Mapping Large-Scale Climate Variability to Hydrological Extremes: An Application of the Linear Inverse Model to Subseasonal Prediction

KAI-CHIH TSENG,[a,b] NATHANIEL C. JOHNSON,[b] ERIC D. MALONEY,[c] ELIZABETH A. BARNES,[c] AND SARAH B. KAPNICK[b]

[a] *Program in Atmospheric and Oceanic Science, Princeton University, Princeton, New Jersey*
[b] *NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey*
[c] *Colorado State University, Fort Collins, Colorado*

(Manuscript received 7 July 2020, in final form 11 February 2021)

ABSTRACT: The excitation of the Pacific–North American (PNA) teleconnection pattern by the Madden–Julian oscillation (MJO) has been considered one of the most important predictability sources on subseasonal time scales over the extratropical Pacific and North America. However, until recently, the interactions between tropical heating and other extratropical modes and their relationships to subseasonal prediction have received comparatively little attention. In this study, a linear inverse model (LIM) is applied to examine the tropical–extratropical interactions. The LIM provides a means of calculating the response of a dynamical system to a small forcing by constructing a linear operator from the observed covariability statistics of the system. Given the linear assumptions, it is shown that the PNA is one of a few leading modes over the extratropical Pacific that can be strongly driven by tropical convection while other extratropical modes present at most a weak interaction with tropical convection. In the second part of this study, a two-step linear regression is introduced that leverages a LIM and large-scale climate variability to the prediction of hydrological extremes (e.g., atmospheric rivers) on subseasonal time scales. Consistent with the findings of the first part, most of the predictable signals on subseasonal time scales are determined by the dynamics of the MJO–PNA teleconnection while other extratropical modes are important only at the shortest forecast leads.

KEYWORDS: Madden-Julian oscillation; Teleconnections; Extreme events; Climate prediction

1. Introduction and motivation

Since the early 1980s, the extratropical teleconnected response induced by the large-scale tropical heating [e.g., El Niño–Southern Oscillation (ENSO) or Madden–Julian oscillation (MJO)] has been well studied through the framework of linear Rossby wave theory (Hoskins and Karoly 1981). In the tropics, due to the weak Coriolis force and horizontal temperature gradient, the convective heating is balanced by the adiabatic cooling of rising motion (Charney 1963). The divergent outflow in the upper troposphere interacts with the potential vorticity gradient around the subtropical jet and generates a Rossby wave source (Sardeshmukh and Hoskins 1988; Hoskins and Ambrizzi 1993), which excites Rossby waves that propagate to the downstream regions. This process allows for the influence of tropical heating extending beyond the barrier of tropical easterlies.

One canonical example of tropical–extratropical teleconnection is the excitation of the Pacific–North American pattern (PNA) by tropical convective heating. The PNA encompasses four high or low pressure centers that follow a wave train that initiates in the subtropical Pacific, strengthens around the Gulf of Alaska, and then propagates to the East Coast of the United States. Studies have found that many PNA events can be excited by a tropical heating pattern associated with the MJO and ENSO (Horel and Wallace 1981; Hsu 1996). By conducting a detailed vorticity budget analysis and idealized simulations in a dry general circulation model (GCM), Mori and Watanabe (2008) and Seo and Lee (2017) demonstrated that the initiation of a PNA can be attributed to the barotropic conversion induced by the interaction between MJO divergent outflow and the subtropical jet. This is evident especially for specific MJO phases (Tseng et al. 2019). On interannual time scales, ENSO is crucial for modulating extratropical circulations. The ENSO teleconnection was first documented by Bjerknes (1969). With a one-point regression map, Horel and Wallace (1981) unearthed the covariability between the PNA pattern and the tropical SST on interannual time scales. Similar teleconnection patterns were identified in the following observational and modeling studies with various statistical approaches, which demonstrate the robustness of the ENSO–PNA teleconnection (Halpert and Ropelewski 1992; Trenberth et al. 1998).

