The Predictors and Forecast Skill of Northern Hemisphere Teleconnection Patterns for Lead Times of 3–4 Weeks

JIAXIN BLACK

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, and National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

NATHANIEL C. JOHNSON

Cooperative Institute for Climate Science, Princeton University, Princeton, New Jersey

STEPHEN BAXTER

NOAA/NCEP/Climate Prediction Center, College Park, Maryland

STEVEN B. FELDSTEIN

Department of Meteorology and Atmospheric Science, College of Earth and Mineral Sciences, The Pennsylvania State University, University Park, Pennsylvania

DANIEL S. HARNOS AND MICHELLE L. L'HEUREUX

NOAA/NCEP/Climate Prediction Center, College Park, Maryland

(Manuscript received 18 October 2016, in final form 31 March 2017)

ABSTRACT

The Pacific-North American pattern (PNA), North Atlantic Oscillation (NAO), and Arctic Oscillation (AO) are three dominant teleconnection patterns known to strongly affect December-February surface weather in the Northern Hemisphere. A partial least squares regression (PLSR) method is adopted in this study to generate wintertime two-week statistical forecasts of these three teleconnection pattern indices for lead times of up to five weeks over the 1980-2013 period. The PLSR approach generates forecasts for the teleconnection pattern indices by maximizing the variance explained by predictor indices determined as linear combinations of predictor fields, which include gridded outgoing longwave radiation (OLR), 300-hPa geopotential height (Z300), and 50-hPa geopotential height (Z50). Overall, the PLSR models yield statistically significant skill at all lead times up to five weeks. In particular, cross-validated correlations between the combined weeks 3-4 PLSR forecasts and verification for the PNA, NAO, and AO indices are 0.34, 0.28, and 0.41, respectively. The PLSR approach also allows the authors to isolate a small number of predictor patterns that help shed light on the sources of prediction skill for each teleconnection pattern. As expected, the results reveal the importance of tropical convection (OLR) for forecast skill in weeks 3-4, but the initial atmospheric flow (Z300) accounts for a substantial fraction of the skill as well. Overall, the results of this study provide promise for improving subseasonal-to-seasonal (S2S) forecasts and the physical understanding of predictability on these time scales.

1. Introduction

Subseasonal forecasting for lead times between 10 and 30 days is in general a difficult task but one that has been receiving increasing focus. At longer lead times, the importance of the initial conditions progressively diminishes, and the impact of slowly varying boundary conditions increases, though they have a more modest impact on the weather and climate. From the end-user perspective, however, skillful subseasonal forecasting is of great importance because it would provide a sound basis for extended decision-making in agriculture and food security, water and energy management, and disaster risk reduction that could save lives and protect property.

DOI: 10.1175/MWR-D-16-0394.1

© 2017 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

Corresponding author: Jiaxin Black, jiaxin.black@noaa.gov

With the realization of the importance of subseasonal forecasting, governmental agencies, together with the scientific community, have increasingly invested resources to improve the skill and to promote the utility of subseasonal forecasts in recent years. In 2009, the World Meteorological Organization (WMO) initiated a Subseasonal to Seasonal (S2S) Prediction Project (https:// www.wcrp-climate.org/s2s-overview), and later laid out a detailed research implementation plan. In addition, the U.S. National Academy of Sciences (NAS) recently released a research agenda to achieve a vision that S2S forecasts (forecasts made 2 weeks to 12 months in advance) will become as widely used in a decade as weather forecasts are today. In the United States, President Obama's administration announced a coordinated effort to develop new extreme-weather outlooks in the 15-30-day range, seeking to produce actionable information products for intermediate time scales at which climate change influences risk. As an initial step toward these efforts in the United States, the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) recently began issuing experimental week 3-4 outlooks for precipitation and temperature (http://www.cpc. ncep.noaa.gov/products/predictions/WK34/), which provide probabilities of temperatures and accumulated precipitation being above normal or below normal. With all these active engagements among governmental agencies, the scientific community, and the private sector, subseasonal forecasts are becoming an increasingly important component in climate prediction.

Skillful subseasonal forecasts generally rely on the skillful prediction of the large-scale atmospheric circulation, which is closely tied to large-scale teleconnection patterns. These teleconnection patterns reflect largescale changes in the atmospheric wave and jet stream patterns, and thus have strong impacts on temperature, precipitation, and storm tracks over vast geographical areas. In the Northern Hemisphere, prominent teleconnection patterns that strongly affect the temperature and precipitation, especially during wintertime, include the Pacific-North American pattern (PNA; Horel and Wallace 1981; Wallace and Gutzler 1981; Barnston and Livezey 1987), North Atlantic Oscillation (NAO; Barnston and Livezey 1987), and Arctic Oscillation (AO; Thompson and Wallace 1998). Recently, Scaife et al. (2014) showed that skillful predictions of the NAO in a new forecast system leads to similarly skillful predictions of European winter surface climate. Therefore, successful subseasonal forecasts over large portions of the Northern Hemisphere rely on being able to skillfully predict these dominant teleconnection patterns.

Studies over the past few decades have identified a few potential sources of skill regarding atmospheric teleconnection pattern forecasts that may operate at lead times of more than two weeks. The first source is tropical deep convection, particularly in association with the Madden-Julian oscillation (MJO; Madden and Julian 1971, 1972). The MJO, which has a natural time scale of 30-70 days, has been shown to have a significant impact on the extratropical circulation, especially during wintertime within a time frame of 1-4 weeks, through the excitation of poleward-propagating Rossby waves (e.g., Hoskins and Karoly 1981; Karoly et al. 1989). Specifically, both observations and idealized models have shown that the PNA arises in response to the tropical convection anomalies associated with MJO (Knutson and Weickmann 1987; Ferranti et al. 1990; Higgins and Mo 1997; Mori and Watanabe 2008; Johnson and Feldstein 2010; Moore et al. 2010; Roundy et al. 2010; Franzke et al. 2011; Seo and Son 2012; Yoo et al. 2012a,b; Riddle et al. 2013). Other studies identified significant relationships between MJO-related tropical convection and North America wintertime surface weather (Mo and Higgins 1998; Higgins et al. 2000; Yao et al. 2011; Rodney et al. 2013; Johnson et al. 2014; Lin 2015; DelSole et al. 2017). Similarly, studies also have linked the MJO with the excitation of the NAO and the closely related AO, as well as the associated surface weather in the Atlantic-European sector and Arctic region (Vecchi and Bond 2004; Cassou 2008; L'Heureux and Higgins 2008; Lin and Brunet 2009; Lin et al. 2009, 2010). The studies above indicate that the MJO may contribute to skillful subseasonal forecasts of atmospheric teleconnection patterns and the associated extratropical surface weather.

In addition to the MJO, large convective anomalies over the tropical Pacific associated with El Niño– Southern Oscillation (ENSO) are also known to excite extratropical teleconnection patterns through the same basic Rossby wave train mechanism as with the MJO, such that El Niño (La Niña) results in a positive (negative) PNA-like response (Horel and Wallace 1981; Wallace and Gutzler 1981; Trenberth et al. 1998; Johnson and Feldstein 2010). Additionally, observational studies (Moron and Gouirand 2003; Pozo-Vázquez et al. 2001, 2005) and model simulations (Merkel and Latif 2002; Gouirand et al. 2007; Li and Lau 2012) indicated that ENSO may excite an NAO-like teleconnection over the North Atlantic–European sector.

A potential source of predictability from the extratropics is stratospheric forcing. Baldwin and Dunkerton (1999) provided evidence of downward propagation of zonal wind anomalies from the stratosphere to the troposphere over the course of ~ 3 weeks, which was followed by an NAO- or AO-like response in the troposphere. Model simulations have reproduced a negative NAO or AO response to weakened stratospheric winds on seasonal and longer time scales (Norton 2003; Scaife et al. 2005; Scaife and Knight 2008). This provides hope that the stratospheric downward coupling could be a useful predictor for subseasonal forecasts, particularly for anomalies associated with the NAO and AO.

The initial tropospheric flow pattern may be another potential source of skill for these lead times, although this source remains relatively unexplored. This promise takes root in studies that identified particular midlatitude wavelike initial flow patterns, which are closely associated with the jet stream waveguide effect (e.g., Branstator 1983, 2002), serving as precursors to extratropical circulation anomalies (Feldstein 2002; Mori and Watanabe 2008; Moore et al. 2010; Franzke et al. 2011; Teng et al. 2013; Risbey et al. 2015; McKinnon et al. 2016; Teng and Branstator 2017). In addition, Goss and Feldstein (2015) demonstrated that the extratropical response to tropical convection, particularly in association with the MJO, is sensitive to the midlatitude initial atmospheric flow. These studies provide hope that the diagnosis of the midlatitude initial atmospheric flow can be leveraged for weeks 3-4 forecasts of teleconnection patterns.

Although previous studies identify potentially important sources of skill for the dominant Northern Hemisphere teleconnection patterns, questions still remain about how much skill we can expect and how each predictor contributes to skill for lead times beyond two weeks. Motivated by these questions, this study employs a statistical approach to predict wintertime prominent Northern Hemisphere teleconnection patterns (i.e., PNA, NAO, and AO), at different lead times within weeks 1-5, and to identify the primary sources of skill for these teleconnection patterns. Specifically, we focus on the skill and its sources at a lead time of weeks 3-4 to be consistent with the coordinated effort within the United States to develop new outlooks in the 15-30-day range, which includes the new CPC experimental week 3-4 outlooks. Our statistical approach, based on partial least squares regression (PLSR), establishes a skill benchmark for dynamical forecast models and provides a small subset of dominant predictor patterns for each teleconnection pattern. We find that the statistical model yields significant skill for all teleconnection patterns at almost all lead times within weeks 1-5, providing promise for improving subseasonal-to-seasonal forecasts. Further investigation suggests that the significant skill at weeks 3-4 can be primarily attributed to the impact of tropical convection and extratropical initial flow.

The remainder of the paper is organized as follows. Section 2 provides the data sources and a detailed description of the statistical method that is used to generate the forecasts. Section 3 evaluates the prediction skill against observations and a representative state-of-theart dynamical model, followed by the exploration of sources of predictability at a lead time of 3–4 weeks in section 4. Finally, section 5 summarizes the results and discusses their implications.

