

Subseasonal Prediction of Wintertime East Asian Temperature Based on Atmospheric Teleconnections

CHANGHYUN YOO

Department of Climate and Energy Systems Engineering, Ewha Womans University, Seoul, South Korea

NATHANIEL C. JOHNSON

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

CHUEH-HSIN CHANG

Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan

STEVEN B. FELDSTEIN

Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, Pennsylvania

YOUNG-HA KIM

Severe Storm Research Center, Ewha Womans University, Seoul, South Korea

(Manuscript received 30 November 2017, in final form 14 September 2018)

ABSTRACT

A composite-based statistical model utilizing Northern Hemisphere teleconnection patterns is developed to predict East Asian wintertime surface air temperature for lead times out to 6 weeks. The level of prediction is determined by using the Heidke skill score. The prediction skill of the statistical model is compared with that of hindcast simulations by a climate model, Global Seasonal Forecast System, version 5. When employed individually, three teleconnections (i.e., the east Atlantic/western Russian, Scandinavian, and polar/Eurasian teleconnection patterns) are found to provide skillful predictions for lead times beyond 4–5 weeks. When information from the teleconnections and the long-term linear trend are combined, the statistical model outperforms the climate model for lead times beyond 3 weeks, especially during those times when the teleconnections are in their active phases.

1. Introduction

There has been rising interest in subseasonal weather and climate prediction for lead times beyond two weeks but less than a month, which lie between the time scales of short-range weather forecasts and long-range monthly prediction. Up to approximately 7–10 days, weather forecasts heavily rely on predictions from numerical weather prediction models, which are formulated by the physical laws and the subgrid-scale parameterizations of the atmosphere. However, the predictability of such a nonlinear forecast system diminishes rapidly due to inevitable errors in the initial conditions, along with

imperfect representations of the atmospheric processes in the model. Theoretically, the nonlinearity of the atmosphere hinders a deterministic prediction beyond approximately 2 weeks (Lorenz 1963), and operational centers push the limit by employing ensemble prediction systems that generate a range of possible outcomes. At longer lead times beyond one month, slowly varying boundary conditions, such as sea surface temperature (SST) and land surface conditions, allow us to assemble a probabilistic forecast on the statistics of weather and hence on climate. At this time scale, both ensemble forecasts using dynamical models and statistical models based on empirical relationships produce skillful predictions over some regions. For subseasonal prediction, however, the combination of initial condition error and

Corresponding author: Changhyun Yoo, cyoo@ewha.ac.kr

the brief time for boundary conditions to take effect makes skillful predictions a challenging problem both for operational centers and for the research community. Nonetheless, the importance of accurate subseasonal prediction has increased due to growing concerns of frequent extreme weather events that make considerable societal and economic impacts.

Despite the challenges of subseasonal prediction, hope may be found in large-scale modes of climate variability, frequently recurring patterns that explain a large portion of climate variance, which are a potential source of skill targeting lead time scales longer than two weeks. For North American subseasonal prediction, the climate modes of particular interest are those that moderate large-scale convection in the tropics. This includes El Niño–Southern Oscillation (ENSO) and the Madden–Julian oscillation (MJO; Madden and Julian 1971, 1972). As the most pronounced pattern of interannual tropical SST variability, ENSO plays an important role in determining the strength of the seasonal mean convection anomalies over the tropical Pacific. For example, during the positive phase of ENSO (El Niño), a dipole anomaly of enhanced and reduced convection can be seen over the central and western Pacific, respectively (e.g., Fig. 1b in Son et al. 2017). A similar spatial pattern of convection anomalies with the opposite sign can be seen during the negative ENSO phase (La Niña). At subseasonal time scales, the MJO is associated with tropical convection anomalies that resemble those associated with ENSO. With a period of about 30–60 days, the MJO convection, coupled with the circulation field, circumnavigates the equator. Consistent with the finding that diabatic heating associated with tropical convection can excite Rossby waves that propagate from the tropics into the extratropics (e.g., Hoskins and Karoly 1981), it has been shown that the responses to both ENSO (e.g., Horel and Wallace 1981; Johnson and Feldstein 2010) and the MJO (e.g., Higgins and Mo 1997; Cassou 2008; Lin et al. 2009; Riddle et al. 2013; Yoo et al. 2015) have a significant impact on the extratropical circulation and surface air temperature (SAT) (Halpert and Ropelewski 1992; Trenberth and Caron 2000; Vecchi and Bond 2004; Yoo et al. 2012; Lin 2015; DelSole et al. 2017) on the subseasonal time scale. These sources of subseasonal predictability take root in the features that poleward propagating Rossby wave activity, excited by tropical convection, takes about one week to travel from the tropics to high latitudes (Hoskins and Karoly 1981), that Rossby waves triggered by tropical convection will persist for 2 to 3 weeks even after the convection ceases (Branstator 2014), and that the dominant low-frequency (period greater than 10 days) extratropical teleconnection patterns,

such as the North Atlantic Oscillation (NAO) and Pacific–North American (PNA) patterns, typically persist for about two weeks (Feldstein 2000), even in the absence of strong tropical convection (Dai et al. 2017).

Motivated by the finding that ENSO and the MJO excite long-lived extratropical teleconnection patterns, Johnson et al. (2014) constructed a statistical forecast model for North American wintertime SAT with lead times up to 6 weeks. The model relies on the composite-based relationships between the extratropical SAT and the above-mentioned two climate modes (ENSO and the MJO) and additionally the long-term linear SAT trend. The phase information of the climate modes is the key to capture changes in the probability distribution of the extratropical SAT response. It turns out that the skill of the statistical model during some phases of ENSO and the MJO exceeds the typical skill of dynamical models for lead times beyond 3–4 weeks. This composite-based model is convenient for investigating subsamples, such as strong MJO events, which may provide increased forecast skill relative to moderate events. Their model and a similar variant currently provide guidance for the 3–4-week operational predictions of the SAT by the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) (<http://www.cpc.ncep.noaa.gov/products/predictions/WK34/>).

Although the connections between both ENSO and MJO and North American climate are well established, it remains to be determined if such subseasonal skill can be realized over East Asia. Building upon Johnson et al. (2014), this study constructs a composite-based statistical model for East Asian wintertime SAT with lead times out to 6 weeks. Here, we attempt to obtain skill from the dominant low-frequency extratropical atmospheric teleconnection patterns. The rationale behind this approach is as follows:

- 1) Low-frequency variability of the atmosphere can be predominantly described in terms of recurrent and persistent teleconnection patterns (Wallace and Gutzler 1981; Barnston and Livezey 1987). Teleconnection patterns may not be ideal candidates for predictions out to 6 weeks because teleconnections persist for about 2 weeks in the absence of strong tropical convection, as discussed above, and one week longer for the PNA when there is strong convection (Dai et al. 2017). However, the amplitude and influence of the teleconnection patterns may extend beyond their decorrelation time scales [i.e., ~ 1 – 2 weeks, measured by the e -folding time scales of autocorrelations, which is a measure of linear

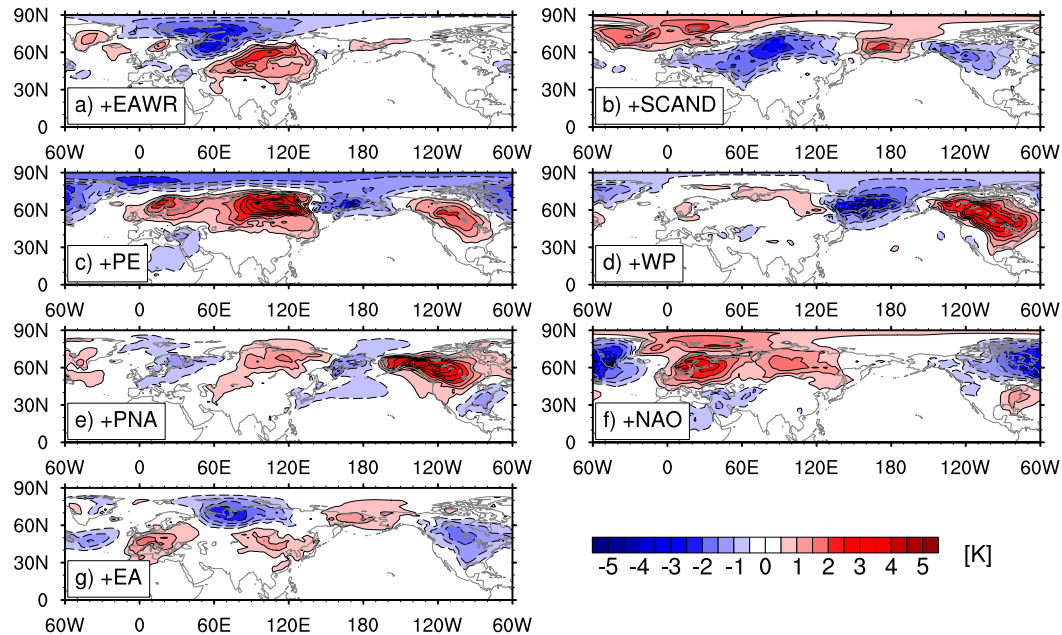


FIG. 1. Composites of the DJF 2-m temperature anomalies for the positive phases of seven teleconnection patterns: (a) EAWR, (b) SCAND, (c) EA, (d) WP, (e) PNA, (f) NAO, and (g) PE. Solid (dashed) contours are positive (negative), and the zero contours are omitted. The contour interval is 0.5 K. The positive phase is defined when the value of an index exceeds 0.5 standard deviations.

temporal persistence, as in Table 2 of [Feldstein \(2000\)](#). As will be shown, this is indeed the case as the surface footprint of the teleconnection patterns remains out to 3–4 weeks or even longer. This suggests that the initial state of teleconnection patterns can provide information for subseasonal forecasts with a lead time of several weeks ahead.