A particular motivation of studying tropical–extratropical teleconnections is that some persistent extratropical extremes, such as anticyclonic blocking (Henderson et al. 2016), atmospheric rivers (Mundhenk et al. 2016), extreme cold-air outbreaks (Lin 2018), and heat waves (Lee and Grotjahn 2019), are associated with the slow variation of the extratropical circulation that can be driven by the interaction between tropical convection and the extratropical atmospheric circulation. Sardeshmukh et al. (2000) found a small shift of the ENSO states can lead to a dramatic increase of extreme events over the extratropics. In addition, the information from the tropics takes time to develop over the extratropical regions. A delayed signal between forcing and response indicates the potential to leverage the tropical information for subseasonal forecasts of extratropical weather. By analyzing the ensemble hindcasts of an operational forecast model, Tseng et al. (2018) found the improved forecast skills of extratropical geopotential

Corresponding author: Kai-Chih Tseng, kaichiht@princeton.edu

DOI: 10.1175/JCLI-D-20-0502.1

height over specific MJO phases on subseasonal time scales (2–5 weeks).

Although the excitation of the PNA by tropical convection has been well studied, the interaction between the tropics and other Pacific–North America region teleconnection patterns (i.e., other leading modes of extratropical variability separate from the PNA) has received comparatively little attention. There are two possible reasons. First, the power spectrum of extratropical circulation is nearly red, and so the slowest-varying modes such as the PNA may dominate any analysis that emphasizes explained variance. Thus, when analyzing the raw data, PNA signal stands out from other extratropical modes. The second possibility is that the PNA is one of a few leading modes that can be strongly driven by tropical convection whereas other modes have little interaction with the tropics. In this study, we examine both possibilities with a linear inverse model (LIM). The LIM is a multivariate linear stochastically forced system, whose parameters are derived from the observed covariance of system components (Penland and Sardeshmukh 1995). Although the LIM is statistically stable, it allows for nonmodal growth on short time scales. In addition, forecasts derived from LIM can have forecast skills comparable to state-of-art climate models for seasonal ENSO prediction and subseasonal prediction (Newman and Sardeshmukh 2017; Albers and Newman 2019), which indicates its process-level relevance to the full-physics climate models. Henderson et al. (2020) further demonstrates that some fundamental dynamics of PNA initiation, such as the tropical–extratropical teleconnection and the internal extratropical dynamics can be captured by a LIM.

To investigate the interactions between tropical convection and other extratropical modes in addition to the PNA, this study is organized as follows. In section 2, detailed descriptions of the data and a LIM is provided. In section 3, we evaluate and confirm the suitability of the LIM for modeling PNA region teleconnection patterns. We then examine and quantify the optimal tropical forcing patterns that evolve into different extratropical modes such as the PNA (details are given in the following sections). In section 4, we propose a two-step linear regression that enables us to apply a LIM to the prediction of hydrological variables, including atmospheric river frequency, on subseasonal time scales without violating the assumptions inherent in the LIM. We next investigate the relationships between different extratropical modes and the prediction of hydrological extremes. Section 5 presents the conclusion and remarks.

2. Data, model, and method

a. Reanalysis and satellite data

We use the European Centre for Medium-Range Weather Forecasts (ECMWF) third generation reanalysis product (ERA-Interim; Dee et al. 2011) from 1979 to 2017. Since the extratropical teleconnection patterns have the strongest signal during boreal winter, we only use the data from November through March. To construct a LIM that allows for studying tropical–extratropical interaction, we include 500-hPa geopotential height (Z500) from ERA-Interim and outgoing longwave radiation data (OLR) obtained from the National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellites (Liebmann and Smith 1996) in the state vectors of the LIM. To examine the role of extratropical modes in modulating hydrological extremes, the precipitable water and the integrated water vapor transport (IVT) are used in this study. The IVT is defined as

$$\mathrm{IVT} = \frac{1}{g}\sqrt{\left(\int_{1000}^{300} qu\,dp\right)^2 + \left(\int_{1000}^{300} qv\,dp\right)^2}, \qquad (1)$$

where q is specific humidity; u and v are zonal and meridional wind, respectively; g is the gravitational acceleration; and the integral is evaluated for the layer between 1000 and 300 hPa.