2. Data and methodology

a. Data sources

We use the daily time series for the PNA, NAO, and AO indices from the NOAA/CPC spanning the period of 1980–2013. To be consistent with the CPC's current weeks 3–4 outlook, a low-pass filter (14-day running mean) is applied to the time series before isolating data from December–February (DJF) for wintertime forecasts. The filtered wintertime indices of the PNA, NAO, and AO are the target predictands, and therefore our forecasts are 2-week averages.

In this study, we generate statistical forecasts for teleconnection pattern indices based on a set of threedimensional meteorological variables referred to as predictor fields or simply predictors. As discussed in the introduction, previous studies have demonstrated that tropical convection anomalies, the midlatitude initial tropospheric flow, and for the NAO and AO, the initial stratospheric flow can have strong impacts on the extratropical circulation anomalies. To capture the first two effects for PNA, NAO, and AO forecasts, we consider tropical outgoing longwave radiation (OLR) and hemispheric 300-hPa geopotential height (Z300) as the first two predictor fields. We use daily NOAA interpolated OLR data (Liebmann and Smith 1996), covering the tropics (30°S-30°N). All pressure-level variables listed below are from the National Centers for Environmental Prediction–Department of Energy Atmospheric Model Intercomparison Project Reanalysis 2 (NCEP-DOE AMIP-II R-2; Kanamitsu et al. 2002). For daily Z300, we focus on the spatial domain covering the entire Northern Hemisphere and the Southern Hemisphere tropics (30°S-90°N). To incorporate the potential source of predictability from stratospheric downward coupling, we also include Northern Hemisphere 50-hPa geopotential height (Z50) as a third predictor field for NAO and AO forecasts (but not for PNA forecasts). All predictor fields have a spatial resolution of $2.5^{\circ} \times 2.5^{\circ}$; however, to increase the computational efficiency, all predictor fields are linearly interpolated onto a $5^{\circ} \times 5^{\circ}$ grid for the statistical forecasts. We conducted tests to verify that the interpolation to the coarse grid did not

have a noticeable impact on forecast skill. The seasonal cycle, which is defined as the calendar day means from 1980 to 2013, is removed from all predictor fields, and the resulting anomalies are standardized prior to generating forecasts.

To examine the linkage between Rossby wave activity propagation and extratropical circulation anomalies, 300-hPa wave activity fluxes (Takaya and Nakamura 2001) are calculated. The horizontal component of the wave activity flux (**W**, hereafter WAF) is given by

$$\mathbf{W} = \frac{1}{2|\mathbf{U}|} \begin{bmatrix} U(\psi_x^2 - \psi\psi_{xx}) + V(\psi_x\psi_y - \psi\psi_{xy}) \\ U(\psi_x\psi_y - \psi\psi_{xy}) + V(\psi_y^2 - \psi\psi_{yy}) \end{bmatrix},$$
(1)

where ψ is the perturbation streamfunction, and the subscripts denote partial derivatives. Here $\mathbf{U} = (U, V)$ is the 300-hPa time-mean horizontal wind vector for the winter. The winter seasonal mean 300-hPa WAF and streamfunction averaged over 1980–2013 are plotted in Fig. 1. The climatological WAF features strong meridional wave activity propagation over the North Pacific, North Atlantic, and Eurasia, emanating from the exit regions of the climatological jets.

For comparison with the statistical forecasts, we also evaluate the prediction skill for a set of 45-day retrospective forecasts (hindcasts) from the NCEP's Climate Forecast System, version 2 model (CFSv2; Saha et al. 2014), a state-of-the-art dynamical forecast model. The CFSv2 model is a fully coupled model consisting of atmospheric (NCEP Global Forecast System), oceanic (Geophysical Fluid Dynamics Laboratory Modular Ocean Model, version 4.0), land surface (Noah land surface model), and sea ice models. The hindcasts are initialized at 6-h intervals from 1999 to 2010 and run out for 45 days. For each day, the ensemble mean of four members is calculated. To construct the index for the AO from model outputs, we follow the CPC's definition for the AO, which is the projection of the daily 1000-hPa geopotential height anomalies poleward of 20°N onto the loading pattern of the AO from 1979 to 2000. For the PNA and NAO, we construct indices from the hindcast outputs using a different, yet similar, definition to that of the CPC. The hindcast PNA index is constructed by projecting the daily 500-hPa geopotential height anomalies onto the leading empirical orthogonal function (EOF) of wintertime monthly 500-hPa geopotential height from 1979 to 2008 for the domain of 10°-80°N and 160°E-60°W. Similarly, the hindcast NAO index is constructed by projecting the daily 500-hPa geopotential height anomalies onto the leading EOF of wintertime monthly 500-hPa geopotential height over the same time period for the domain of 20°-85°N and 80°W-0°. As in



FIG. 1. Wintertime (DJF) mean 300-hPa streamfunction $(10^7 \text{ m}^2 \text{ s}^{-1}, \text{shadings})$ and wave activity flux $(\text{m}^2 \text{ s}^{-2}, \text{vectors})$ over 1980–2013.

the statistical forecasts, we focus on evaluating the 2-week-averaged CFSv2 hindcasts for the PNA, NAO, and AO indices during winter months.

b. Generation of statistical forecasts for teleconnection pattern indices

We generate the 2-week statistical forecasts for the PNA, NAO, and AO indices in winter months from 1980 to 2013. The forecast period covers lead times from weeks 1–2 to approximately weeks 4–5.

The method of PLSR (Wold 1966) is adopted to generate the statistical forecasts. Several studies have demonstrated the utility of PLSR in geosciences for diagnostic and forecast purposes (Kalela-Brundin 1999; McIntosh et al. 2005; Abudu et al. 2010; Smoliak et al. 2010; Wallace et al. 2012; Smoliak et al. 2015). For a univariate predictand, PLSR essentially finds linear combinations of predictors (three-dimensional fields in this case) that maximize the variance explained in a predictand time series (a teleconnection pattern index in this application) through an iterative process. Therefore, in the present application, PLSR might be considered a means of finding "optimal" regressors (hereafter referred to as PLS components) to be used in an iterative linear regression, where these PLS components are based on linear combinations of high-dimensional dataset. This approach is attractive in the present application because we have prior understanding of the important predictor fields (OLR, Z300, and Z50) but not necessarily the specific predictor patterns that are optimal for exciting each teleconnection pattern at lead times beyond two weeks. PLSR provides a means of determining these dominant predictor patterns and associated PLS components, and for developing a linear model to predict the teleconnection pattern index based on these predictor patterns.

The iterative PLSR procedure is described as follows. Starting with a set of ranked area-weighted predictor fields, we

 calculate the grid-by-grid correlation coefficients between the standardized anomalies of the first predictor field and the predictand, obtaining a correlation map (i.e., predictor pattern);

- obtain a time series (i.e., the first PLS component of the first predictor) by projecting the first predictor field onto the correlation map obtained in step 1;
- use conventional least squares fitting and regress the predictand on the first PLS component (the time series in step 2) to obtain the first partial regression;
- linearly remove the first PLS component from both predictand and all predictor fields. The residual predictand and predictor fields become the new predictand and predictor fields;
- 5) repeat steps 1–4 to obtain higher ranked PLS components of the first predictor, until a stoppage criterion is met; and
- repeat steps 1–5 to obtain PLS components for the other predictors.

The final regression that links the predictor fields to the predictand is the sum of all partial regressions from each PLS component of each predictor. Each successive iteration explains more of the predictand variance. The PLSR approach also allows us to isolate PLS components that are linearly independent from each other, which may facilitate attribution of different physical processes that provide sources of prediction skill for the teleconnection patterns (limitations are discussed in section 5). For more discussion of the PLSR methodology, see, for example, Abdi (2010) and Smoliak et al. (2015). A natural question that therefore arises is, how many PLS components do we retain for each variable? We discuss the criteria we use in section 2c.

c. Cross validation

As mentioned previously in step 5 of the PLSR approach, a stoppage criterion has to be defined to determine the optimal number of PLS components of each predictor. If too many PLS components or predictor variables are retained, then our PLSR model will be overfitted, resulting in poor forecasts when applied to independent data. We choose to conduct a crossvalidated screening procedure to determine the optimal number of components to retain for each predictor. Whenever a screening procedure is applied, however, care must be taken to guard against artificial skill, which refers to biased estimates of skill arising due to the inclusion/exclusion of certain predictors without any cross validation on the screening procedure (DelSole and Shukla 2009). As pointed out by Michaelsen (1987), both the screening and the model building procedure must be cross validated; therefore, we adopt a double cross-validation approach in this study to ensure that the data used in our forecast evaluations remain completely independent of the model building process. The first (inner) cross validation determines the optimal number of PLS components of each predictor for the construction of the PLSR model, while the second (outer) cross validation validates the forecast skill of the constructed PLSR model in predicting an independent sample.

For the first (inner) cross validation, we calculate the incremental explained variance based on a repetition of leaveone-year-out sample data until each year is "left out" once. The basic idea is that we increase the number of retained PLS components for each predictor field until the incremental variance becomes negative (i.e., the crossvalidated forecasts become worse as measured by explained variance). The incremental variance explained by the *k*th PLS component of the *l*th predictor, $v'_{k,l}$, is defined by

$$\begin{aligned} & v_{k,l}' = v_{k,l}, \quad k = 1, \\ & v_{k,l}' = v_{k,l} - v_{k-1,l}, \quad k > 1, \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{v}_{k,l} &= \left(1 - \frac{\boldsymbol{v}_{k,l}^*}{\boldsymbol{v}_y}\right) \times 100\%, \quad l = 1, \\ \boldsymbol{v}_{k,l} &= \left(1 - \frac{\boldsymbol{v}_{k,l}^*}{\boldsymbol{v}_{l-1}^*}\right) \times \frac{\boldsymbol{v}_{l-1}^*}{\boldsymbol{v}_y} \times 100\%, \quad l > 1. \end{aligned}$$
(3)

Here v_{y} is the *total* variance of the original predictand, v_{kl}^* is the *residual* variance of the predictand after removing the predicted time series using k PLS components for the *l*th predictor, and v_{l-1}^* is the *residual* variance after removing the predicted time series using the previous l - 1 predictor(s). A negative v'_{kl} indicates an increase in the residual variance and decreasing forecast performance in terms of the explained variance of the predictand. We note that it is possible for v'_{kl} to be negative for k = 1, which would indicate that even the leading PLS component does not yield skillful forecasts. In that case, we reject the PLS component for that particular predictor. In contrast, positive incremental variance indicates additional variance explained by that PLS component, and so the PLS component is retained. In this sense, the sum of incremental variance explained in the predictand is maximized by an optimal combination of PLS components and predictors.