- 2) The geographical location of East Asia may not be ideally suited to benefit from ENSO or the MJO, although a recent study finds the MJO influencing East Asia via a local Hadley circulation ([Seo et al. 2016](#)). The poleward propagating Rossby waves, excited by ENSO or MJO, propagate downstream to North America. This location of these Rossby waves is influenced by the strong potential vorticity gradient associated with the East Asian jet, which acts as a waveguide ([Branstator 2002](#)), while at the same time playing an important role in contributing to the Rossby wave source ([Sardeshmukh and Hoskins 1988](#)). In fact, an analysis of the low-frequency vorticity equation budget associated with the MJO suggests that the Rossby waves are initially triggered near the jet center over the northwestern subtropical Pacific but escape to the extratropics near the central Pacific at the jet exit region ([Lukens et al. 2017](#)). We will show that for East Asian SAT, the prediction skill tends to originate from the teleconnection patterns located upstream.

Therefore, in the present study, we investigate the level of prediction skill that can be obtained by employing Northern Hemisphere teleconnection patterns, targeting the East Asian wintertime [December–February (DJF)] SAT, into a statistical forecast model. Consistent with the results of [Johnson et al. \(2014\)](#), we show that the statistical model can outperform a dynamical model at lead times exceeding 3 weeks. The present study demonstrates that some individual teleconnection indices can provide substantial skill for East Asia. In addition, we show that a linear combination of the teleconnection patterns, as well as the linear trend, can enhance forecast skill.

The remainder of the study is organized as follows. In [section 2](#), the data and methods that are used to produce and evaluate the forecasts are presented. [Section 3](#) illustrates the weekly SAT forecasts from the dynamical and cross-validated statistical models. A discussion and the conclusions are presented in [section 4](#).

2. Data and methods

a. Data for the statistical model

To build the statistical model, we use the daily SAT and 500-hPa geopotential height of the European Centre for Medium-Range Weather Forecasts (ECMWF)

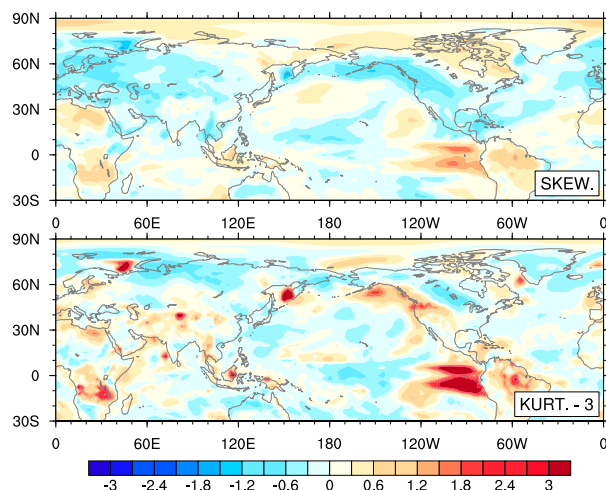


FIG. 2. Skewness and kurtosis map of the DJF 7-day running averaged SAT anomaly distribution of the daily ERA-Interim from 1979–2014. Value 3 is subtracted from the kurtosis so that value zero indicates a mesokurtotic distribution.

interim reanalysis (ERA-Interim; Dee et al. 2011) dataset for the period from 1979 through 2014. The data, which have a native horizontal resolution of ~ 60 km, have been regridded to a 2.5° latitude \times 2.5° longitude grid. The error due to the spatial interpolation is documented in appendix A. Anomalies are computed at each grid point by subtracting the seasonal cycle, where the

seasonal cycle is defined as the first four harmonics of the calendar day means. For the SAT, to remove the variability at the weather time scale, a 7-day running average is applied to the anomalies.

The Northern Hemisphere teleconnection patterns are obtained from the NOAA CPC website (<http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml>). The patterns are defined through the use of rotated principal component analysis on the monthly mean standardized 500-hPa geopotential height anomalies for 20° – 90° N and the years 1950–2000. Note that some patterns of the CPC, such as the NAO and PNA, are robustly reproduced when analyzed using the period of 1979–2014 (not shown). However, the other patterns of the CPC show only partial matches to those from the recent period, presumably due to a changing climate. Nonetheless, for reproducibility and operational use of our results, we choose to use the CPC indices. Furthermore, the sensitivity to the number of retained unrotated EOFs was examined. It was found that 7 similar rotated EOFs were found when 8, 10, and 12 unrotated EOFs were used. Also, note that for each teleconnection pattern, the same spatial pattern with opposite sign is used for the positive and negative phases. This may contribute to a linear bias in the statistical relationship in the HSS values for various lead times.

Among the 10 teleconnection patterns monitored by the NOAA/CPC, we explore the seven patterns that

DJF HSS (GloSea5)

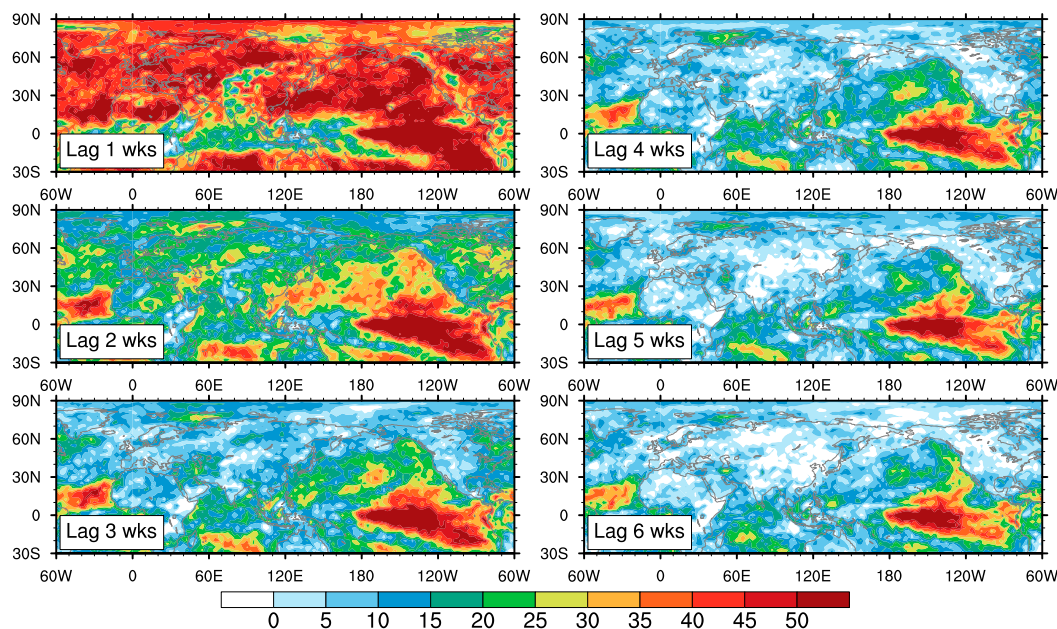


FIG. 3. The HSS of the 7-day running averaged SAT anomalies in GloSea5 hindcast experiments evaluated during DJF for each grid point. Forecast lead times from lag +1 week to lag +6 weeks are shown. The shading interval is 5.

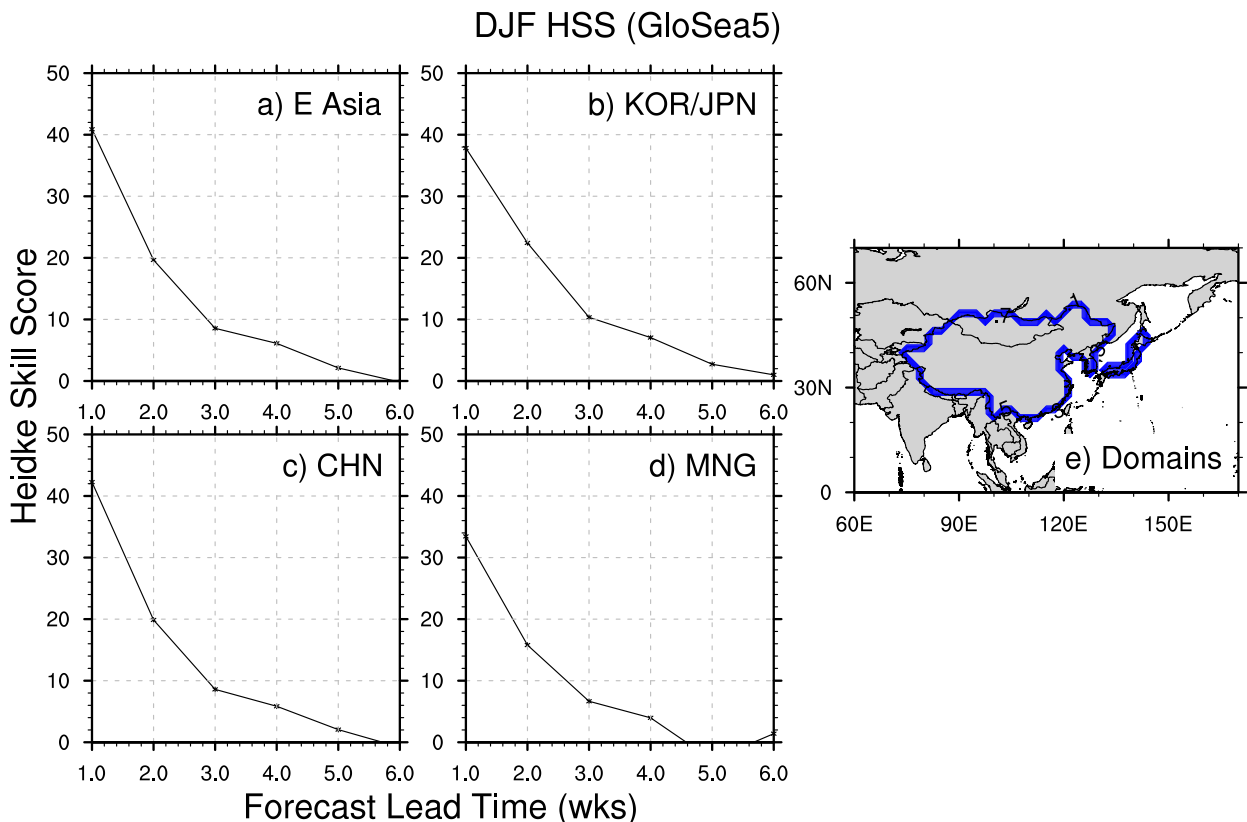


FIG. 4. The DJF HSS of GloSea5 hindcast simulations for lead times of 1–6 weeks, shown in Fig. 3, is averaged over (a) East Asia and its three subdomains: (b) the Korean peninsula and Japan, (c) China, and (d) Mongolia. (e) The domain for all of East Asia is marked with a blue contour.

