1900–2001 and 1950–2001 using patterns of May and late summer/autumn North Atlantic SST. Two warm season snow cover predictors also produce significant skill 1972–2001. Twentieth-century NAO_{DJF} hindcast skill is nonstationary and highest since 1972 for all predictors except those derived using SSTs. This coincides with a period of large decadal NAO_{DJF} variability. The highest skill from T_{SP} also coincides with positive trends in the NAO_{DJF} index during the early and late twentieth century. However, T_{SP} performs equally well predicting above- or below-median NAO_{DJF} events.

Evidence is presented supporting a physical link between summer T_{SP} and the upcoming winter NAO. First, there is a significant contemporaneous association during summer between T_{SP} and Northern Hemisphere snow cover extent. During subsequent months, the atmospheric response to T_{SP} is centered on the midlatitude North Atlantic. Circulation anomalies over the ocean are associated with a persistent pattern of North Atlantic SST into October. SST persistence is strong off Newfoundland, which coincides with the main region of North Atlantic cyclogenesis. This suggests that T_{SP} either produces a direct NAO_{DJF} response to local SST gradients or initiates a feedback onto NAO_{DJF} from a third variable.

Further investigation using coupled dynamical models is required to fully understand the link between summer and winter North Atlantic climate. The large repository of global historical temperature observations allow for more extended analyses of the T_{SP}/NAO_{DJF} mechanism than is possible with seasonal NAO_{DJF} predictive links involving shorter time series (e.g., snow cover). Future studies must also examine whether the nonstationarity in this link is explained by observed variations in twentieth-century climate. Experiments using coupled GCMs with realistic snow–atmosphere interaction and subpolar teleconnections are required to further investigate the dynamics of how snow cover, T_{SP}, and SST link to NAO_{DJF}.

Our findings suggest that the NH subpolar regions may provide extended-range NAO_{DJF} seasonal predictability in addition to the midlatitudes or the Tropics. This contrasts with recent thinking based on atmospheric GCM experiments, which indicate that variations in tropical SSTs are of primary importance for explaining the NAO_{DJF} trend 1950–2000 (Hurrell et al. 2004). We believe that these results offer exciting new possibilities for the future of seasonal forecasting in the extratropics.

Acknowledgments. CGF was supported by a Research Studentship from the U.K. Natural Environment Research Council. We thank B. Lloyd-Hughes for helpful discussions. We acknowledge NOAA–CIRES Climate Diagnostics Center, Boulder, Colorado, for NCEP–NCAR global reanalysis project data; the Snow Data Resource Center at Rutgers University for snow extent records; the Met Office Hadley Centre for HadISST and HadSLP data; the Climatic Research Unit, University of East Anglia, for CRUTEM2 and NAO data; and the Climate Analysis Section, NCAR, Boulder, Colorado, for Hurrell (1995) NAO data.

REFERENCES

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

Basnett, T. A., and D. E. Parker, 1997: Development of the Global Mean Sea Level Pressure data set GMSLP2. Climate Research Tech. Note 79, Hadley Centre, Met Office, Exeter, Devon, United Kingdom, 16 pp. + appendixes.

Bloomfield, P., 2000: *Fourier Analysis of Time Series*. John Wiley and Sons, 288 pp.

Bojariu, R., and L. Gimeno, 2003: The role of snow cover fluctuations in multiannual NAO persistence. *Geophys. Res. Lett.,* **30,** 1156, doi:10.1029/2002GL015651.

Cassou, C., C. Deser, L. Terray, J. W. Hurrell, and M. Drévillon, 2004: Summer sea surface temperature conditions in the North Atlantic and their impact upon the atmospheric circulation in early winter. *J. Climate,* **17,** 3349–3363.

Cohen, J., and D. Entekhabi, 1999: Eurasian snow cover variability and northern hemisphere climate predictability. *Geophys. Res. Lett.,* **26,** 345–348.

Czaja, A., and C. Frankignoul, 2002: Observed impact of Atlantic SST anomalies on the North Atlantic Oscillation. *J. Climate,* **15,** 606–623.

Davis, R. E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.,* **6,** 249–266.

Deser, C., and M. S. Timlin, 1997: Atmosphere–ocean interaction on weekly timescales in the North Atlantic and Pacific. *J. Climate,* **10,** 393–408.

Drévillon, M., L. Terray, P. Rogel, and C. Cassou, 2001: Midlatitude SST influence on European winter climate variability in the NCEP reanalysis. *Climate Dyn.,* **18,** 331–344.

Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol.,* **21,** 1111–1152.

Hilmer, M., and T. Jung, 2000: Evidence for a recent change in the link between the North Atlantic Oscillation and Arctic sea ice export. *Geophys. Res. Lett.,* **27,** 989–992.

Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperature and precipitation. *Science,* **269,** 676–679.

——, and H. van Loon, 1997: Decadal variations in climate associated with the North Atlantic Oscillation. *Climatic Change,* **36,** 301–326.

——, M. P. Hoerling, A. S. Phillips, and T. Xu, 2004: Twentieth century North Atlantic climate change. Part I: Assessing determinism. *Climate Dyn.,* **23,** 371–389.

Johansson, A., A. Barnston, S. Saha, and H. Van den Dool, 1998:

On the level and origin of seasonal forecast skill in northern Europe. *J. Atmos. Sci.,* **55,** 103–127.

Jones, P. D., and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *J. Climate,* **16,** 206–223.

——, T. Jónsson, and D. Wheeler, 1997: Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland. *Int. J. Climatol.,* **17,** 1433–1450.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kitaev, L., A. Kislov, A. Krenke, V. Razuvaev, R. Martunganov, and I. Konstantinov, 2002: The snow cover characteristics of northern Eurasia and their relationship to climatic parameters. *Boreal Environ. Res.,* **7,** 437–445.

Kvamsto, N. G., P. Skeie, and D. B. Stephenson, 2004: Impact of Labrador sea-ice on the North Atlantic Oscillation. *Int. J. Climatol.,* **24,** 603–612.

Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.,* **111,** 46–59.

Manly, B. F. J., 1997: *Randomization, Bootstrap and Monte Carlo Methods in Biology*. 2d ed. Chapman and Hall, 424 pp.

Michaelsen, J., 1987: Cross-validation in statistical climate forecast models. *J. Climate Appl. Meteor.,* **26,** 1589–1600.

Palmer, T. N., and Coauthors, 2004: Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER). *Bull. Amer. Meteor. Soc.,* **85,** 853–872.

Peng, S., and J. S. Whittaker, 1999: Mechanisms determining the atmospheric response to midlatitude SST anomalies. *J. Climate,* **12,** 1393–1408.

——, W. A. Robinson, and S. Li, 2003: Mechanisms for the NAO response to the North Atlantic tripole. *J. Climate,* **16,** 1987–2004.

Polvani, L. M., and D. W. Waugh, 2004: Upward wave activity flux as precursor to extreme stratospheric events and subsequent anomalous surface weather regimes. *J. Climate,* **17,** 3548–3554.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. K. Alexander, and D. P. Rowell, 2003: Global analyses of sea surface temperature, sea ice and night marine air temperature since the late nineteenth century. *J. Geophys. Res.,* **108,** 4407, doi:10.1029/2002JD002670.

Robinson, D. A., K. F. Dewey, and R. R. Heim, 1993: Global snow cover monitoring: An update. *Bull. Amer. Meteor. Soc.,* **74,** 1689–1696.

Robock, A., M. Mu, K. Y. Vinnikov, and D. Robinson, 2003: Land surface conditions over Eurasia and Indian summer monsoon rainfall. *J. Geophys. Res.,* **108,** 4131, doi:10.1029/2002JD002286.

Rodwell, M. J., and C. K. Folland, 2002: Atlantic air-sea interaction and seasonal predictability. *Quart. J. Roy. Meteor. Soc.,* **128,** 1413–1443.

——, D. P. Rowell, and C. K. Folland, 1999: Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature,* **398,** 320–323.

Saunders, M. A., and B. Qian, 2002: Seasonal predictability of the winter NAO from North Atlantic sea surface temperatures. *Geophys. Res. Lett.,* **29,** 2049, doi:10.1029/2002GL014952.

——, ——, and B. Lloyd-Hughes, 2003: Summer snow extent heralding of the winter North Atlantic Oscillation. *Geophys. Res. Lett.,* **30,** 1378, doi:10.1029/2003GL017401.

Sutton, R. T., and D. L. R. Hodson, 2003: Influence of the ocean on North Atlantic climate variability 1871–1999. *J. Climate,* **16,** 3296–3313.

Trigo, R. M., T. J. Osborn, and J. M. Corte-Real, 2002: The North Atlantic Oscillation influence on Europe: Climate impacts and associated physical mechanisms. *Climate Res.,* **20,** 9–17.

Walker, G. T., and E. W. Bliss, 1932: World weather V. *Memo. Roy. Meteor. Soc.,* **4,** 53–84.

Wanner, H., S. Bronnimann, C. Casty, D. Gyalistras, J. Luterbacher, C. Schmultz, D. B. Stephenson, and E. Xoplaki, 2001: North Atlantic Oscillation—Concepts and studies. *Surv. Geophys.,* **22,** 321–382.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 467 pp.