Both the PLS components and predictors are processed sequentially. For example, if OLR is the first predictor, we first determine the number of OLR components to retain. Then, we proceed to determine the number of PLS components of the second predictor (i.e., Z300). By doing so, the first Z300 PLS component captures the influence of Z300 *that is linearly independent* of all retained OLR components. We note that the change in forecast skill in terms of correlation due to different sequential ordering of the predictors is on the order of 0.01, so the sequential order of predictors has little effect on the final skill. However, the predictor order does impact the physical attribution owing to collinearity between predictor fields. We elaborate on this point in section 4 when we explore the sources of skill.

Now that we have constructed a PLSR model with an optimal number of PLS components and predictors, we have to validate its forecast skill on an independent sample, which is the purpose of the second (outer) cross validation, performed similarly by withholding data one year at a time.

To further illustrate the entire double-cross-validation procedure, let us consider the forecast for the year 1985. The PLSR forecasts are validated for December 1984– February 1985, and the PLSR model is constructed from all DJFs excluding 1985 (the training data). To construct the PLSR model, the number of retained PLS components for each predictor is determined with Eq. (2) through a cross validation with the training data (the first cross validation). After constructing the PLSR model, the forecasts are made for DJF 1985. This same procedure is carried out for each DJF until we have generated forecasts for each year. We then evaluate the forecasts for each DJF (the second cross validation).

The procedure described above indicates that data from years beyond the forecast year are included to build and evaluate the statistical model. The justification for this approach is that if the time scale of the growth and decay of atmospheric teleconnection patterns is on the order of 10 days (e.g., Feldstein 2000), and if the processes responsible for the growth and decay of these patterns do not change with time, then the data from future years represent independent samples that are just as valid as the data from previous years for training the model. However, future data will not be available for real-time forecasts, and so to test the validity of our assumption, we repeat the PLSR forecasts of teleconnection pattern indices for the period of 1999-2010 (to be consistent with the CFSv2 output), using data only from 1980 until the year prior to the forecast year. The results (not shown) indicate a reduction (~ 0.1) in the correlations for the PNA and NAO, especially for longer lead times, but a modest increase (~ 0.04) of the AO correlation skill. Therefore, the effect of removing future years from the training data is mixed, and it remains unclear whether such reduction in performance for the PNA and NAO is due to the exclusion of future information or simply due to a reduced sample size for model training. Since the double cross validation should ensure the realistic

TABLE 1. The optimal number (averaged over all years) of PLS components to be retained for each predictor. The results are from forecasts at a lead time of weeks 3–4.

Teleconnection pattern	Predictors	Averaged optimal number of PLS components
PNA	OLR	1
	Z300	2.56
NAO	OLR	0
	Z300	1.88
	Z50	0.12
AO	OLR	0.09
	Z300	2.06
	Z50	0.06

assessments on the forecast skill of PLSR approach, we retain the approach that uses all available years for model training.

We also note that the double cross validation determines that the final forecasts of the teleconnection patterns may be based on different combinations of predictors and/or PLS components over different years. This variation reflects the property that the prediction model may be subject to underfitting or overfitting for individual years due to the uncertainty in the optimal number of PLS components for each predictor. However, because the forecast years remain separate from the model-building process, our estimates of forecast skill are likely to be more realistic than if we chose a fixed number of PLS components and did not account for the uncertainty in this number. For making physical interpretations, we focus only on those relationships that are robust across most years (see section 4). We list in Table 1 the optimal number of PLS components for the weeks 3-4 forecasts averaged over all years for each predictor and each teleconnection pattern. It is worth emphasizing again that since the optimal number can vary from one year to another due to the double cross validation, the averaged optimal number may not be an integer. Overall, we find that a larger number of PLS components of Z300 are desired for the construction of the PLSR model. However, it is surprising to see that although tropical convection anomalies are believed to strongly impact the downstream PNA and NAO or AO response, relatively few of the screened PLS components of OLR are retained for the forecasts.

3. Prediction skill of PLSR forecasts

In the previous section, we discussed how the PLSR model is constructed with an optimal number of PLS components and predictors for each year's forecast. In this section, we evaluate the prediction skill of the constructed PLSR models by first calculating the



FIG. 2. Boxplots of the correlation between forecasts and observed wintertime (DJF) 2-week-averaged (top) PNA, (middle) NAO, and (bottom) AO index time series at different forecast lead times. The boxplots are generated based on the reshuffled data using Monte Carlo approach described in section 3. The blue (green) lines denote the correlation of PLSR forecasts evaluated during1980–2013 (1999–2010). The yellow lines denote the correlation of CFSv2 dynamical forecasts evaluated during 1999–2010. The purple solid and black dashed lines denote the correlation from persistence and cross-validated climatological forecasts, respectively.

forecast biases, and verifying that the biases are on the order of 0.001-0.01 at all lead times for all three teleconnection patterns. We then calculate the correlation coefficients r between forecasted and observed teleconnection pattern indices at different lead times. We show in Fig. 2 the correlations between PLSRforecasted and observed teleconnection pattern indices (blue lines), and their box plots based on the resampled data using a Monte Carlo approach, where the calendar years of the predictand and the forecasted time series are randomly reshuffled 1000 times. The PLSR forecasts noticeably outperform the persistence forecasts (purple lines) and climatological forecasts (black dashed lines) at all lead times for the PNA and AO, while for the NAO, the PLSR forecasts perform close to the climatological forecasts and slightly worse than the persistence forecasts at longer lead times. Here, climatological

forecasts are defined by the cross-validated calendar day means of the teleconnection pattern index. The prediction skill of PLSR forecasts decreases relatively quickly for short lead times, as we would expect based on the loss of skill associated with the initial conditions. Beyond week 3, however, the PLSR forecast skill remains rather stable. Specifically, at weeks 3-4, the correlations between PLSR forecasts and observations for the PNA, NAO, and AO are 0.34, 0.28, and 0.41, all of which outlie the maximum correlation based on the resampled data and therefore can be interpreted as significantly different from zero. The prediction skill of the CFSv2 hindcasts (yellow lines) over 1999-2010 and that of the PLSR forecasts validated over the same period as CFSv2 (green lines) for each teleconnection pattern are also displayed in Fig. 2. The correlations validated over a shorter period for PLSR forecasts are highly consistent



FIG. 3. The root-mean-square errors (RMSEs) of the forecasts of wintertime (DJF) (top) PNA, (middle) NAO, and (bottom) AO index time series at different forecast lead times, evaluated over 1980–2013 unless stated otherwise. The purple solid (black dashed) lines indicate the results from persistence (cross-validated climatological) forecasts. The blue and yellow lines denote the results from PLSR forecasts and CFSv2 output (evaluated over 1999–2010), respectively.

with those validated over the full period. For shorter lead times (weeks 1–3), CFSv2 hindcasts show higher prediction skill, as we might expect, given the importance of dynamical processes associated with the initial condition details at these shorter lead times. For longer lead times (~4 weeks onward), however, the PLSR forecasts are able to perform comparably to, or even outperform the CFSv2. Overall, we find that the PNA and NAO forecast skill of CFSv2 and PLSR shows moderate to strong improvements in weeks 3–4 compared to those evaluated using CFSv1 by Johansson (2007); such improvements in correlation can be up to $\sim +0.2$ ($\sim +0.3$) for PLSR (CFSv2).

In addition, we evaluate and illustrate in Fig. 3 the root-mean-square errors (RMSEs) of the PLSR forecasts evaluated over 1980–2013 (blue lines) and compare them with those from CFSv2 hindcasts (yellow lines), and cross-validated climatological (black dashed lines) and persistence (purple lines) forecasts. We note that the RMSEs of PLSR forecasts evaluated over

TABLE 2. Correlation coefficients between PLSR forecasts and observed teleconnection pattern indices in weeks 3–4. The second through sixth columns give the correlation of PLSR forecasts where no running mean (rm), 7-day, and 14-day running mean is applied to the predictors to forecast 2-week averaged predictands. The subcolumns under "14-rm" list correlation coefficients of forecasts for full record (1980–2013), last 25% (2005–13), and first 25% (1980–88) of the teleconnection pattern indices. The last column gives the correlation of PLSR forecasts where unfiltered predictors are used.

		2-week averaged predictands Predictors				
Teleconnection		14 rm				Daily predictands
pattern	Norm 7 rm	7 rm	Full forecasts	Last 25% forecasts	First 25% forecasts	and predictors
PNA	0.18	0.19	0.34	0.26	0.33	0.16
NAO	0.22	0.24	0.28	0.25	0.12	0.11
AO	0.42	0.43	0.41	0.54	0.16	0.33

1999–2010, which are almost identical to those evaluated over the full period, are therefore not displayed. Figure 3 shows that the PLSR forecasts have lower RMSEs than the climatological and persistence forecasts at all lead times for the PNA and AO. In weeks 3–4, PLSR forecasts have comparable RMSEs to those of CFSv2 for all three teleconnection patterns. Overall, the significant correlations and the low RMSEs of the PLSR forecasts in weeks 3–5 provide promise that the sources of skill at these lead times may provide statistical guidance for improving subseasonal and seasonal forecasts.

We tested the sensitivity of forecast performance to certain analysis choices, including the predictor/predictand filtering and the training/validation data partitioning, with a focus again on weeks 3-4. The evaluations of each set of test forecasts are listed in Table 2. We first investigate whether different time intervals over which the predictors are averaged affects the prediction skill. We generate two other sets of test forecasts: one using daily predictors and the other using 1-week averaged (7-day running mean applied) predictors to predict 2-week averaged predictands. For the PNA, the prediction skill of the PLSR forecasts is significantly increased by applying a 2-week average to both the predictands and predictors; on the other hand, the forecast skill for the NAO and AO does not change much with different time intervals over which the predictors are averaged (the second through fourth columns in Table 2). Because the skill is higher or comparable with 14-day-averaged predictors, we use 14-day running mean predictor data for the remainder of the study.