can be defined in all seasons. These are the east Atlantic (EA), east Atlantic/western Russia (EAWR), NAO, polar/Eurasian (PE), PNA, Scandinavia (SCAND), and west Pacific (WP) teleconnection patterns. Daily indices are available only for the NAO and PNA, while for other patterns indices are provided as monthly values. Therefore, we compute daily indices by projecting the daily 500-hPa geopotential height anomalies onto the corresponding monthly teleconnection patterns, where the patterns are defined for each month by generating correlation maps between the monthly indices and the monthly 500-hPa geopotential height anomalies at each grid point. When tested for the NAO and PNA, this temporal downscaling technique produces daily indices that show linear correlations of about 0.8 with those of the CPC, which suggests roughly one-third of the CPC index variance is unexplained by our downscaled indices (appendix B).

For all teleconnection patterns, the positive (negative) phase is defined to have taken place when the index is greater (less) than 0.5 (−0.5) standard deviations. When the index is in between these values, it is

categorized as the neutral phase. This threshold of 0.5 standard deviations is subjectively chosen to balance the sample size with the signal amplitude and has been used in previous studies to define phases (e.g., Nakamura and Wallace 1990). We have used terciles to evenly distribute the number of days in each phase, and the results are similar to those of the 0.5 standard deviation thresholds (not shown). Figure 1 illustrates the positive phase composite of the DJF SAT anomalies for each teleconnection index. One can notice that the positive phases of the EAWR, SCAND, and PE show large SAT anomalies over East Asia. The readers are referred to the NOAA/CPC website for the geopotential height fields associated with the teleconnection patterns.

b. Hindcast simulations of GloSea5

To compare the statistical model performance with that of a dynamical model, we examine hindcast simulations of the Global Seasonal Forecast System, version 5 (GloSea5), a global atmosphere–land–ocean–sea ice coupled model (MacLachlan et al. 2015) composed of HadGEM3, JULES, and NEMO of the UK Met Office,

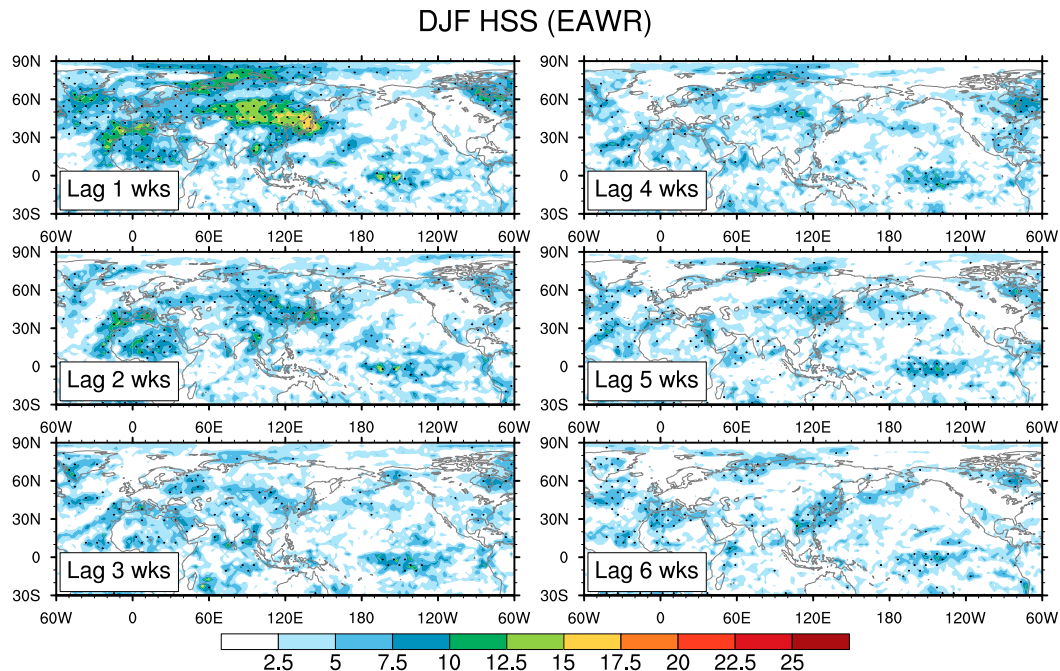


FIG. 5. As in Fig. 3, but for the HSS obtained from the statistical model using the initial state of the EAWR teleconnection pattern for six different weekly lead times during DJF. The shading interval is 2.5. Stippling indicates the values that exceed the 95% confidence level.

as well as CICE of Los Alamos National Laboratory. Three 60-day ensemble simulations, which are initialized on days 1, 9, 17, and 25 of each month from 1991 through 2010, are provided by the Korea Meteorological Administration. This 20-yr time period is shorter than the 36-yr time period of the statistical model. However, we note that the results of the statistical model in this study are similar when evaluated for the shorter time period. That is, the pattern correlation between the statistical model results for 1979–2014 (which are shown later in Fig. 5) and the same model for 1991–2010 exceeds the value of 0.72 throughout all lead times.

The atmosphere model outputs, which have a horizontal resolution of 0.56° latitude and 0.83° longitude, are interpolated onto the 2.5° horizontal resolution grid for comparison with the ERA-Interim data. The model bias is corrected by subtracting the difference between the lead-dependent model climatology and the ERA-Interim climatology of the same period. A 7-day running average is applied to the bias-corrected anomalies and for anomalies in the ERA-Interim data.

c. The composite-based model

Following Johnson et al. (2014), we first find the upper and lower terciles at each grid point in order to categorize above and below average SATs. The terciles are defined by the SAT anomaly distribution for each

calendar day with a 21-calendar-day window, where the window includes 10 calendar days before and after the target calendar day. Here the window size is chosen to be more generous than the 7-day window chosen in Johnson et al. (2014) in order to allow for a greater sample size, but the results of this study are not sensitive to the size of the window.

The next step is to identify the changes in the SAT probability distribution associated with the initial state of the climate modes for six weekly lags. For this a calculation, we use all days that the teleconnection indices have an amplitude that exceeds 0.5 standard deviations. To do so, we calculate the lagged response in the mean and variance of the SAT anomalies associated with each climate mode for each grid point and time lag with the 21-calendar day window. That is, for a given climate mode in the initial state, we generate the probabilistic forecast by producing a Gaussian distribution based on the calculated mean and variance for each grid point and time lag. The Gaussian assumption may break down over regions where the SAT distribution is considerably skewed, and/or heavily or lightly tailed. This includes the U.S. Pacific Northwest and Europe, where the SAT distribution is negatively skewed, and the tropical eastern Pacific, where the distribution is positively skewed (Fig. 2). Large values of kurtosis also can be seen in the skewed regions. Over large areas of East Asia, it can be