To identify connections between PNA region teleconnection patterns and hydrological extremes, we apply an atmospheric river (AR) detection algorithm developed by Mundhenk et al. (2016) to identify AR activity. Specifically, the detection algorithm incorporates the anomalous IVT intensity and geometric information (e.g., aspect ratio and total area) to identify a plumelike feature with an appropriate spatial scale, where each time step is scrutinized independently. In this study, the 94th percentile of the anomalous IVT intensity over the Pacific basin is used as the minimum threshold of detecting ARs. This criterion has been used in previous AR studies (e.g., Mundhenk et al. 2016, 2018). An AR event is recorded when the criteria are satisfied or exceeded. The details can be found in the appendix of Mundhenk et al. (2016), and the detection algorithm is available online (https://mountainscholar.org/handle/10217/170619). All variables are converted into daily anomalies by removing the first three harmonics of the seasonal cycle and then regridding at 2.5° × 2.5° spatial resolution due to the highest available resolution of OLR data. The AR detection is originally conducted at a higher resolution (1° × 1°) and then interpolated to lower resolution to ensure the AR frequency is not underestimated in this study. The analysis domain for AR detection spans the North Pacific and North America (20°–80°N, 150°E–30°W).

b. MJO index

In addition to the OLR data, we also use the Real-Time Multivariate MJO (RMM) index as a measure of the MJO activity (Wheeler and Hendon 2004). The RMM index is defined as the principal components (PCs) of the first two leading empirical orthogonal functions (EOFs) of the combined 200-hPa zonal wind, 850-hPa zonal wind and OLR fields averaged over 15°S–15°N. (The daily RMM index is available at http://www.bom.gov.au/climate/mjo/graphics/rmm.74toRealtime.txt.) In this study, we also replace the OLR-based state vectors with the RMMs to examine the degree to which the RMM index captures the tropical forcing of PNA region teleconnection patterns.

c. LIM

Whitaker and Sardeshmukh (1998) showed that many features of extratropical synoptic eddy statistics can be captured by a linear stochastically forced system. Cash and Lee (2001)

and Henderson et al. (2020) further demonstrated the optimal growth of the PNA pattern and the associated tropical–extratropical interaction are reasonably simulated by a LIM. A LIM uses the statistics of coarse-grained variables $\mathbf{x}$ to approximate the dynamical evolution of the observed system:

$$\frac{d}{dt}\mathbf{x} = \mathbf{B}\mathbf{x} + \xi, \tag{2}$$

where $\mathbf{x}$ is a state vector, including daily Z500 and OLR anomalies, $\mathbf{B}$ is a stable linear operator, which represents the interaction among different system variables and ξ is random white noise forcing. It is worth mentioning that the linear approximation of nonlinear processes is also included in matrix $\mathbf{B}$. In this study, the coarse-grained state vector (i.e., $\mathbf{x}$) consists of principal components (PCs) of the leading Z500 and OLR EOFs. Details are provided in section 3. With (2), the homogeneous solution of $\mathbf{x}$ (i.e., the linear predictable components) can be written as

$$\mathbf{x}_\tau = \mathbf{G}_\tau \mathbf{x}_0 = e^{\mathbf{B}\tau}\mathbf{x}_0. \tag{3}$$

In (3), the subscripts 0 and τ indicate the deterministic evolution of $\mathbf{x}$ from lag = 0 to lag = τ and $\mathbf{G}_\tau$ is a propagator operator that provides the best estimation of $\mathbf{x}_\tau$ with input $\mathbf{x}_0$ in a least squares error sense. Mathematically, $\mathbf{G}_\tau$ is the so-called Jacobian matrix or Green's function, which can be estimated by using two covariance matrices $\mathbf{C}_\tau$ and $\mathbf{C}_0$:

$$\mathbf{C}_\tau \mathbf{C}_0^{-1} = \mathbf{G}_\tau, \tag{4}$$

where $\mathbf{C}_\tau$ and $\mathbf{C}_0$ are defined as $\mathbf{x}_\tau\mathbf{x}_0^{\mathrm{T}}$ and $\mathbf{x}_0\mathbf{x}_0^{\mathrm{T}}$, respectively (the superscript T is the transpose operator). In general, $\mathbf{B}$ in (3) is independent of τ due to the linear assumption of (2). According to this relationship, Penland and Sardeshmukh (1995) proposed a τ test to examine the reasonableness of the linearity assumption. Specifically, if a LIM is applicable to a given dataset, the diagonal element of $\mathbf{G}_\tau$ will decay exponentially with the increase of τ, which gives a stable estimation of $\mathbf{B}$; otherwise, the linear assumption in (1) breaks down. The τ test can be written as

$$\mathbf{B} = \ln(\mathbf{G}_\tau)/\tau. \tag{5}$$

Equation (5) also implies the system of interest can be approximated by a Markov process. Since we are interested in the growth of different extratropical modes, a generalized eigenvalue analysis is applied to $\mathbf{G}_\tau$ under different norms $\mathbf{N}$:

$$\mathbf{G}_\tau^{\mathrm{T}}\mathbf{N}\mathbf{G}_\tau\mathbf{p} - \mu(\tau)\mathbf{p} = 0. \tag{6}$$

In (6), $\mu(\tau)$ is the eigenvalue or the growth rate of the system under the norm $\mathbf{N}$ and $\mathbf{p}$ is the corresponding eigenvector or the optimal initial pattern. Henderson et al. (2020) specified a PNA norm (i.e., $\mathbf{N}_{\mathrm{PNA}}$) that ensures $\mathbf{x}_\tau$ will be a PNA pattern. In this study, we extend their results by specifying the $\mathbf{N}$ with the norms of different extratropical modes and explore the corresponding initial tropical heating patterns. The $\mathbf{N}$ matrix can be derived by calculating the regression coefficient among different variables in state vectors:

$$\mathbf{N}_n = \mathbf{r}_n\mathbf{r}_n^{\mathrm{T}} + \varepsilon\mathbf{I}. \tag{7}$$

In (7), $\mathbf{r}_n$ is a diagonal matrix where the diagonal elements are the regression coefficient of each system variable with the PC of n extratropical leading modes. For example, if $n = 1$, the first diagonal element of $\mathbf{r}_1$ is the regression coefficient between PC1 and PC1 and the second diagonal element is the regression coefficient between PC1 and PC2. $\mathbf{N}_n$ is simply the squared amplitude of the projection of different modes onto the final state $\mathbf{x}_\tau$. To ensure the numerical stability when solving the eigenvalue problems of (6), we include an identity matrix (i.e., $\mathbf{I}$) multiplied by a small constant $\varepsilon = 10^{-9}$ in the original $\mathbf{N}$ matrix. Similar approaches have been used in Henderson et al. (2020), Vimont et al. (2014), and Tziperman et al. (2008).

3. Tropical–extratropical interaction versus internal extratropical dynamics

a. τ test

In this study, we set up two different LIMs (8). In the first, the state vectors consist of the PCs from the first 10 leading EOFs of Z500 and two RMMs [$\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{10}$ and RMMs in (8a)]. As discussed below, the 10 leading Z500 EOFs represent the extratropical Pacific atmospheric teleconnection patterns, including the PNA. In the second LIM, we replace the RMMs with PCs of the first 10 leading EOFs of tropical OLR in (8b). Since we are interested in tropical–extratropical interactions, the Z500 EOFs are calculated over the extratropical Pacific domain (20°–80°N, 120°E–90°W) and OLR EOFs are calculated over a tropical band (15°S–15°N). The first 10 EOFs of Z500 and OLR explain 80% and 53% of total variances, respectively, and the results remain qualitatively unchanged when additional modes are included (not shown). While the first LIM focuses on the MJO-related teleconnection, the second LIM more generally addresses which tropical convection patterns can lead to the growth of different extratropical modes.