We calculated the difference in skill between daily and 14-day mean predictands. Christiansen (2005) previously found that statistical forecasts can be improved substantially by averaging the predictand over 10 or more days for extended-range lead times (up to 25 days). Consistently, at a lead time of weeks 3–4, the prediction skill of 2-week forecasts (the second through fourth columns) of the teleconnection pattern indices is higher than that of the daily forecasts (last column in Table 2).

We also attempted to increase forecast skill by incorporating an interaction term, which is defined as the dot product of any two predictor fields. However, adding interaction terms did not yield additional skill (not shown). Additionally, to evaluate the importance of underlying low-frequency variability and its impact on prediction skill, we generated two sets of test forecasts using the first (last) 75% of the data to predict the last (first) 25% of the data and their prediction skill is listed in the fifth and sixth columns of Table 2. We find that for the PNA, the prediction skill does not vary much among the two sets of test forecasts, and they are relatively close to those from the forecasts using full record. However, for the NAO and AO, the forecasts of the last 25% of the data (column 5) are much more skillful than the forecasts of the first 25% of the data (column 6). The considerably better forecast skill of the NAO and AO for the last 25% of the data suggests considerable lowfrequency variability of the predictand-predictor relationships and forecast skill. Similar results were also found in dynamical models (e.g., Scaife et al. 2014; Kang et al. 2014). The source of this variability is not immediately clear.

Although the PLSR forecasts for the PNA, NAO, and AO yield robust and significant skill in weeks 3–4 and even longer lead times, we also note some unexpected differences among the forecasts of the teleconnection patterns. For example, the source of skill for the PNA is thought to be most closely related to tropical Pacific convection, and Table 1 confirms that an OLR component is always retained for PNA forecasts but not for NAO or AO. The PLSR forecasts of the PNA, however, do not have the highest correlations. In fact, both the PLSR and CFSv2 forecasts demonstrate the highest skill for the AO forecasts. In addition, the forecast skill of the NAO is considerably lower than that of the AO, despite the NAO often being viewed as a local manifestation of the planetary-scale AO. In fact, the correlation between the two indices during DJF is relatively modest (0.69) during 1980–2013. These results suggest that a large fraction of the AO variance unassociated with the NAO may be predictable in weeks 3–4. Overall, the promising forecast skill out to week \sim 5 based on the PLSR model potentially can help to improve forecasts on subseasonal and seasonal time scales.

4. Sources of prediction skill from the PLSR model

The PLSR model is shown to be significantly skillful in forecasting the teleconnection pattern indices in weeks 3–4; therefore, our next goal is to investigate the sources of prediction skill in the statistical model. As indicated by Table 1, an advantage of the PLSR approach is that it decomposes the forecasts and their skill into a generally small number (<3) of PLS components and their corresponding predictor patterns. This helps us focus on the dominant physical processes that provide skill.

While the double cross validation largely avoids overfitting and artificial skill due to non-cross-validated predictor screening, it yields predictor patterns (spatial patterns associated with their corresponding PLS components; the correlation maps described in step 1 of the previous section) that can be different from one year to another, which would be a disadvantage for physical interpretations of sources of predictability. Therefore, we choose to carry out lagged regression analysis as an alternative approach described as follows. We first specify a priori the number of PLS components to retain, which means the inner cross validation used to determine the optimal number of PLS components to be retained is absent. We generate leave-one-year-out cross-validated forecasts following the procedure outlined in section 2b. We then calculate how much skill is gained by adding an additional PLS component. Next, we subtract the forecasted teleconnection pattern index without this additional component from the forecasts with this additional component to obtain a residual index time series. For example, if we retain one OLR component followed by adding one Z300 component, then the residual index time series for the OLR component (OLR_1) would simply be the forecasted index time series using only OLR₁; the residual index time series for the Z300 component $(Z300_1)$ would be the forecasted index time series using both OLR₁ and Z300₁ minus the forecasted index time series using only OLR₁. We then regress meteorological fields on the residual index time series at different negative lag times to examine the evolutions of the meteorological variables associated with a particular PLS component. Because of the 2-week averaging, the start time centered at

TABLE 3. Evaluation of PLS forecasts when adding each PLS component successively in weeks 3–4. The subscripts denote the rank of PLS component of each predictor. Here r and Δr stand for the correlation and change in correlation by adding the PLS component in question. The incremental variance [as defined in Eq. (2)] explained by each PLS component of each predictor is listed in the fifth column. The PLS components explaining positive incremental variance are highlighted in boldface fonts; for these PLS components, in the last column we list the pattern correlation (r_p) between Z300 anomalies at forecast validation time and the corresponding teleconnection pattern.

Pattern	PLS component	r	Δr	Incremental variance (%)	r_p
PNA	OLR1	0.26	_	6.18	0.77
	Z3001	0.28	+0.02	0.12	0.52
	Z300 ₂	0.37	+0.08	2.00	0.20
	$Z300_{3}$	0.39	+0.02	-1.36	_
NAO	Z3001	0.21	_	3.92	0.65
	Z300 ₂	0.31	+0.10	-1.74	_
	$Z50_{1}$	0.29	-0.02	-1.45	_
	Z50 ₂	0.33	+0.04	5.03	0.70
AO	OLR_1	0.20		-2.29	_
	OLR ₂	0.29	+0.09	4.99	0.37
	Z3001	0.28	-0.01	-1.07	_
	Z300 ₂	0.39	+0.11	6.65	0.58
	Z300 ₃	0.40	+0.01	-4.67	_
	Z501	0.42	+0.02	0.68	0.31
	Z50 ₂	0.41	-0.01	-0.87	—

lag = -22 days corresponds to forecast initialization at a lead time of weeks 3–4. Similarly, lag = 0 days corresponds to the 2-week period that is centered at the forecast validation time. By carrying out the procedure in this manner, we identify cross-validated patterns that not only pass the screening procedure but also reflect sources of predictability that are *linearly independent* from each other and that contribute to *additional* skill in the final forecasts. To test the statistical significance of the regression coefficients, we adopt an approach with adjusted standard error and adjusted degrees of freedom as outlined in Santer et al. (2000).

In Table 3, we evaluate the prediction skill of PLSR forecasts r, the change in prediction skill Δr , and the explained incremental variance by successively adding an extra PLS component in the lagged regression analysis described above. It is worth emphasizing that since we specify whether or not to retain a particular PLS component in this procedure, the first (inner) cross validation, which is used to determine the optimal number of PLS components and predictors, is absent. As a result, the PLSR forecasts might seemingly have higher skill than those with double cross validation reported in the final forecasts. Let us take PNA forecasts as an example. If we choose to retain three Z300 PLS components, r = 0.39 (Table 3), which is higher than r = 0.34 with double cross validation mentioned in



FIG. 4. Lagged regression coefficients of 2-week-averaged OLR anomalies (W m⁻²) associated with OLR₁ for PNA forecasts. The stippling denotes statistical significance at the 10% level. The following conventions apply from this figure and all subsequent figures: the coefficients are scaled by the standard deviation of the basis time series for regression; the number of lag days indicates the centered day of a 2-week period, therefore, lag = 0 day corresponds to the 2-week period centered at the forecast validation time, while lag = -22 days corresponds to forecast initialization at a lead time of weeks 3–4.

section 3. However, we see from Table 1 that the optimal number of Z300 for PNA is smaller than three, which means that the third PLS component of Z300 is not retained for each year's forecast. Therefore, a seemingly higher skill would be concluded if we chose to retain three Z300 PLS components; however, we have no a priori reason to assume three retained components, and so we emphasize the concept that without cross validating the model building procedure, we would introduce artificial skill.

Overall, we find that the number of PLS components that explain positive incremental variance for each teleconnection pattern index in Table 3 is mostly consistent with the optimal number listed in Table 1. In addition, PLS components that explain more positive incremental variance typically help improve the forecasts more, as reflected by both the higher Δr^1 and the positive pattern correlations (r_p) between the associated Z300 regression patterns and the corresponding teleconnection loadings at forecast validation time. Therefore, we focus on the regression coefficients associated with the PLS components that explain positive incremental variance (highlighted in bold) in Table 3, unless noted otherwise.

a. PNA

Consistent with expectations rooted in previous studies, we find that OLR is a skillful predictor of the PNA for weeks 3–4. The OLR regression pattern

associated with the first OLR PLS component (OLR₁) features a dipole pattern with enhanced convection in the central equatorial Pacific and suppressed convection over the Maritime Continent (Fig. 4). Such a pattern of OLR anomalies is reminiscent of a prominent El Niño signature, which is supported by the relatively large lagged correlation (>0.6) between the Niño-3.4 index and the residual index time series for OLR₁ throughout the weeks 3-4 period, although the influence of MJO convection may be mixed within the regression pattern. Accordingly, the extratropical response is a classic wave train pattern originating over the eastern tropical Pacific and propagating downstream into North America and North Atlantic, projecting onto a positive PNA pattern that persists through weeks 3-4 (Fig. 5, contours). The regression patterns for the WAF (Fig. 5, vectors) show strong WAF anomalies in the eastern Pacific, which are approximately collocated with the Z300 anomalies, associated with positive PNA forecasts; meanwhile, there are also strong WAF anomalies downstream in the North Atlantic and Europe associated with the positive PNA. Such features are consistent with the 300-hPa streamfunction composites of the positive PNA reported in previous studies (e.g., Feldstein 2002; Mori and Watanabe 2008; Franzke et al. 2011). In general, it can be inferred from OLR_1 that the statistical model is able to capture and take advantage of the persistence of the PNA associated with tropical heating, predominantly ENSO episodes, to generate skillful forecasts.

The skill of PNA forecasts for weeks 3–4 based solely on a single OLR PLS component is r = 0.26 (Table 3). This value is substantially lower than the skill with all screened PLS components of predictors (r = 0.34), which means that the initial extratropical flow (Z300)

¹Since the regression model is cross validated, there is not necessarily a one-to-one correspondence between *r* and explained variance or between Δr and incremental variance in Table 3.