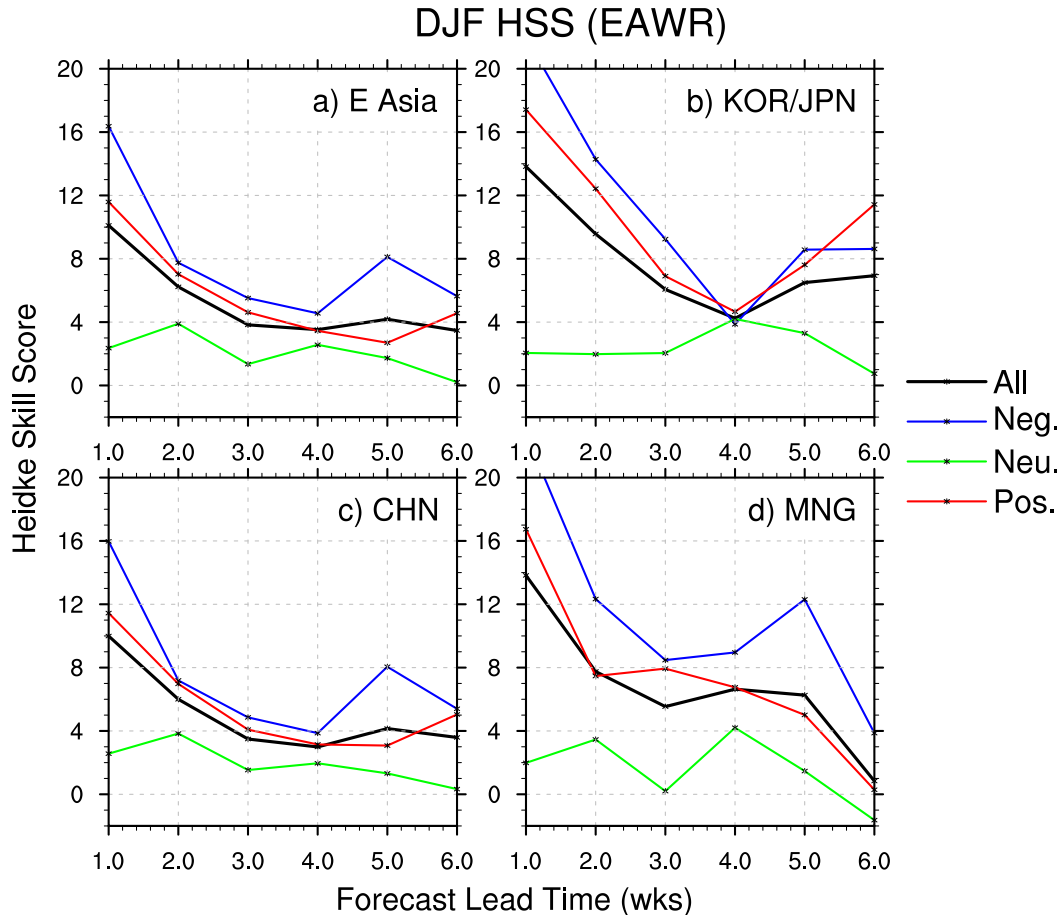


FIG. 6. As in Figs. 4a–d, with DJF HSS averaged over the four domains (black lines), but the HSS values are obtained from the statistical model for the EAWR teleconnection pattern, shown in Fig. 5. The domain averages are also separately performed for the positive (red), neutral (green), and negative (blue) EAWR phases.

seen that the SAT distribution deviates from Gaussian. However, an advantage of this Gaussian assumption is that the probability distribution can be stored and reconstructed using the two parameters. For the leave-one-year-out cross-validation approach, the statistics are calculated for every year using all data excluding the chosen year. The predictions are then made for the year that was left out of the training data.

The Heidke skill score (HSS) is used to evaluate the probabilistic forecasts. The HSS is a common metric used by operational forecast centers such as NOAA CPC for probabilistic forecast verification. Using the terciles and probabilistic distributions, we assign one of the three forecast categories (top, middle, or bottom terciles) based on the highest of the three forecast probabilities. The HSS is determined by the number of categories forecasted correctly H , the expected number of categories forecasted correctly just by chance E , and the total number of forecasts T , and is expressed as

$$\text{HSS} = \frac{H - E}{T - E} \times 100 (\%).$$

For a random forecast, E has an expected value of $T/3$. Extreme values of HSS are -50 (all forecasts being wrong) and 100 (a perfect set of forecasts). An HSS value of 0 corresponds to the expected score from a random set of forecasts, and positive values indicate some level of skill. We note that when computed using the joint distribution of forecasts and observations, which does not use the assumption of $E = T/3$ [e.g., Eq. (8.22) in Wilks (2011)], the HSS values in our results are robust (i.e., the pattern correlations between skill maps are ~ 0.98 for all lead times). Also, the Peirce skill score (PSS), a score similar to the HSS but formulated to have an unbiased denominator, shows only negligible differences from the HSS (not shown). The pattern correlations between the HSS and PSS are again ~ 0.98 for all lead times.

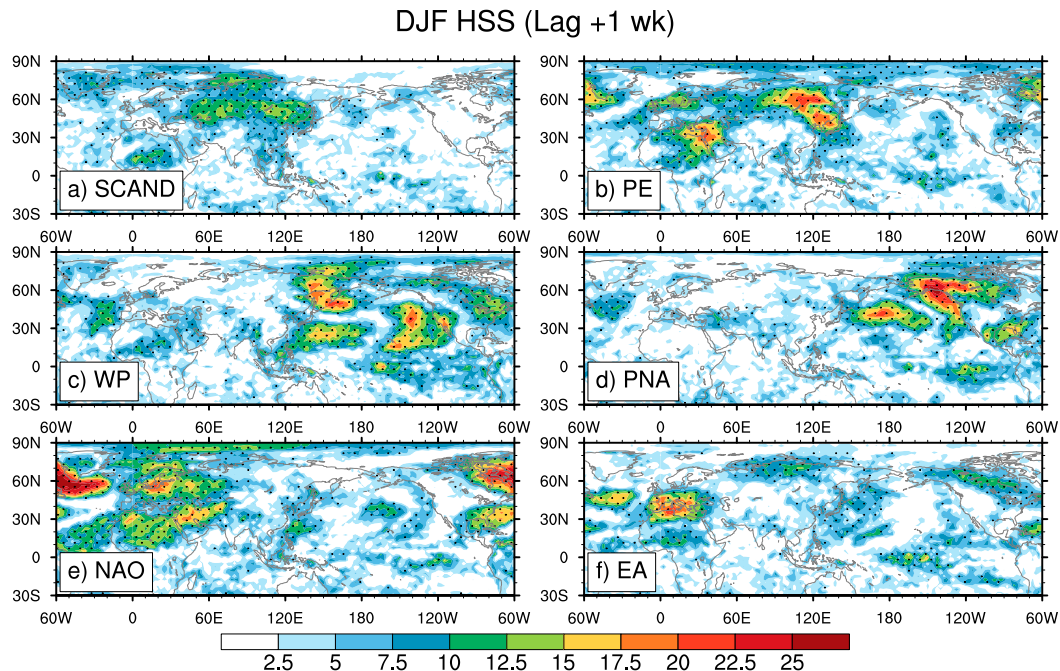


FIG. 7. As in Fig. 5a, the DJF HSSs for lag +1 week are shown except that the statistical models here depend on the initial state of the SCAND, EA, WP, PNA, NAO, and PE teleconnection patterns, respectively. The shading interval is 2.5. As in Fig. 5, stippling is used to indicate the values that exceed the 95% confidence level.

3. Results

a. The wintertime HSS of GloSea5

We start by examining the HSS of the GloSea5 hind-cast experiments for lead times of 1–6 weeks (Fig. 3). This establishes a skill baseline to place the statistical forecast model results in context. Here and in all subsequent figures, we focus on the skill during DJF. The forecast categories for weekly lead times are determined from the ensemble averages of the three deterministic forecasts.

Figure 3 illustrates the simple fact that the SAT in the dynamical model is highly forecastable for the first few weeks, but the forecastability quickly degrades with time. This is particularly true for SAT over land where the time scale of atmospheric processes dominates. In contrast, over the ocean, the HSS tends to persist longer, probably owing to the longer memory of the initial state. For example, the value of the HSS at lag +1 week exceeds 50 over large areas of land, but at lag +6 weeks, the HSS drops to approximately 0 over most of land areas, indicating that at this forecast time the model's prediction is no better than a random guess. In contrast, there are large swaths of ocean that retain large HSS values at lag +6 weeks, with the HSS showing its largest persistence over the tropical eastern Pacific. For lags beyond 3 weeks, the HSS values are the smallest in the

midlatitudes. One possible contribution to this result is the fact that the signal of the MJO is stronger in the tropics and Arctic than it is in the midlatitudes (Yoo et al. 2014).

When the HSS is averaged over the domains of our interest, a dramatic decline of the HSS can be clearly seen (Figs. 4a–d). The domain of East Asia and its three subdomains (i.e., the Korean Peninsula and Japan, China, and Mongolia) do not include grid points over ocean and are shown by blue contours in Fig. 4e. In all four domains, the HSS values start roughly near 40 at lag +1 week, and they quickly decline to near 20 at lag +2 weeks. Eventually, all HSS values approach 0 in all domains by lags +5 and +6 weeks. We note that the values and their behaviors with time are similar to that of CFSv2 over North America (Fig. 3c in Johnson et al. 2014). The rapid decline of the dynamical model skill is an indicator of why a statistical model can be competitive with or even superior to the dynamical model for lead times of 3 weeks and beyond, as the HSS values range between 4 and 7 at those lead time scales depending on the activity of the ENSO and MJO (Figs. 3a,b in Johnson et al. 2014).

b. Individual teleconnection indices

In this subsection, we present the statistical model results based on each teleconnection pattern in Fig. 1.