Figure 1 shows the first 10 EOFs (regression maps) of extratropical geopotential height variability used in this study. To derive the regression maps, we first apply EOF analysis to the North Pacific–western North America domain extending from 20° to 80°N and from 120°E to 90°W and acquire the corresponding PCs. The daily Z500 anomalies are then regressed upon the PCs to produce the regression maps shown in Fig. 1. We will show their relationships with hydrological variables in later sections. Since EOFs maximize explained variance (in both space and time), it is not surprising that the first two leading EOFs, which explain about 35% of Z500 variability in the domain, are of planetary scale (wavenumbers 1–4). EOF1 is characterized by a dipole structure over the extratropical Pacific and the North America. The maximum amplitude of EOF1 is spatially collocated with the regions where the Z500 has the greatest variance on synoptic time scales. This pattern is similar to the North Pacific Oscillation–west Pacific pattern (NPO–WP) identified by Walker and Bliss (1932). EOF2 shows a canonical PNA pattern (Wallace and Gutzler 1981) with three high and low pressure centers extending from the subtropical Pacific to the southeast of North

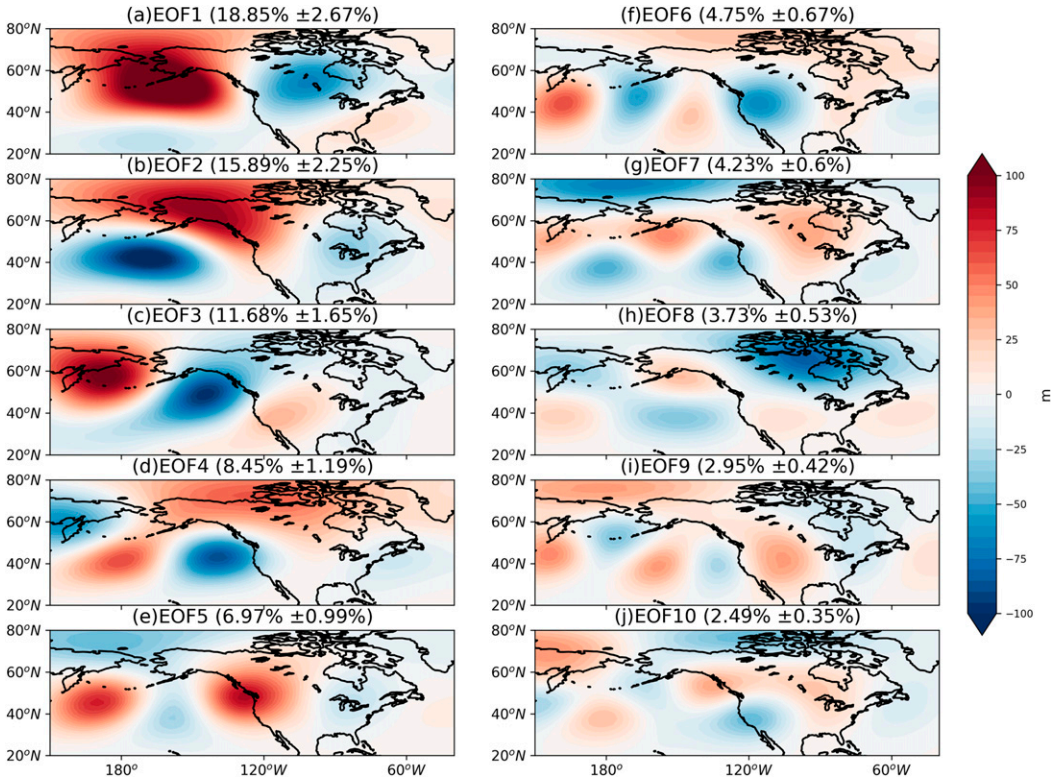

FIG. 1. The regression pattern of Z500 with first 10 leading PCs over the domain 20°–80°N, 120°E–90°W (unit: m). The number in the parentheses is variance explained by each leading mode and the corresponding 95% confidence bounds according to North et al. (1982).