FIG. 5. Lagged regression coefficients of 2-week- averaged Z300 anomalies (m, contours) and WAF ($m^2 s^{-2}$, vectors) associated with OLR₁ for PNA forecasts. The green stippling and thickened vectors denote the 10% significance of Z300 anomalies and WAF, respectively.

must account for the additional gain in skill. We next examine the lagged regression coefficients of Z300 associated with the first Z300 PLS component $(Z300_1;$ Fig. 6, contours). The overall evolution of the patterns demonstrates a phase transition from the negative PNA at weeks 3-4 lead to the positive PNA at forecast validation time. Based on the previous findings that MJO phases 1-3 (5-7) are followed 7-10 days later by the negative (positive) PNA (Mori and Watanabe 2008; Johnson and Feldstein 2010; Moore et al. 2010; Roundy et al. 2010; Franzke et al. 2011; Yoo et al. 2012b; Riddle et al. 2013; Goss and Feldstein 2015), we expect that the phase transition of the PNA would be closely related to the phase change of the MJO. However, there is little change in the associated OLR anomaly patterns (not shown). On the other hand, we see in Fig. 6 a clear circumglobal wavelike pattern in the midlatitudes, which was shown to be closely associated with the waveguide of troposphere jets (e.g., Branstator 1983, 2002). In the North Atlantic, we see a positive NAO pattern and a dipole pattern resembling Ural blocking (UB) over Eurasia, which seems to capture the close connection between positive NAO events and UB events that was previously reported (e.g., Luo et al. 2016a,b). Meanwhile, the positive center of a negative PNA-like structure over the northern Pacific at forecast initialization gradually moves equatorward, and a negative center forms over the Bering Sea. These two anomaly centers, together with the disturbances over the North America–North Atlantic sector, form a positive PNAlike structure. Overall, the extratropical initial flow, independent of the MJO, seems to be responsible for the phase change of the PNA captured in the regression patterns.

The second Z300 PLS component (Z300₂) accounts for even more skill than Z300₁ (Table 3). The regression patterns of the Z300 anomalies associated with Z300₂ (Fig. 7, contours) show that wavelike anomalies are mostly located in mid- and high latitudes, which align quite well with zones of strong WAF anomalies (Fig. 7, vectors). Meanwhile, the associated regression patterns of tropical OLR anomalies are weak and disorganized (not shown). These findings suggest that Z300₂ captures processes internal to the extratropical atmosphere. Specifically, a negative NAO-like pattern that extends eastward toward the Ural Mountains can be seen, with the anticyclonic anomaly center being dominant, though it is more confined in space. More importantly, we see in



FIG. 6. Lagged regression coefficients of 2-week- averaged Z300 anomalies (m, contours) and WAF (m² s⁻², vectors) associated with Z300₁ for PNA forecasts on selected lag days.

Fig. 7 a typical pattern of upper-tropospheric variability near the time-averaged jet—a meridionally confined and zonally elongated north–south dipole over Eurasia in the vicinity of the Asia jet. Meanwhile, we see in midlatitudes that wave activity propagates across the North Pacific and North America, revealing a pattern that is reminiscent of a positive PNA pattern up to approximately lag -10 days. The waveguide structure over East Asia and the North Pacific, including the downstream PNA pattern, are consistent with previous studies (e.g., Hoskins and Ambrizzi 1993; Risbey et al. 2015; Teng and Branstator 2017), indicative of the fundamental role of the midlatitude background flow in PNA development (Feldstein 2002; Mori and Watanabe 2008; Franzke et al. 2011; Risbey et al. 2015; Teng and Branstator 2017).

b. AO

Next, we examine the regression patterns associated with the PLS components of OLR, Z300, and Z50 that account for most of the AO forecast skill for weeks 3–4. AO forecasts based on the first OLR PLS component (OLR₁) only yield positive skill in terms of correlation (r = 0.20) but negative skill in terms of explained incremental variance (-2.29%) (Table 3). Positive correlation skill but negative explained variance indicates

overconfident forecasts (i.e., the variance of the forecasts is too high relative to the correlation). The negative incremental variance is also the reason that OLR is rarely retained as a predictor of the AO (Table 1). However, considering the relatively high prediction skill when including OLR₁, and the strong pattern correlation between its associated regression pattern of Z300 and the AO ($r_p = 0.54$) at forecast validation time, we choose to combine OLR_1 and OLR_2 (OLR_C = $OLR_1 + OLR_2$) to examine the impact of tropical convection on AO forecasts. The regression patterns of OLR anomalies associated with OLR_C for the AO (Fig. 8) depict convection anomalies that resemble a prolonged MJO phase \sim 1–3 with weakening amplitude over time. In response, a clear wave train can be observed that originates in the tropical Pacific and meanders downstream into North America, the North Atlantic, and eventually into Eurasia, projecting onto a negative PNA, a positive NAO and AO pattern, and a UB pattern (Fig. 9, contours). Consistently, strong WAF anomalies (Fig. 9, vectors) can be observed near the centers of the Z300 anomalies. These findings reflect the expected source of predictability for the NAO and AO from MJOrelated tropical Pacific convection (e.g., Cassou 2008; L'Heureux and Higgins, 2008; Lin et al. 2009).



FIG. 7. Lagged regression coefficients of 2-week- averaged Z300 anomalies (m, contours) and WAF ($m^2 s^{-2}$, vectors) associated with Z300₂ for PNA forecasts. The green stippling and thickened vectors denote the 10% significance of Z300 anomalies and WAF, respectively.

Apart from the tropical influence, the second Z300 PLS component (Z300₂) also contributes to substantial skill in terms of improvement in correlation ($\Delta r = +0.11$) and explained variance (+6.65%, Table 3). We therefore examine the Z300 regression pattern associated with Z300₂, which shows a northern Eurasian wave train and a prominent West Pacific (WP)–North Pacific Oscillation (NPO) pattern (Wallace and Gutzler 1981;

Nigam 2003; Linkin and Nigam 2008; Nigam and Baxter 2015) at forecast initialization time (Fig. 10, top-left contours). Large regression coefficients of WAF are collocated with the northern Eurasian wave train (Fig. 10, vectors), and both persist through the first half of the lagged regression period. These results appear to agree with Branstator (2002), which showed that the circumglobal waveguide pattern induced by time mean



FIG. 8. Lagged regression coefficients of 2-week- averaged OLR anomalies (W m⁻²) associated with OLR_C for AO forecasts. The stippling denotes statistical significance at the 10% level.



FIG. 9. Lagged regression coefficients of 2-week- averaged Z300 anomalies (m, contours) and WAF ($m^2 s^{-2}$, vectors) associated with OLR_C for AO forecasts. The green stippling and thickened vectors denote the 10% significance of Z300 anomalies and WAF, respectively. The polar projections cover the entire Northern Hemisphere.



FIG. 10. Lagged regression coefficients of 2-week- averaged Z300 anomalies (m, contours) and WAF ($m^2 s^{-2}$, vectors) associated with Z300₂ for AO forecasts. The green stippling and thickened vectors denote the 10% significance of Z300 anomalies and WAF, respectively. The polar projections cover the entire Northern Hemisphere.

tropospheric jet has a noticeable north-south dipole structure over the North Atlantic and projects onto the positive NAO. Over Eurasia, a pattern reminiscent of UB pattern can be seen, though the anticyclonic anomaly is displaced westward and continues to move westward during the lagged regression period. Furthermore, it is intriguing to see from Fig. 10 (contours) a twocenter AO pattern evolving into a monopole pattern. Specifically, two negative Z300 anomaly centers exist at forecast initialization time, with one over the Bering Strait (which appears to also be the northern branch of the WP/NPO) and the other over western Russia. While the negative center over western Russia appears to undergo stationary decay, the negative center over the Bering Strait seems to gradually move eastward, merging with a newly emerging weak negative center over the Davis Strait and southern Greenland, forming a single negative Arctic anomaly. Meanwhile, a midlatitude zonal belt of positive Z300 anomalies persists throughout the entire composite period.

We note that in the final AO forecasts, virtually no OLR PLS components are retained (Table 1); in this case, we expect the lagged regression patterns associated with the first Z300 PLS component $(Z300_1)$ to capture the influence of tropical convection. To verify this, we repeat the lagged regression analysis with no OLR components retained, and examine the regression patterns associated with each PLS component. The regression patterns of OLR associated with Z3001 bear prominent signatures of La Niña (not shown), and the corresponding Z300 regression patterns resemble those in Fig. 9. Similarly, when no OLR components are retained, the Z300 regression patterns associated with the second Z300 PLS component (Z300₂) also resemble those in Fig. 10. In addition, the maximum skill of the AO forecasts with the PLS components specified a priori $(r_{\text{max}} = 0.42, \text{Table 3})$ is highly consistent with that of the final forecast (r = 0.41). Based on these findings, we can conclude that (i) when OLR components are not retained, as in the final forecasts of the AO, the subsequent $Z300_1$ indeed captures the influence of tropical convection; (ii) the decision whether or not to retain OLR components has virtually no impact on the AO forecast skill.

Although many studies focus on the role of the stratosphere in the long-lead predictability of the NAO and AO, we find that the addition of a Z50 anomaly predictor field only contributes minor improvements to the AO forecast skill (Table 3). Nevertheless, it is worthwhile to learn how the first Z50 PLS component (Z50₁) contributes to AO skill in weeks 3–4. Since previous studies have shown that the coupling between stratosphere and troposphere is manifested by downward propagation of westerly anomalies, we examine the regression pattern of the meridional-vertical cross section of the zonally averaged zonal wind anomalies associated with $Z50_1$ (Fig. 11). At forecast initialization time, there are significant westerly anomalies throughout the troposphere between 45° and 70°N, which dominates the entire mid-/high-latitude region of the Northern Hemisphere. To the south of the westerly anomalies, there is a thin vertical channel of weaker easterly anomalies. The subsequent evolution of the lagged regression patterns generally captures the persistence of the features described above and the gradual retraction of the wind anomalies toward the stratosphere. Overall, the regression patterns of the zonal wind anomalies are consistent with the findings previously reported in observational studies and model simulations of the stratospheric influence on the troposphere (e.g., Black 2002; Song and Robinson 2004; Scaife et al. 2005), but downward propagation of the stratospheric zonal wind anomalies does not appear to play a major role in weeks 3-4 NAO and AO forecasts, at least in the linear statistical model that we use.