DJF HSS over East Asia

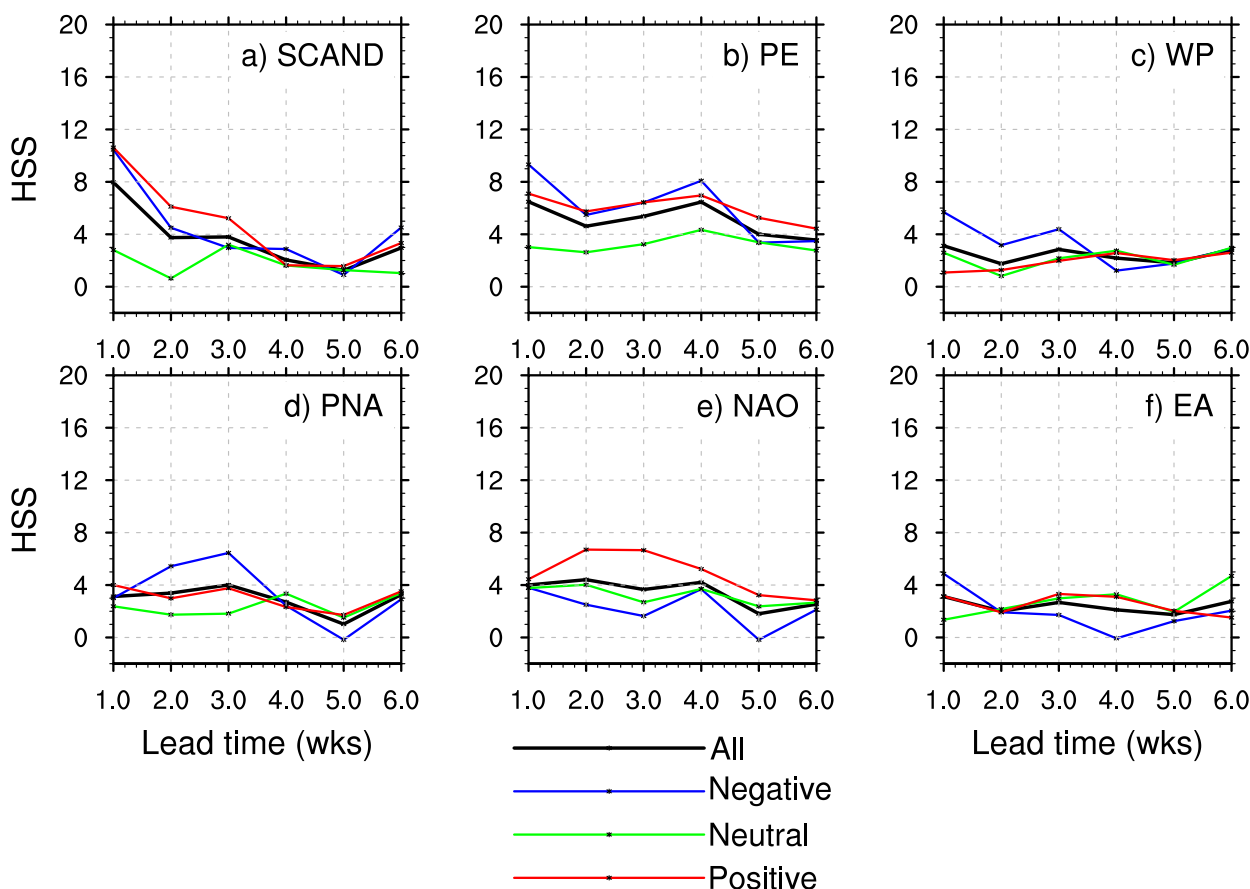


FIG. 8. As in Fig. 6a, the DJF HSSs averaged for the lead times of 1–6 weeks over East Asia (black) except that the results are for the following six teleconnection patterns: SCAND, EA, WP, PNA, NAO, and PE. The domain averages are performed for the positive (red), neutral (green), and negative (blue) phases of the corresponding patterns.

We first carefully describe the EAWR results for all six lags. Then the results of the rest of the teleconnection patterns are shown.

Figure 5 illustrates the enhanced prediction skill using only the initial state of the EAWR as the predictor, averaged for all phases. The HSS is averaged for DJF as in the dynamical model results in Fig. 3. The statistical significance of the HSS values is examined at the 95th percentile using Monte Carlo resampling at each grid point. To preserve the autocorrelation, the resampling is performed by reshuffling individual seasons, instead of individual days. The HSS values are computed 500 times with randomly resampled forecasts and verification fields. The p values of the Monte Carlo test averaged over East Asia range from 0.03 to 0.18 for the EAWR during the 6-week lead times. A quick visual inspection makes it clear that the signal of the EAWR is apparently visible over central and East Asia and the Middle East, coinciding with the location of the EAWR mode (Fig. 1a).

Near 50°N , 160°E , the maximum HSS values at lag +1 week are above 15, and the enhanced skill remains quite persistent for 2–4 weeks over East Asia. At lag +5 and +6 weeks, the HSS values over East Asia are between 2.5 and 7.5. One can also see enhanced skill over the equatorial and subtropical eastern Pacific. This is possibly due to the fact that the EAWR is linked with ENSO. As a result, the skill for the eastern Pacific is quite high, with HSS values ranging from 2.5–10 at lag +1 week to 5–10 at lag +6 weeks.

When averaged over East Asia for all EAWR phases (black line in Fig. 6a), it can be seen that the EAWR provides potentially useful skill. All phases refer to both positive and negative phases and the neutral phase, where the neutral EAWR phase is defined, as mentioned in section 2, when the index falls between -0.5 and 0.5 standard deviations. The value of the HSS value starts off at 10 at lag +1 week, and slowly drops from 6 to 3.5 from lag +2 to +6 weeks. Compared to the results

DJF HSS over KOR/JPN

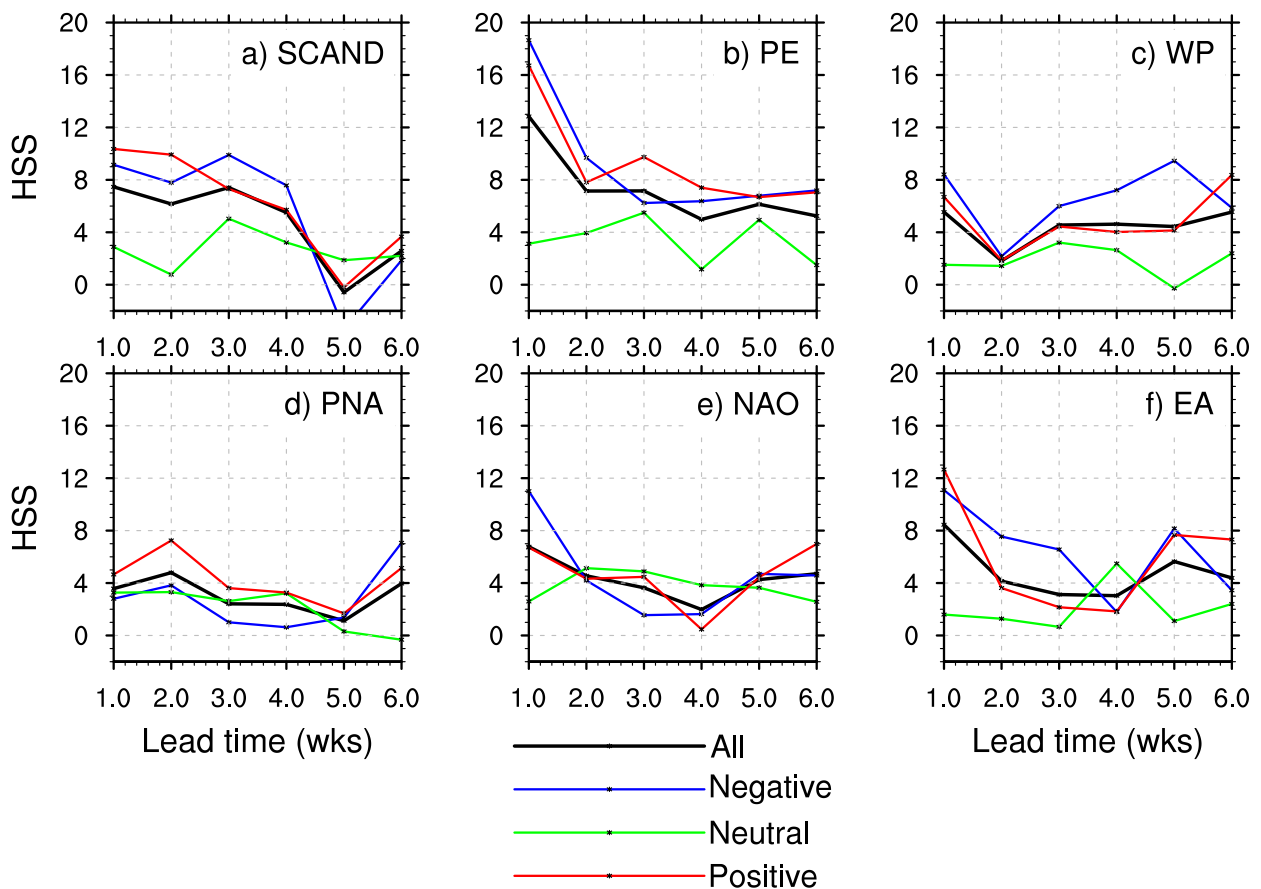


FIG. 9. As in Fig. 8, but for the domain of the Korean Peninsula and Japan.

from GloSea5 (Fig. 4a), these values may not seem very impressive, especially for the earlier lags. However, it is important to note that the EAWR model outperforms GloSea5 at lag +5 weeks, although very marginally, and provides skill beyond that of a random forecast out to lag +6 weeks. Over the Korean Peninsula and Japan, the benefits of using the EAWR are better. The domain-averaged HSS values begin with 14 at lag +1 week and decline to 7–4 at later lags (black line in Fig. 6b). For Mongolia, the values begin at 14 at lag +1 week and come down to 1 at lag +6 weeks (black line in Fig. 6d). The results for China are not too different from those of East Asia because in terms of area, China occupies most of East Asia (black line in Fig. 6c).

When the EAWR is active (i.e., excluding the neutral phase), the EAWR becomes an even better source of prediction skill. In Fig. 6, it can be seen that the positive (red lines) and negative (blue lines) phases generally result in clear elevations in skill compared to the neutral EAWR phase (green lines). For example, over East Asia, when the initial state of the EAWR is negative, the

values of HSS start from 16 at lag +1 week and decline to 5–6 (blue line in Fig. 6a). Such a boost in the HSS value is also apparent for the Korean Peninsula and Japan for both the positive and negative phases (red and blue lines in Fig. 6b, respectively). In all domains, the neutral EAWR almost always provides little skill (green lines in Fig. 6).