America. The other EOF patterns generally show more zonally oriented features of smaller scales in both the zonal and meridional directions, which likely correspond to synoptic wave propagation:

$$\mathbf{x} = \begin{bmatrix} z_1 \\ \vdots \\ z_{10} \\ \mathrm{RMM1} \\ \mathrm{RMM2} \end{bmatrix} \tag{8a}$$

and

$$\mathbf{x} = \begin{bmatrix} z_1 \\ \vdots \\ z_{10} \\ \mathrm{OLR}_1 \\ \vdots \\ \mathrm{OLR}_{10} \end{bmatrix}. \tag{8b}$$

To quantify the decaying time scales of each extratropical mode, a τ test is applied to the first LIM (with RMMs supplying the tropical state information). Figure 2 illustrates the results of this test by showing the diagonal elements of $\mathbf{G}_\tau$ as a function of τ, which represents the autocorrelation functions of each variable extending to a lag of 14 days. The blue line in each figure is the corresponding exponential fit, and the legend shows the exponential fitting equation (real part only). The rate at which the autocorrelation decreases with tau specifies the intrinsic time scale of each mode, with an exponential decay indicating a red noise spectrum. The intrinsic time scale of each mode then can be calculated as the e-folding time of the exponentially decaying curves. As the forecast lead time approaches infinity, the linear predictable signals will approach zero, indicating we are predicting a climatological value (i.e., 0), which is the red noise assumption of LIM. From Fig. 2, one can find the exponential curves are closely collocated with the autocorrelation functions estimated by the LIM, which supports a red spectrum for each mode. Moreover, the stability of the decorrelation time scales with respect to choice of lag validates the use of LIM for predicting daily Z500. This feature also implies the Markov process can describe the evolution of the leading extratropical modes. Another apparent feature in Fig. 2 is that the decorrelation time scales are much longer for the first two EOFs (5.4 and 7.2 days, respectively) than the others (~3 days). This indicates that lower-frequency extratropical Pacific Z500 variability with time scales longer than a week is dominated by the first two EOFs. We also examine the decorrelation time scales of each extratropical mode in the second LIM, which replaces the RMMs with the 10 leading EOFs of tropical OLR. The result is nearly identical to the first

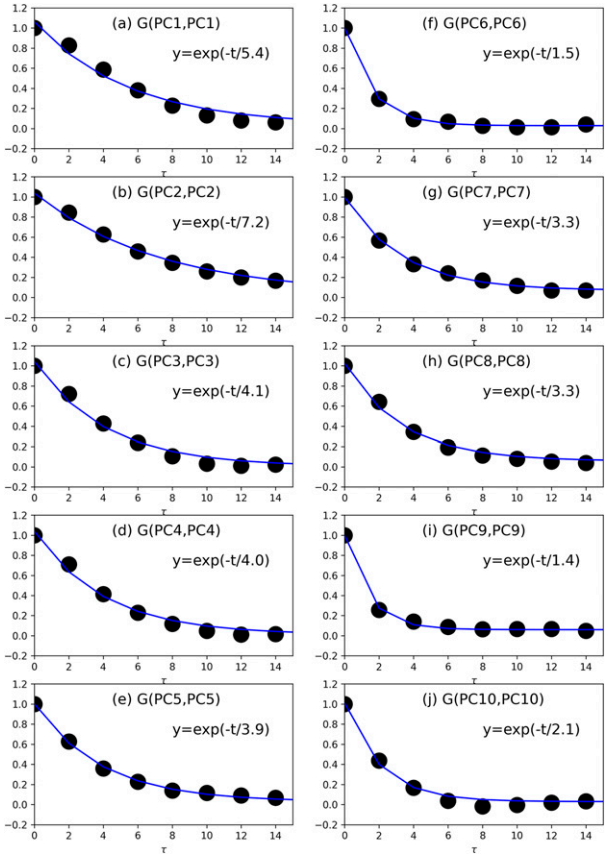

FIG. 2. The first 10 diagonal elements of $\mathbf{G}_\tau$ as a function of τ (days). The legend and the blue curves show the exponential fitting function (i.e., the real part) of each extratropical leading mode.