It is worth noting that due to the limitations of the methodology, the minor improvement from $Z50_1$ to the AO forecast skill may not accurately quantify the stratospheric contribution to AO predictions for weeks 3-4. In particular, if stratospheric variability is linearly related to the OLR or to the Z300 predictor patterns, then the stratospheric influence may be attributed to the OLR and Z300 predictors because those predictors were screened prior to Z50. Upon examining the regression patterns of zonal wind anomalies associated with OLR_{C} and $Z300_2$, we find that the former have similar features as in Fig. 11, although much less significant. Nevertheless, some linear relation between OLR predictor patterns and stratospheric variability is indicated. Therefore, the full potential of the usage of a stratospheric predictor for weeks 3-4 NAO and AO forecasts is not completely clear and is therefore worth additional pursuit.

c. NAO

Because the NAO is closely related to the AO but the prediction skill is noticeably lower than those of the PNA and AO (Table 3), we only briefly discuss the results for the NAO. Although the PLS components of OLR do not pass the screening procedure (Table 1), we find that the first Z300 PLS component (Z300₁) captures the influence of tropical Pacific convection on NAO forecasts, predominantly associated with La Niña episodes based on the pattern resemblance and the relatively high anticorrelation (\sim -0.5) between the Niño-3.4 index and the residual index time series for Z300₁ (not shown). The inability of the OLR PLS components to pass the screening procedure might again be due to



FIG. 11. Lagged regression coefficients of meridional–vertical cross section of zonally and 2-week-averaged zonal wind anomalies (m s⁻¹) in the Northern Hemisphere associated with Z50₁ for AO forecasts. The contour lines are at 0.5 m s⁻¹ interval, with zero line highlighted in bold. The stippling denotes statistical significance at the 10% level.

the fact that the OLR field tends to be noisier with more spatial degrees of freedom than the Z300 field, and so it may be more difficult for the OLR PLS components to pass the screening procedure described in section 2. While the MJO influence on the NAO response is often emphasized, the linkage between ENSO and the NAO has been somewhat inconsistent (e.g., Toniazzo and Scaife 2006; Brönnimann 2007). Therefore, it is surprising to observe a well-captured ENSO influence on weeks 3–4 NAO forecasts.

The influence of the midlatitude background flow is, however, not as evident in the regression patterns associated with $Z300_1$. Despite that the second Z50 PLS component (Z50₂) accounts for considerable amount of incremental variance, its inclusion in the forecasts does not significantly improve the skill (Table 3). Additionally, zonal wind anomaly regression patterns associated with Z50₂ are qualitatively similar to those for the AO (Fig. 11), although much weaker and less significant in the lower troposphere. The low forecast improvement and insignificant regression patterns associated with Z50₂ could again, at least partially, be due to the collinearity among the predictor fields as previously discussed.

5. Summary and discussion

In this study, we examine the subseasonal wintertime forecast skill and sources of predictability of the dominant Northern Hemisphere teleconnection patterns, as determined through a partial least squares regression (PLSR) approach. Specifically, we generate 2-week forecasts of the wintertime PNA, NAO, and AO indices at different lead times of up to five weeks. We consider potential predictor fields of tropical OLR anomalies, tropical Northern Hemispheric Z300 anomalies for PNA forecasts, as well as Northern Hemispheric Z50 anomalies for NAO and AO forecasts, because we believe that these three fields can capture the impacts of tropical convective heating, the midlatitude background flow, and stratospheric downward coupling, which have been previously demonstrated to be fundamental to the development of these teleconnection patterns. The PLSR forecasts perform comparably to, or even outperform the benchmark of dynamical models at lead times beyond approximately three weeks. This suggests the plausibility of leveraging statistical relationships and developing statistical or hybrid dynamical-statistical tools to improve predictions. The sources of predictability in weeks 3-4 are investigated using lagged regression analysis. As expected, tropical Pacific convection, including anomalies related to ENSO and MJO activity, is an important source of skill. In addition, the initial state of the midlatitude flow seems to contribute substantially to forecast skill. However, more studies are needed to understand the mechanisms through which the initial flow exerts its impact onto the subsequent downstream circulation.

In addition to the predictors mentioned above, we note that other potential predictors have been previously explored. For the NAO and AO, there have been studies (e.g., Rodwell et al. 1999; Rodwell and Folland 2002; Hurrell et al. 2003; Folland et al. 2012; Scaife et al. 2014; Smith et al. 2016) highlighting the source of predictability from extratropical sea surface temperatures (SST). Although most of these studies focused on the predictability on longer, seasonal time scales, we examined if ocean memory could provide any prediction skill on subseasonal time scales. We replaced tropical OLR anomalies with SST anomalies north of 30°S as the first predictor, in the hope of capturing both the impacts of tropical convection and extratropical ocean forcing. However, the PLS components associated with SST anomalies did not contribute any additional skill. Similarly, Arctic sea ice concentration anomalies, another potential predictor that has been hypothesized as a source of subseasonal to seasonal NAO and AO prediction skill (Alexander et al. 2004; Folland et al. 2012; Scaife et al. 2014; Smith et al. 2016), was also added to the candidate PLSR predictors. Again, at lead times of 3-4 weeks, Arctic sea ice concentration anomalies did not contribute additional skill.

We also attempted to apply the PLSR approach to summertime forecasts, although we expected lower prediction skill since the Rossby wave source, midlatitude westerlies that facilitate Rossby wave propagation, and ENSO signals are all much weaker during summer. Indeed, the overall prediction skill for teleconnection pattern indices is considerably lower, and the skill drastically decreases beyond week 2 (not shown). To refine the PLSR approach for potential operational use, exploration of additional predictors may be necessary, as is true for all other seasons.

Despite the promising prediction skill of the PLSR approach, several limitations of the method should be recognized. First, the potential collinearity among two or more predictor fields can make it very difficult to accurately attribute the source of predictability to a particular predictor. For example, the lagged regression results in section 4 show that there are cases where it is difficult to completely disentangle the impact of OLR and Z300 because of their close linkage. Second, being a linear method, the PLSR approach is unable to capture the nonlinearity, which might be crucial when studying the teleconnections of ENSO and the combined MJO and ENSO influence (Hoerling et al. 2001; Toniazzo and Scaife 2006; Roundy et al. 2010; Moon et al. 2011; Johnson and Kosaka 2016) as well as stratospheric– tropospheric coupling. Last, there is a tendency for PLSR to overfit because not only the regression coefficients, but also the PLS components are sample dependent. With the limited training data, the overfitting could prohibit more PLS components from passing the screening procedure and emerging as robust regressors.

Finally, although the results indicate that our statistical approach has comparable prediction skill to a current dynamical forecast model at lead times of \sim 3–5 weeks, we do not disregard the usefulness of dynamical models in subseasonal-to-seasonal predictions. For example, Vitart and Molteni (2010) found that both MJO and its teleconnections are better represented in a set of reforecasts with a coupled atmosphere-ocean model. Xiang et al. (2015) showed that the upper bound of prediction skill for the MJO can reach out to 42 days in a new version of the Geophysical Fluid Dynamics Laboratory (GFDL) coupled model. These findings provide promise that ongoing improvements on dynamical forecasts of the MJO will likely carry through to improve the dynamical forecasts for extratropical teleconnection patterns. As dynamical models continue to improve, we expect that dynamical forecast model performance may unambiguously surpass that of statistical models while also providing important tools for understanding the mechanisms of subseasonal predictability. Nevertheless, the approach we have taken establishes a statistical forecast benchmark while remaining parsimonious, which may be useful as a source of forecast guidance. Because we are able to generate skillful forecasts with a small subset of predictor patterns, we may shed light on the large-scale conditions that are associated with skillful predictions on the dominant teleconnection patterns of lead times of 3-4 weeks. Although the forecast skill is fairly modest by weeks 3-4, the identification of large-scale precursors to teleconnection pattern development may allow us to identify "forecasts of opportunity" when the expected forecast skill is higher than normal. Further, the general approaches that we have used may be extended to develop hybrid statistical-dynamical forecast tools (e.g., using dynamical forecast model output as predictor fields), which potentially can combine the advantages from both dynamical and statistical models to help improve subseasonal-to-seasonal forecast performance.

Acknowledgments. This study is supported by NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections Program Award NA14OAR4310189. The authors thank Dr. Peter Jan van Leeuwen, Dr. James Risbey, and an anonymous reviewer for their constructive suggestions and comments that resulted in the improved discussion and presentation of the results. The raw daily teleconnection pattern indices are provided by Climate Prediction Center. The NCEP– DOE Reanalysis 2 data are provided by the Physical Sciences Division from NOAA's Earth System Research Laboratory.