Next, we move on to the results for the other six teleconnection patterns, the SCAND, PE, WP, PNA, NAO, and EA patterns. Figure 7 shows the HSS values of each pattern for lag +1 weeks only. Consistent with the results of the EAWR, elevation in skill is apparent over the regions where the teleconnection patterns are most influential. For example, high HSS values for SCAND, EA, and NAO are clearly marked over the North Atlantic (Figs. 7a,b,e), while the increase of skill for the WP and PNA takes place over the North Pacific (Figs. 7c,d). The results for the other lags are also consistent with the results of the EAWR, which illustrate a persistent yet stable decline of the initial HSS values with lag (not shown).

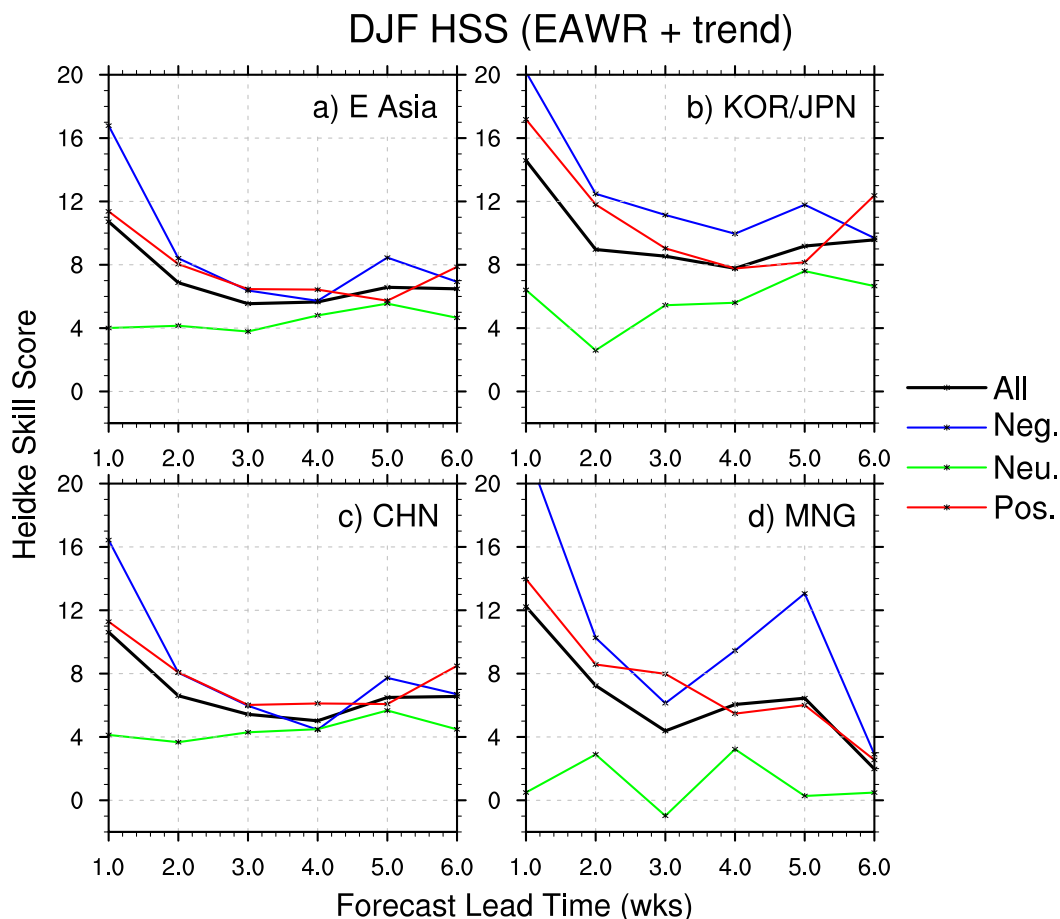


FIG. 10. As in Fig. 6, but with the long-term linear SAT trend included as an additional predictor.

For East Asia, a comparison between the six modes reveals that two teleconnection patterns, SCAND and PE, contribute the most to the prediction skill at lag +1 week (Figs. 7a–c). However, at lag +6 weeks, all the six teleconnection patterns provide skill with HSS values generally ranging from 1–4. When domain averaged for East Asia, the SCAND and PE predictors yield high HSS values at lag +1 week of about 8 and 6, respectively, and at lag +6 weeks values of 3–4 (black lines in Figs. 8a,b). The other four modes (i.e., WP, PNA, NAO, and PE) contribute slightly less to the skill (Figs. 7c–f). Over East Asia, the HSS values of WP, PNA, NAO, and PE range from approximately 1 to 4 throughout the lead times (Figs. 8c–f).

Over the Korean Peninsula and Japan, SCAND and PE (Figs. 9a,b) provide better values than the other four teleconnection patterns (Figs. 9c–f). For SCAND, the HSS values begin with 8 at lag +1 week and drops to 1–3 at lag +5 and +6 weeks. The HSS values obtained from PE starts with 6 at lag +1 weeks and decrease to 4 at lag +6 weeks. For the other four modes, the skill

averaged over the Korea Peninsula and Japan lies between 1 and 4 throughout the lead times.

The better performance of EAWR, SCAND, and PE over East Asia can be expected, given that the three modes are located upstream of the East Asia domain (Figs. 1a–c). In contrast, the PNA, for instance, is located too far downstream. Interestingly, the results demonstrate that the leading-order atmospheric teleconnection patterns in terms of explained variance may not be the most important factor to consider for predictability of local climate. In other words, modes such as the NAO and PNA are known to explain the largest portion of the hemispheric geopotential height variance and thus may be expected to provide the most skill when averaged over the hemisphere. Instead, we find here that adjacency and upstream location may be key considerations for potential predictors.

When averaged over the positive and negative phases at lag +1 week, the HSS fields of the active SCAND phase (i.e., the positive and the negative phases) show maximum values above 10 for East Asia (by averaging

DJF HSS over East Asia

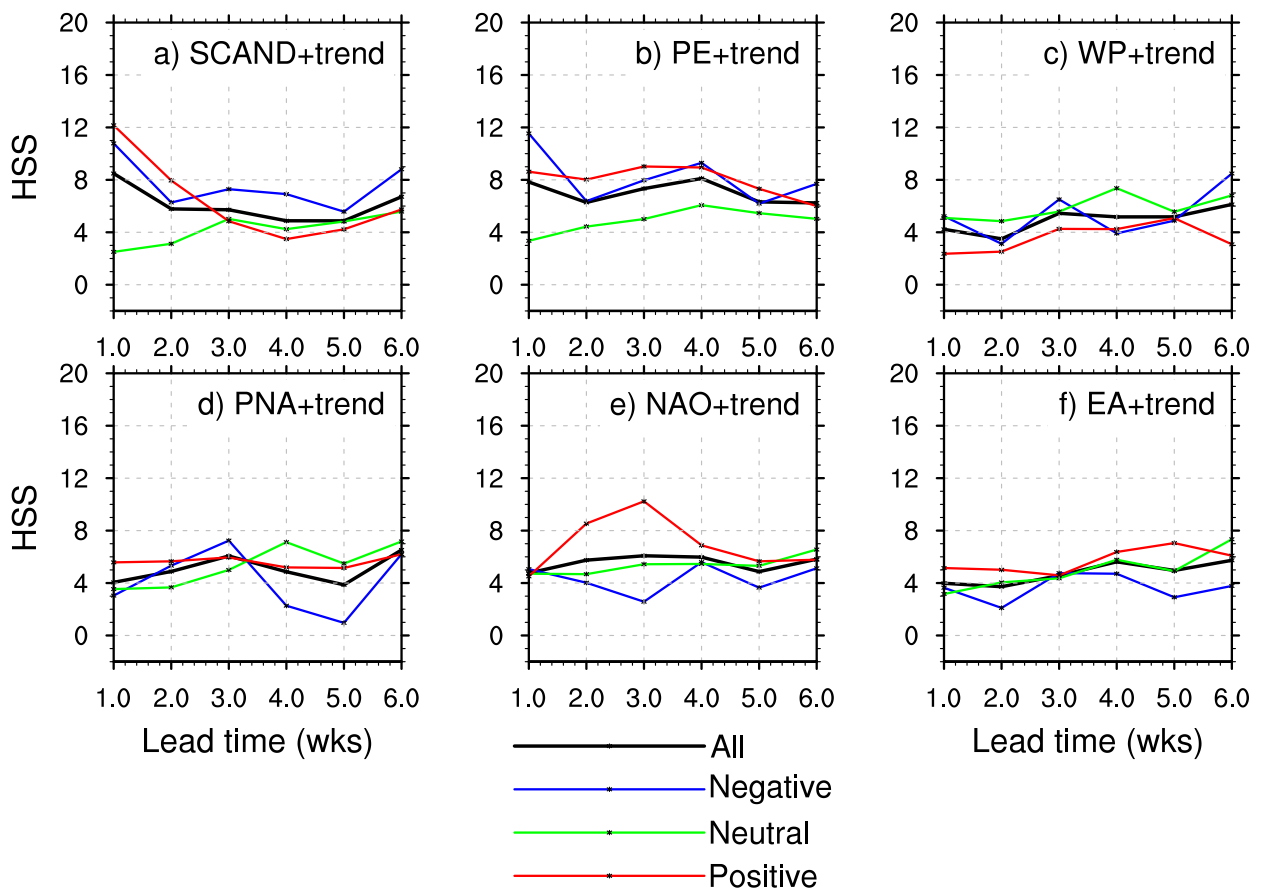


FIG. 11. As in Fig. 8, but with the long-term linear SAT trend included as an additional predictor.

the red and blue lines Fig. 8a) and near 10 for the Korean Peninsula and Japan (Fig. 9a). The PE also show elevated maximum HSS values near 8 for East Asia (red and blue lines in Fig. 8b) and 18 for the Korean Peninsula and Japan (Fig. 9b) during the active phases. A value near 10 or above from a single teleconnection pattern is an impressive result, considering for North America that a HSS value of about 7 was achieved only when both the ENSO and MJO were collectively active (Fig. 3b in Johnson et al. 2014).