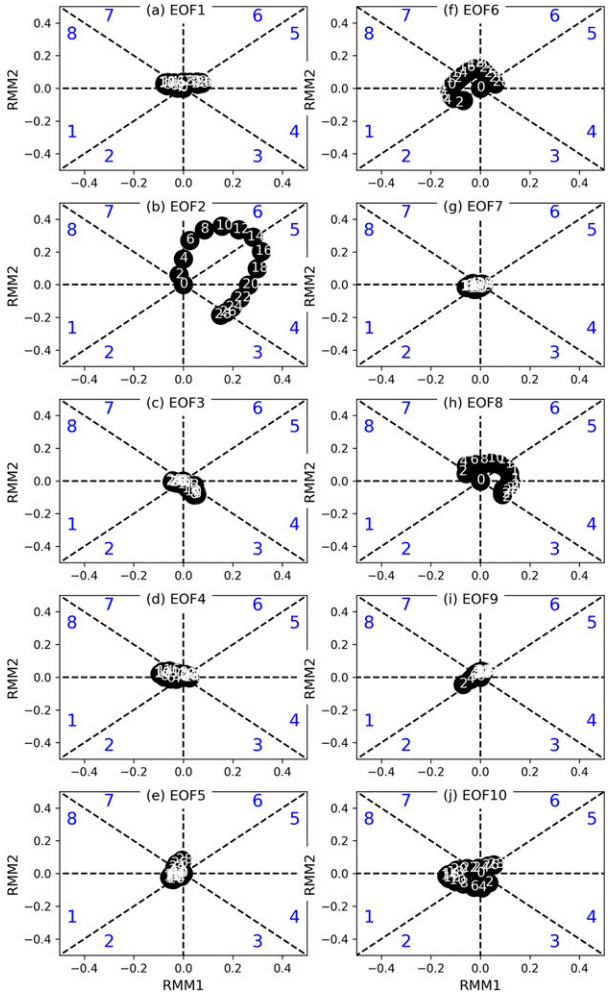

FIG. 3. The optimal MJO forcing for the development of different extratropical modes shown in the Wheeler and Hendon (2004) RMM phase space. The number in each dot indicates the time scale in days for initial optimal forcing (i.e., $\mathbf{x}_0$) growing into the ultimate response pattern (i.e., $\mathbf{x}_\tau$). Results are from the RMM-based LIM.

LIM (not shown), which reinforces the robustness of the analysis.

The τ test, in general, provides the decaying time scales of different forced responses, that is, decaying from the maximum amplitude to a factor of e^{-1}. However, all time scales shown in Fig. 2 are either shorter than or merely extending to the lower end of subseasonal time scales (2–5 weeks). This indicates that the hope for subseasonal predictions within the PNA regions cannot rely on the persistence of extratropical modes and must instead involve other processes. Tropical convective forcing represents one of these most well-known processes (Vitart 2017), as the signal from the tropics takes time to develop over the extratropical regions. It is therefore important to take the delayed time scales between tropical forcing and extratropical responses into account. Thus, in the following section, we will examine the initial optimal forcing pattern in the tropics with varying τ to explore and quantify the importance of different extratropical modes in subseasonal prediction.

b. *The initial optimal forcing in the tropics*

In this section, we show the results of initial optimal forcing patterns in the tropics, which lead to the growth of different extratropical Z500 modes, as a function of τ in both RMM-based and OLR-based LIMs. The initial optimal forcing pattern is derived by solving the eigenvectors of the generalized eigenvalue problems in (6). The pattern indicates the initial perturbation (i.e., $\mathbf{x}_0$), which gives rise to the maximum growth rate under a given norm $\mathbf{N}$ in (7). For example, if one specifies a norm of $\mathbf{N}_{z_1}$, the ultimate pattern of $\mathbf{x}_\tau$ will be the pattern of EOF1 shown in Fig. 1. Similarly, a norm of $\mathbf{N}_{z_2}$ yields a pattern that evolves and amplifies into the PNA pattern. Thus, by replacing $\mathbf{N}$ with various norms (e.g., $\mathbf{N}_{z_1}$, $\mathbf{N}_{z_2}$, $\mathbf{N}_{z_3}$,...), we can identify the corresponding tropical forcing pattern for each extratropical mode. Figure 3 shows the optimal initial RMMs (i.e., x_0) projected onto an MJO phase diagram from the RMM-based LIM. The numbers within each dot on the phase diagram are the time difference between the initial pattern and the ultimate pattern (i.e., $\tau - 0$). For example, when $\tau = 14$, the dots are located in the first quadrant (between MJO phases 5