REFERENCES

- Abdi, H., 2010: Partial least squares regression and projection on latent structure regression (PLS Regression). Wiley Interdiscip. Rev.: Comput. Stat., 2, 97–106, doi:10.1002/wics.51.
- Abudu, S., J. P. King, and T. C. Pagano, 2010: Application of partial least-squares regression in seasonal streamflow forecasting. J. Hydrol. Eng., 15, 612–623, doi:10.1061/ (ASCE)HE.1943-5584.0000216.
- Alexander, M. A., U. S. Bhatt, J. E. Walsh, M. S. Timlin, J. S. Miller, and J. D. Scott, 2004: The atmospheric response to realistic Arctic sea ice anomalies in an AGCM during winter. *J. Climate*, **17**, 890–905, doi:10.1175/1520-0442(2004)017<0890: TARTRA>2.0.CO;2.
- Baldwin, M. P., and T. J. Dunkerton, 1999: Propagation of the Arctic oscillation from the stratosphere to the troposphere. J. Geophys. Res., 104, 30 937–30 946, doi:10.1029/1999JD900445.
- Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115**, 1083–1126, doi:10.1175/ 1520-0493(1987)115<1083:CSAPOL>2.0.CO;2.
- Black, R. X., 2002: Stratospheric forcing of surface climate in the Arctic Oscillation. J. Climate, 15, 268–277, doi:10.1175/ 1520-0442(2002)015<0268:SFOSCI>2.0.CO;2.
- Branstator, G., 1983: Horizontal energy propagation in a barotropic atmosphere with meridional and zonal structure. J. Atmos. Sci., 40, 1689–1708, doi:10.1175/1520-0469(1983)040<1689: HEPIAB>20.CO;2.
- —, 2002: Circumglobal teleconnections, the jet stream waveguide, and the North Atlantic Oscillation. J. Climate, 15, 1893–1910, doi:10.1175/1520-0442(2002)015<1893:CTTJSW>2.0.CO;2.
- Brönnimann, S., 2007: Impact of El Niño–Southern Oscillation on European climate. *Rev. Geophys.*, 45, RG3003, doi:10.1029/ 2006RG000199.
- Cassou, C., 2008: Intraseasonal interaction between the Madden-Julian oscillation and the North Atlantic oscillation. *Nature*, 455, 523–527, doi:10.1038/nature07286.
- Christiansen, B., 2005: Downward propagation and statistical forecast of the near-surface weather. J. Geophys. Res., 110, D14104, doi:10.1029/2004JD005431.
- DelSole, T., and J. Shukla, 2009: Artificial skill due to predictor screening. J. Climate, 22, 331–345, doi:10.1175/2008JCLI2414.1.
- —, L. Trenary, M. Tippett, and K. Pegion, 2017: Predictability of week-3–4 average temperature and precipitation over the contiguous United States. J. Climate, **30**, 3499–3512, doi:10.1175/JCLI-D-16-0567.1.
- Feldstein, S. B., 2000: The timescale, power spectra, and climate noise properties of teleconnection patterns. J. Climate, 13, 4430-4440, doi:10.1175/1520-0442(2000)013<4430:TTPSAC>2.0.CO;2.
- —, 2002: Fundamental mechanisms of the growth and decay of the PNA teleconnection pattern. *Quart. J. Roy. Meteor. Soc.*, **128**, 775–796, doi:10.1256/0035900021643683.

- Ferranti, L., T. N. Palmer, F. Molteni, and E. Klinker, 1990: Tropical–extratropical interaction associated with the 30–60 day oscillation and its impact on medium and extended range prediction. J. Atmos. Sci., 47, 2177–2199, doi:10.1175/ 1520-0469(1990)047<2177:TEIAWT>2.0.CO;2.
- Folland, C. K., A. A. Scaife, J. Lindesay, and D. B. Stephenson, 2012: How potentially predictable is northern European winter climate a season ahead? *Int. J. Climatol.*, **32**, 801–818, doi:10.1002/joc.2314.
- Franzke, C., S. B. Feldstein, and S. Lee, 2011: Synoptic analysis of the Pacific-North American teleconnection pattern. *Quart. J. Roy. Meteor. Soc.*, **137**, 329–346, doi:10.1002/ qj.768.
- Goss, M., and S. B. Feldstein, 2015: The impact of the initial flow on the extratropical response to Madden–Julian oscillation convective heating. *Mon. Wea. Rev.*, **143**, 1104–1121, doi:10.1175/ MWR-D-14-00141.1.
- Gouirand, I., V. Moron, and E. Zorita, 2007: Teleconnections between ENSO and North Atlantic in an ECHO-G simulation of the 1000-1990 period. *Geophys. Res. Lett.*, 34, L06705, doi:10.1029/2006GL028852.
- Higgins, R. W., and K. C. Mo, 1997: Persistent North Pacific circulation anomalies and the tropical intraseasonal oscillation. *J. Climate*, 10, 223–244, doi:10.1175/1520-0442(1997)010<0223: PNPCAA>2.0.CO;2.
- —, J.-K. E. Schemm, W. Shi, and A. Leetmaa, 2000: Extreme precipitation events in the western United States related to tropical forcing. *J. Climate*, **13**, 793–820, doi:10.1175/ 1520-0442(2000)013<0793:EPEITW>2.0.CO;2.
- Hoerling, M. P., A. Kumar, and T. Xu, 2001: Robustness of the nonlinear climate response to ENSO's extreme phases. *J. Climate*, 14, 1277–1293, doi:10.1175/1520-0442(2001)014<1277: ROTNCR>2.0.CO:2.
- Horel, J. D., and J. M. Wallace, 1981: Planetary-scale atmospheric phenomena associated with the Southern Oscillation. *Mon. Wea. Rev.*, **109**, 813–829, doi:10.1175/1520-0493(1981)109<0813: PSAPAW>2.0.CO;2.
- Hoskins, B. J., and D. J. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographical forcing. J. Atmos. Sci., 38, 1179–1196, doi:10.1175/ 1520-0469(1981)038<1179:TSLROA>2.0.CO;2.
- —, and T. Ambrizzi, 1993: Rossby wave propagation on a realistic longitudinally varying flow. J. Atmos. Sci., 50, 1661–1671, doi:10.1175/1520-0469(1993)050<1661:RWPOAR>2.0.CO;2.
- Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck, 2003: An overview of the North Atlantic Oscillation. *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, J. W. Hurrell et al., Eds., Amer. Geophys. Union, 1–35, doi:10.1029/134GM01.
- Johansson, Å., 2007: Prediction skill of the NAO and PNA from daily to seasonal time scales. J. Climate, 20, 1957–1975, doi:10.1175/JCLI4072.1.
- Johnson, N. C., and S. B. Feldstein, 2010: The continuum of North Pacific sea level pressure patterns: Intraseasonal, interannual, and interdecadal variability. J. Climate, 23, 851–867, doi:10.1175/2009JCLI3099.1.
 - —, and Y. Kosaka, 2016: The impact of eastern equatorial Pacific convection on the diversity of boreal winter El Niño teleconnection patterns. *Climate Dyn.*, **47**, 3737–3765, doi:10.1007/ s00382-016-3039-1.
- —, D. C. Collins, S. B. Feldstein, M. L. L'Heureux, and E. E. Riddle, 2014: Skillful wintertime North American temperature forecasts out to 4 weeks based on the state of ENSO and

the MJO. *Wea. Forecasting*, **29**, 23–38, doi:10.1175/ WAF-D-13-00102.1.

- Kalela-Brundin, M., 1999: Climatic information from tree-rings of Pinus sylvestris L. and a reconstruction of summer temperatures back to AD 1500 in Femundsmarka, eastern Norway, using partial least squares regression (PLS) analysis. *Holocene*, 9, 59–77, doi:10.1191/095968399678118795.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, 83, 1631–1643, doi:10.1175/BAMS-83-11-1631.
- Kang, D., and Coauthors, 2014: Prediction of the Arctic Oscillation in boreal winter by dynamical seasonal forecasting systems. *Geophys. Res. Lett.*, **41**, 3577–3585, doi:10.1002/2014GL060011.
- Karoly, D. J., R. A. Plumb, and M. Ting, 1989: Examples of the horizontal propagation of quasi-stationary waves. J. Atmos. Sci., 46, 2802–2811, doi:10.1175/1520-0469(1989)046<2802: EOTHPO>2.0.CO:2.
- Knutson, T. R., and K. M. Weickmann, 1987: 30–60 day atmospheric oscillations: Composite life cycles of convection and circulation anomalies. *Mon. Wea. Rev.*, **115**, 1407–1436, doi:10.1175/1520-0493(1987)115<1407:DAOCLC>2.0.CO:2.
- L'Heureux, M. L., and R. W. Higgins, 2008: Boreal winter links between the Madden–Julian oscillation and the Arctic Oscillation. J. Climate, 21, 3040–3050, doi:10.1175/2007JCLI1955.1.
- Li, Y., and N.-C. Lau, 2012: Impact of ENSO on the atmospheric variability over the North Atlantic in late winter—Role of transient eddies. J. Climate, 25, 320–342, doi:10.1175/ JCLI-D-11-00037.1.
- Liebmann, B., and C. A. Smith, 1996: Description of a complete outgoing longwave radiation dataset. *Bull. Amer. Meteor. Soc.*, **77**, 1275–1277.
- Lin, H., 2015: Subseasonal variability of North American wintertime surface air temperature. *Climate Dyn.*, 45, 1137–1155, doi:10.1007/s00382-014-2363-6.
- —, and G. Brunet, 2009: The influence of the Madden–Julian oscillation on Canadian wintertime surface air temperature. *Mon. Wea. Rev.*, **137**, 2250–2262, doi:10.1175/ 2009MWR2831.1.
- —, —, and J. Derome, 2009: An observed connection between the North Atlantic Oscillation and the Madden–Julian oscillation. J. Climate, 22, 364–380, doi:10.1175/2008JCLI2515.1.
- —, —, and R. Mo, 2010: Impact of the Madden–Julian oscillation on wintertime precipitation in Canada. *Mon. Wea. Rev.*, **138**, 3822–3839, doi:10.1175/2010MWR3363.1.
- Linkin, M. E., and S. Nigam, 2008: The North Pacific Oscillation– west Pacific teleconnection pattern: Mature-phase structure and winter impacts. J. Climate, 21, 1979–1997, doi:10.1175/ 2007JCLI2048.1.
- Luo, D., Y. Xiao, Y. Yao, A. Dai, I. Simmonds, and C. Franzke, 2016a: Impact of Ural blocking on winter warm Arctic–cold Eurasian anomalies. Part I: Blocking-induced amplification. *J. Climate*, **29**, 3925–3947, doi:10.1175/JCLI-D-15-0611.1.
- —, —, Y. Diao, A. Dai, C. Franzke, and I. Simmonds, 2016b: Impact of Ural blocking on winter warm Arctic–cold Eurasian anomalies. Part II: The link to the North Atlantic Oscillation. J. Climate, 29, 3949–3971, doi:10.1175/JCLI-D-15-0612.1.
- Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. J. Atmos. Sci., 28, 702–708, doi:10.1175/1520-0469(1971)028<0702: DOADOI>2.0.CO;2.
- —, and —, 1972: Description of global-scale circulation cells in the tropics with a 40–50 day period. J. Atmos.

Sci., **29**, 1109–1123, doi:10.1175/1520-0469(1972)029<1109: DOGSCC>2.0.CO;2.