Having demonstrated that skillful predictions for East Asia beyond 4–5-week lead times can be achieved by utilizing the Northern Hemisphere teleconnection patterns, we explore whether the long-term linear SAT trend further enhances the skill for the region. In Johnson et al. (2014), it was shown that the long-term trend, when combined with information of the ENSO and MJO phases, adds an HSS value of approximately 2 over North America (see their Fig. 3a), and hence the trend can be an important source of skill. Importantly, the increase of skill by the long-term trend takes place at all

lead times, which therefore helps the statistical model to surpass the performance of the dynamical model at earlier lead times.

Consistent with Johnson et al. (2014), we find a boost in HSS values over East Asia when the long-term trend is included as an additional predictor. The forecasts based on both the EAWR and trend show HSS values near 6–7 at lag +5 and +6 weeks (black line in Fig. 10a). This corresponds to an increase of HSS values by 2–3, compared to the forecasts solely based on the EAWR (black line in Fig. 6a). The increase can be similarly identified in all the other subdomains (Figs. 6b–d and 10b–d). Also, the boost in HSS values can generally be seen for all lead times in all subdomains when the skill is compared by each phase (colored lines in Figs. 6 and 10).

We also combine the long-term trend with the other teleconnection patterns. As for the result with the EAWR, as in Johnson et al. (2014) the enhancement of skill is evident for all the teleconnection patterns over East Asia (Fig. 11) and its subdomains (not shown), although for some teleconnection patterns, such as the PE and

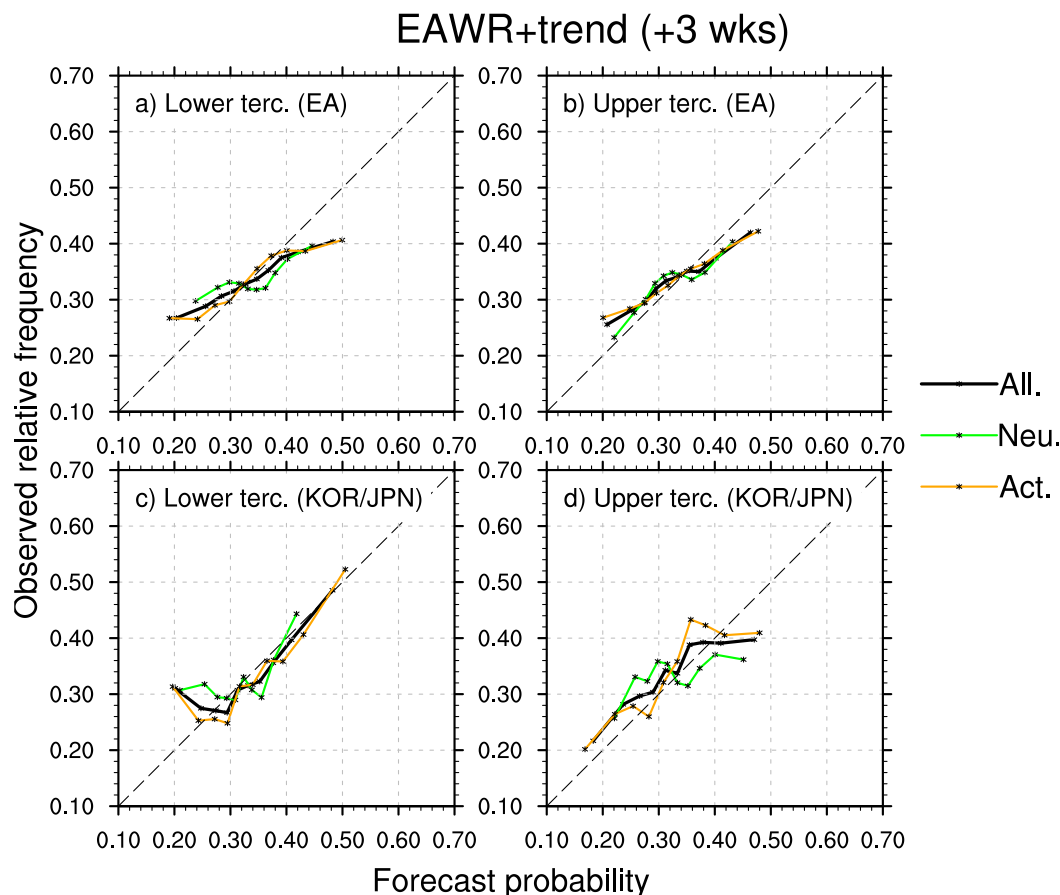


FIG. 12. Calibration functions for (a),(c) lower and (b),(d) upper tercile forecasts for a 3-week lead time using the initial state of the EAWR and the long-term trend, for (top) East Asia and (bottom) the Korean Peninsula and Japan. Each point contains 10% of the forecasts for the corresponding category. The black lines show the results for all phases, while the orange and green lines are when all the four teleconnection patterns are in active and neutral phases, respectively.

NAO, the HSS values at shorter leads are compromised. For instance, based on the PE and long-term trend, the HSS value starts at 7 (Fig. 11b) instead of 13 as for the case solely based on the PE (Fig. 9b). However, the increase of HSS values at longer leads is an important advantage for surpassing the skill of dynamical models.

Last, we examine the reliability of the forecasts for upper and lower terciles. Figure 12 shows the calibration functions of the model using the EAWR and the trend for lag +3 weeks as a representative result since the statistical model and GloSea5 show similar performance at that lag. Here, as in Johnson et al. (2014), information on forecast probability frequency is incorporated into the calibration function. That is, we have binned the forecast probability frequency so that each bin contains 10% of the forecasts. The relative frequency of occurrence of the verified category is therefore computed for each forecast probability bin. Overall, calibration functions are well centered over the diagonal, indicating that

the forecasts are not likely to be biased (black lines). Also, the calibration functions generally show slopes slightly less than 1, representing overconfidence of the forecasts. Over the Korean Peninsula and Japan, the confidence of the forecasts (Figs. 12c,d) appears worse than that for East Asia. When sampled for active phases, the forecasts are better calibrated (orange lines) than all phases (black lines). The neutral phase forecasts tend to be poorly calibrated (green lines), which is not surprising, given that the signal is weak.

4. Summary and discussion

We have constructed a composite-based statistical model to predict the East Asian wintertime SAT for lead times out to 6 weeks. The modeling technique was originally reported by Johnson et al. (2014) using the ENSO, MJO, and linear trend for North America. The present study attempted to utilize low-frequency

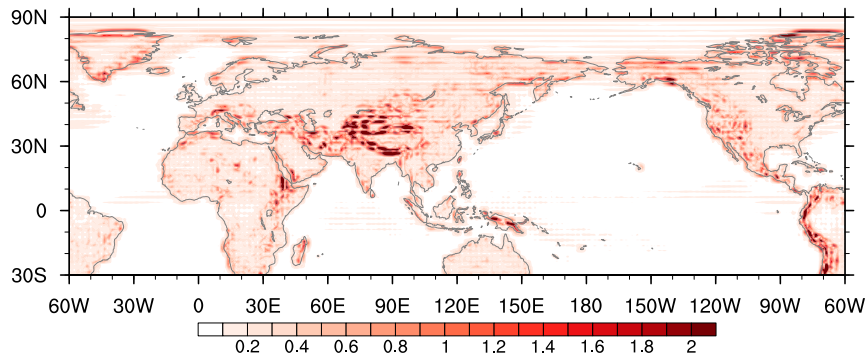


FIG. A1. Time mean for the year 2000 of half of the absolute difference between the original ERA-Interim data and the reconstructed data through bilinear interpolation at the reduced 512×256 Gaussian grid.

atmospheric circulation anomalies represented by seven Northern Hemisphere atmospheric teleconnection patterns (i.e., the EAWR, SCAND, EA, WP, NAO, PNA, and PE patterns) as a potential source of prediction skill. The effect of the linear trend is also tested in combination with the teleconnections.

The study demonstrates that the following three teleconnections make a substantial contribution to East Asian subseasonal prediction of SAT. These are the EAWR, SCAND, and PE patterns, which are positioned upstream of the region. When such modes are employed individually, the statistical model generates forecasts that show comparable skill with that of dynamical model at a lead time of 4 weeks and longer. This is especially true when the amplitude of the teleconnection pattern is large. More interestingly, our results show that when the initial information of the teleconnection patterns is combined with the long-term trend, the skill is substantially enhanced for lead times beyond 3 weeks. This is a great advantage because it is at these longer lead times that the dynamical model suffers from a rapid decline of prediction skill.

It is not clear at this point exactly how the teleconnections provide skill beyond their decorrelation time scales of ~ 2 weeks. One may speculate that the skill may be largely due to a small number of much longer persistent events. It is also possible that the SAT anomalies associated with the teleconnections can remain for a longer period of time through feedback processes with the land surface. In addition, combinations of teleconnections may be utilized to boost prediction skill for lead times beyond 3 weeks. We plan to address these questions in our future study.

Our approach using atmospheric teleconnections has implications for other regions of the globe. Over Europe, the Middle East, and northern Africa, forecasts based on the initial state of EA, SCAND, EAWR, and

the NAO, along with the long-term trend, yield HSSs exceeding 10 at lead times of 3–4 weeks (not shown). The Gaussian assumption may need to be relaxed for some of the regions due to less normality in the temperature distributions.

Acknowledgments. The authors thank Dr. Byung-Kwon Park for making the GloSea5 hindcast experiments available for the study and two anonymous reviewers for their constructive comments. The authors also acknowledge the European Center for Medium-Range Weather Forecasts for providing the ERA-Interim reanalysis products via their data server. CY acknowledges the support of the Korea Polar Research Institute Grant PN17081. NCJ acknowledges the support from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce (Award NA14OAR4320106). C-H C is supported by the Ministry of Science and Technology Grant MOST 106-2111-M-002-003-MY2. SBF acknowledges the support of the National Science Foundation Grants AGS-1401220 and AGS-1822015.

APPENDIX A

Uncertainty of the Spatial Interpolation

The ERA-interim data used in the study are originally on a reduced 512×256 longitude/latitude Gaussian grid. We use bilinear interpolation to regrid the data to the 2.5° horizontal resolution. The choice of bilinear interpolation is made because the 2-m temperature is highly spatially correlated and hence is smoothly varying in space, and the purpose of the regridding is to remove small-scale features, as our focus is to examine SAT forecasts at spatial scales that are much greater than the 2.5° horizontal resolution. Furthermore, because we use the smoothed data as our ground truth, we believe that

TABLE B1. The five-number summary of the difference and the linear correlation between the temporally downscaled index minus the CPC daily index for the DJF PNA and NAO.

	PNA	NAO
Minimum	−2.38	−2.34
1st quartile	−0.6	−0.61
Median	−0.22	−0.20
3rd quartile	0.17	0.20
Maximum	1.91	1.83
R with CPC daily index	0.82	0.79

the errors arising from the interpolation are not necessarily related to errors in our forecasts. That being said, we quantify the possible error of the spatial interpolation scheme by computing the half of the absolute difference between the original data and the reconstructed data at the original 512×256 Gaussian grid. Here the reconstructed data have gone through the interpolation twice, from the original to coarse and then back to the original grid. Figure A1 shows the time-averaged values of the difference for the year 2000. The data loss is most clear over mountainous regions. Also, over some non-mountainous regions, such as Siberia, Iceland, the Sahara, and the Sahel, the data loss can be identified.

APPENDIX B

Uncertainty of the Temporal Downscaling

The uncertainty of the temporal downscaling technique on the monthly CPC teleconnection indices is documented here. The daily PNA and NAO indices are constructed using this technique and are then compared with the daily indices obtained from the CPC. The five-number summary of the difference between the temporally downscaled index and the CPC daily index for DJF is computed (Table B1). The five-number summary of the difference suggests that the CPC daily index tends to be greater than the value of the downscaled one. Linear correlations between the indices are 0.82 and 0.79 for the PNA and NAO, respectively. Based on the correlation values, we estimate that the amounts of variance overlap between the two indices are 67.24% and 62.41% for the PNA and NAO, respectively.

REFERENCES

- Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115**, 1083–1126, [https://doi.org/10.1175/1520-0493\(1987\)115<1083:CSAPOL>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2).
- Branstator, G., 2002: Circumglobal teleconnections, the jet stream waveguide, and the North Atlantic Oscillation. *J. Climate*, **15**, 1893–1910, [https://doi.org/10.1175/1520-0442\(2002\)015<1893:CTTJSW>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1893:CTTJSW>2.0.CO;2).
- , 2014: Long-lived response of the midlatitude circulation and storm tracks to pulses of tropical heating. *J. Climate*, **27**, 8809–8826, <https://doi.org/10.1175/JCLI-D-14-00312.1>.
- Cassou, C., 2008: Intraseasonal interaction between the Madden–Julian oscillation and the North Atlantic Oscillation. *Nature*, **455**, 523–527, <https://doi.org/10.1038/nature07286>.
- Dai, Y., S. B. Feldstein, B. Tan, and S. Lee, 2017: Formation mechanisms of the Pacific–North American teleconnection with and without its canonical tropical convection pattern. *J. Climate*, **30**, 3139–3155, <https://doi.org/10.1175/JCLI-D-16-0411.1>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- DeSole, T., L. Trenary, M. K. Tippet, and K. Pegion, 2017: Predictability of week–3–4 average temperature and precipitation over the contiguous United States. *J. Climate*, **30**, 3499–3512, <https://doi.org/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, [https://doi.org/10.1175/1520-0442\(2000\)013<4430:TTPSAC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<4430:TTPSAC>2.0.CO;2).
- Halpert, M. S., and C. F. Ropelewski, 1992: Surface temperature patterns associated with the Southern Oscillation. *J. Climate*, **5**, 577–593, [https://doi.org/10.1175/1520-0442\(1992\)005<0577:STPAWT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<0577:STPAWT>2.0.CO;2).
- Higgins, R. W., and K. C. Mo, 1997: Persistent North Pacific circulation anomalies and the tropical intraseasonal oscillation. *J. Climate*, **10**, 223–244, [https://doi.org/10.1175/1520-0442\(1997\)010<0223:PNPCAA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<0223:PNPCAA>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, [https://doi.org/10.1175/1520-0493\(1981\)109<0813:PSAPAW>2.0.CO;2](https://doi.org/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 orographic forcing. *J. Atmos. Sci.*, **38**, 1179–1196, [https://doi.org/10.1175/1520-0469\(1981\)038<1179:TSLROA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1981)038<1179:TSLROA>2.0.CO;2).
- 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, <https://doi.org/10.1175/2009JCLI3099.1>.
- , 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, <https://doi.org/10.1175/WAF-D-13-00102.1>.
- Lin, H., 2015: Subseasonal variability of North American wintertime surface air temperature. *Climate Dyn.*, **45**, 1137–1155, <https://doi.org/10.1007/s00382-014-2363-6>.
- , G. Brunet, and J. Derome, 2009: An observed connection between the North Atlantic Oscillation and the Madden–Julian oscillation. *J. Climate*, **22**, 364–380, <https://doi.org/10.1175/2008JCLI2515.1>.
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, [https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2).
- Lukens, K. E., S. B. Feldstein, C. Yoo, and S. Lee, 2017: The dynamics of the extratropical response to Madden–Julian Oscillation convection. *Quart. J. Roy. Meteor. Soc.*, **143**, 1095–1106, <https://doi.org/10.1002/qj.2993>.

- MacLachlan, C., and Coauthors, 2015: Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 1072–1084, <https://doi.org/10.1002/qj.2396>.
- 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, [https://doi.org/10.1175/1520-0469\(1971\)028<0702:DOADOI>2.0.CO;2](https://doi.org/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, [https://doi.org/10.1175/1520-0469\(1972\)029<1109:DOGSCC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2).
- Nakamura, H., and J. M. Wallace, 1990: Observed changes in baroclinic wave activity during the life cycles of low-frequency circulation anomalies. *J. Atmos. Sci.*, **47**, 1100–1116, [https://doi.org/10.1175/1520-0469\(1990\)047<1100:OCIBWA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1990)047<1100:OCIBWA>2.0.CO;2).
- 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, <https://doi.org/10.1007/s00382-012-1493-y>.
- Sardeshmukh, P. D., and B. J. Hoskins, 1988: The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.*, **45**, 1228–1251, [https://doi.org/10.1175/1520-0469\(1988\)045<1228:TGOGRF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<1228:TGOGRF>2.0.CO;2).
- Seo, K.-H., H.-J. Lee, and D. M. W. Frierson, 2016: Unraveling the teleconnection mechanisms that induce wintertime temperature anomalies over the Northern Hemisphere continents in response to the MJO. *J. Atmos. Sci.*, **73**, 3557–3571, <https://doi.org/10.1175/JAS-D-16-0036.1>.
- Son, S. W., Y. Lim, C. Yoo, H. Hendon, and J. Kim, 2017: Stratospheric control of the Madden–Julian oscillation. *J. Climate*, **30**, 1909–1922, <https://doi.org/10.1175/JCLI-D-16-0620.1>.
- Trenberth, K. E., and J. M. Caron, 2000: The Southern Oscillation revisited: Sea level pressures, surface temperatures, and precipitation. *J. Climate*, **13**, 4358–4365, [https://doi.org/10.1175/1520-0442\(2000\)013<4358:TSORSL>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<4358:TSORSL>2.0.CO;2).
- 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, <https://doi.org/10.1029/2003GL018645>.
- 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, [https://doi.org/10.1175/1520-0493\(1981\)109<0784:TITGHF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1981)109<0784:TITGHF>2.0.CO;2).
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. Academic Press, 676 pp.
- Yoo, C., S. Lee, and S. B. Feldstein, 2012: Mechanisms of Arctic surface air temperature change in response to the Madden–Julian oscillation. *J. Climate*, **25**, 5777–5790, <https://doi.org/10.1175/JCLI-D-11-00566.1>.
- , S. B. Feldstein, and S. Lee, 2014: The prominence of a tropical convective signal in the wintertime Arctic temperature. *Atmos. Sci. Lett.*, **15**, 7–12, <https://doi.org/10.1002/asl2.455>.
- , S. Park, D. Kim, J.-H. Yoon, and H.-M. Kim, 2015: Boreal winter MJO teleconnection in the Community Atmosphere Model version 5 with the Unified Convection parameterization. *J. Climate*, **28**, 8135–8150, <https://doi.org/10.1175/JCLI-D-15-0022.1>.