- McIntosh, P., A. Ash, and M. Smith, 2005: From oceans to farms: The value of a novel statistical climate forecast for agricultural management. J. Climate, 18, 4287–4302, doi:10.1175/ JCLI3515.1.
- McKinnon, K. A., A. Rhines, M. P. Tingley, and P. Huybers, 2016: Long-lead predictions of eastern United States hot days from Pacific sea surface temperatures. *Nat. Geosci.*, 9, 389–394, doi:10.1038/ngeo2687.
- Merkel, U., and M. Latif, 2002: A high resolution AGCM study of the El Niño impact on the North Atlantic/European sector. *Geophys. Res. Lett.*, 29, doi:10.1029/2001GL013726.
- Michaelsen, J., 1987: Cross-validation in statistical climate forecast models. J. Climate Appl. Meteor., 26, 1589–1600, doi:10.1175/ 1520-0450(1987)026<1589:CVISCF>2.0.CO;2.
- Mo, K. C., and R. W. Higgins, 1998: Tropical convection and precipitation regimes in the western United States. *J. Climate*, **11**, 2404–2423, doi:10.1175/1520-0442(1998)011<2404: TCAPRI>2.0.CO:2.
- Moon, J.-Y., B. Wang, and K.-J. Ha, 2011: ENSO regulation of MJO teleconnection. *Climate Dyn.*, **37**, 1133–1149, doi:10.1007/s00382-010-0902-3.
- Moore, R. W., O. Martius, and T. Spengler, 2010: The modulation of the subtropical and extratropical atmosphere in the Pacific basin in response to the Madden–Julian oscillation. *Mon. Wea. Rev.*, 138, 2761–2779, doi:10.1175/2010MWR3194.1.
- Mori, M., and M. Watanabe, 2008: The growth and triggering mechanisms of the PNA: A MJO-PNA coherence. J. Meteor. Soc. Japan, 86, 213–236, doi:10.2151/jmsj.86.213.
- Moron, V., and I. Gouirand, 2003: Seasonal modulation of the El Niño-Southern Oscillation relationship with sea level pressure anomalies over the North Atlantic in October–March (1873–1996). Int. J. Climatol., 23, 143–155, doi:10.1002/ joc.868.
- Nigam, S., 2003: Teleconnection. *Encyclopedia of Atmospheric Sciences*, J. R. Holton, J. A. Pyle, and J. A. Curry, Eds., Elsevier Science, 2243–2269.
- —, and S. Baxter, 2015: Teleconnections. *Encyclopedia of Atmospheric Sciences*, 2nd ed. G. North, Ed., Elsevier Science, 90–109, doi:10.1016/B978-0-12-382225-3.00400-X.
- Norton, W. A., 2003: Sensitivity of northern hemisphere surface climate to simulation of the stratospheric polar vortex. *Geophys. Res. Lett.*, **30**, 1627, doi:10.1029/2003GL016958.
- Pozo-Vázquez, D., M. J. Esteban-Parra, F. S. Rodrigo, and Y. Castro-Diaz, 2001: The association between ENSO and winter atmospheric circulation and temperature in the North Atlantic region. J. Climate, 14, 3408–3420, doi:10.1175/ 1520-0442(2001)014<3408:TABEAW>2.0.CO;2.
- —, S. R. Gámiz-Fortis, J. Tovar-Pescador, M. J. Esteban-Parra, and Y. Castro-Diaz, 2005: North Atlantic winter SLP anomalies based on the autumn ENSO state. J. Climate, 18, 97–103, doi:10.1175/JCLI-3210.1.
- Riddle, E. E., M. B. Stoner, N. C. Johnson, M. L. L'Heureux, D. C. Collins, and S. B. Feldstein, 2013: The impact of the MJO on clusters of wintertime circulation anomalies over the North American region. *Climate Dyn.*, 40, 1749–1766, doi:10.1007/ s00382-012-1493-y.
- Risbey, J., T. O'Kane, D. Monselesan, C. Franzke, and I. Horenko, 2015: Metastability of Northern Hemisphere teleconnection modes. J. Atmos. Sci., 72, 35–54, doi:10.1175/JAS-D-14-0020.1.
- Rodney, M., H. Lin, and J. Derome, 2013: Subseasonal prediction of wintertime North American surface air temperature

during strong MJO events. *Mon. Wea. Rev.*, **141**, 2897–2909, doi:10.1175/MWR-D-12-00221.1.

- Rodwell, M. J., and C. K. Folland, 2002: Atlantic air-sea interaction and seasonal predictability. *Quart. J. Roy. Meteor. Soc.*, **128**, 1413–1443, doi:10.1002/qj.200212858302.
- —, D. P. Rowell, and C. K. Folland, 1999: Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature*, **398**, 320–323, doi:10.1038/18648.
- Roundy, P. E., K. MacRitchie, J. Asuma, and T. Melino, 2010: Modulation of the global atmospheric circulation by combined activity in the Madden–Julian oscillation and the El Niño– Southern Oscillation during boreal winter. J. Climate, 23, 4045–4059, doi:10.1175/2010JCLI3446.1.
- Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. J. Climate, 27, 2185–2208, doi:10.1175/ JCLI-D-12-00823.1.
- Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor, 2000: Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *J. Geophys. Res.*, **105**, 7337–7356, doi:10.1029/1999JD901105.
- Scaife, A. A., and J. R. Knight, 2008: Ensemble simulations of the cold European winter of 2005–2006. *Quart. J. Roy. Meteor. Soc.*, **134**, 1647–1659, doi:10.1002/qj.312.
- —, —, G. K. Vallis, and C. K. Folland, 2005: A stratospheric influence on the winter NAO and North Atlantic surface climate. *Geophys. Res. Lett.*, **32**, L18715, doi:10.1029/ 2005GL023226.
- —, and Coauthors, 2014: Skillful long-range prediction of European and North American winters. *Geophys. Res. Lett.*, 41, 2514–2519, doi:10.1002/2014GL059637.
- Seo, K.-H., and S.-W. Son, 2012: The global atmospheric circulation response to tropical diabatic heating associated with the Madden–Julian oscillation during northern winter. J. Atmos. Sci., 69, 79–96, doi:10.1175/2011JAS3686.1.
- Smith, D. M., A. A. Scaife, R. Eade, and J. R. Knight, 2016: Seasonal to decadal prediction of the winter North Atlantic Oscillation: Emerging capability and future prospects. *Quart. J. Roy. Meteor. Soc.*, **142**, 611–617, doi:10.1002/ qj.2479.
- Smoliak, B. V., J. M. Wallace, M. T. Stoelinga, and T. P. Mitchell, 2010: Application of partial least squares regression to the diagnosis of year-to-year variations in Pacific Northwest snowpack and Atlantic hurricanes. *Geophys. Res. Lett.*, 37, L03801, doi:10.1029/2009GL041478.
- —, —, P. Lin, and Q. Fu, 2015: Dynamical adjustment of the Northern Hemisphere surface air temperature field: Methodology and application to observations. J. Climate, 28, 1613–1629, doi:10.1175/JCLI-D-14-00111.1.
- Song, Y., and W. A. Robinson, 2004: Dynamical mechanisms for stratospheric influences on the troposphere. J. Atmos. Sci., 61, 1711–1725, doi:10.1175/1520-0469(2004)061<1711: DMFSIO>2.0.CO;2.
- Takaya, K., and H. Nakamura, 2001: A formulation of a phaseindependent wave-activity flux for stationary and migratory quasigeostrophic eddies on a zonally varying basic flow. J. Atmos. Sci., 58, 608–627, doi:10.1175/1520-0469(2001)058<0608: AFOAPI>2.0.CO:2.
- Teng, H., and G. Branstator, 2017: Causes of extreme ridges that induce California droughts. J. Climate, 30, 1477–1492, doi:10.1175/JCLI-D-16-0524.1.
- —, —, H. Wang, G. A. Meehl, and W. M. Washington, 2013: Probability of US heat waves affected by a subseasonal

planetary wave pattern. Nat. Geosci., 6, 1056–1061, doi:10.1038/ ngeo1988.

- Thompson, D. W., and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, 25, 1297–1300, doi:10.1029/ 98GL00950.
- Toniazzo, T., and A. A. Scaife, 2006: The influence of ENSO on winter North Atlantic climate. *Geophys. Res. Lett.*, 33, L24704, doi:10.1029/2006GL027881.
- Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.*, 103, 14291–14324, doi:10.1029/97JC01444.
- Vecchi, G. A., and N. A. Bond, 2004: The Madden-Julian Oscillation (MJO) and northern high latitude wintertime surface air temperatures. *Geophys. Res. Lett.*, **31**, L04104, doi:10.1029/ 2003GL018645.
- Vitart, F., and F. Molteni, 2010: Simulation of the Madden–Julian oscillation and its teleconnections in the ECMWF forecast system. *Quart. J. Roy. Meteor. Soc.*, **136**, 842–855, doi:10.1002/ qj.623.
- Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon.*

Wea. Rev., **109**, 784–812, doi:10.1175/1520-0493(1981)109<0784: TTTGHF>2.0.CO;2.

- —, Q. Fu, B. V. Smoliak, P. Lin, and C. M. Johanson, 2012: Simulated versus observed patterns of warming over the extratropical Northern Hemisphere continents during the cold season. *Proc. Natl. Acad. Sci. USA*, **109**, 14337–14342, doi:10.1073/pnas.1204875109.
- Wold, H., 1966: Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*, P. R. Krishnaiah, Ed., Academic, 391–420.
- Xiang, B., M. Zhao, X. Jiang, S. J. Lin, T. Li, X. Fu, and G. A. Vecchi, 2015: The 3–4-week MJO prediction skill in a GFDL coupled model. J. Climate, 28, 5351–5364, doi:10.1175/ JCLI-D-15-0102.1.
- Yao, W., H. Lin, and J. Derome, 2011: Submonthly forecasting of winter surface air temperature in North America based on organized tropical convection. *Atmos.–Ocean*, **49**, 51–60, doi:10.1080/07055900.2011.556882.
- Yoo, C., S. Lee, and S. B. Feldstein, 2012a: Arctic response to an MJO-like tropical heating in an idealized GCM. J. Atmos. Sci., 69, 2379–2393, doi:10.1175/JAS-D-11-0261.1.
- —, —, and —, 2012b: Mechanisms of Arctic surface air temperature change in response to the Madden–Julian oscillation. J. Climate, 25, 5777–5790, doi:10.1175/JCLI-D-11-00566.1.