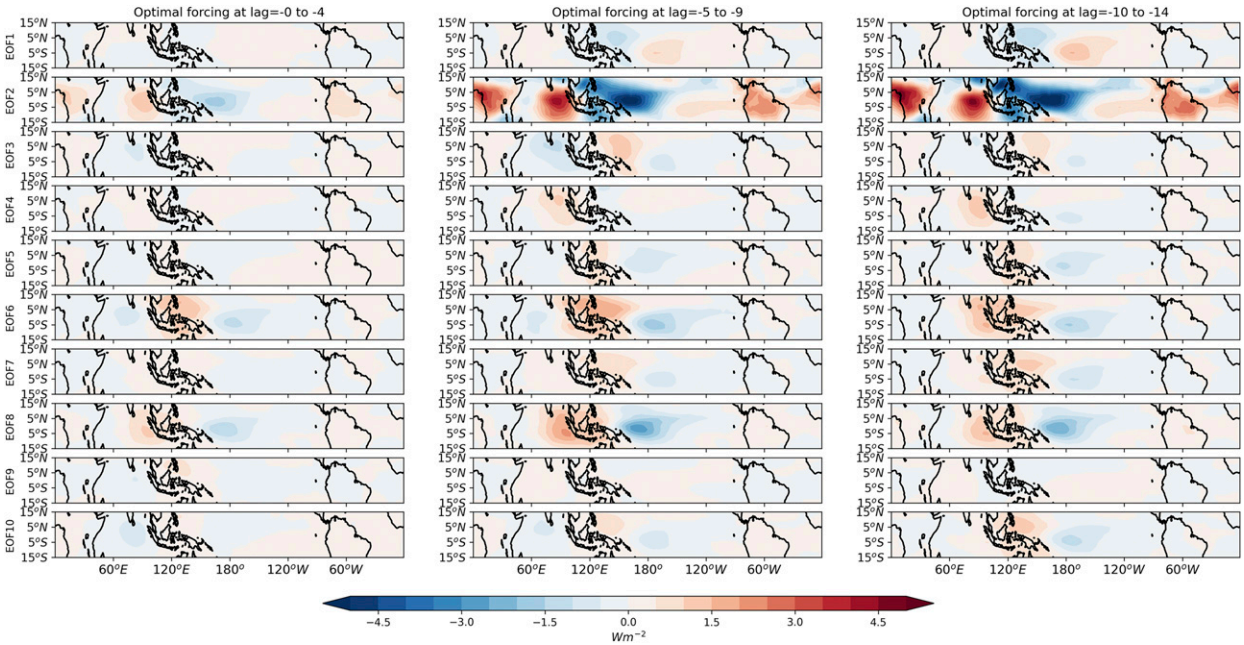

FIG. 4. The optimal tropical OLR forcing patterns for the development of each extratropical mode (rows) as a function of lag (columns). Unit is W m^{-2}. Results are from the OLR-based LIM.

and 6) for EOF2 (Fig. 3b). This indicates when we observe MJO phase 5/6 forcing, the forced positive phase PNA signal is expected to reach its maximum amplitude 14 days later. (For the forcing of the negative PNA, the optimal initial RMMs would be rotated 180° to MJO phases 1/2.) The distance of a point from the origin represents the strength of interaction between the MJO forcing and the extratropical response.

From Fig. 3, one can find that among 10 leading Z500 EOFs, only the second EOF (i.e., the PNA pattern) can be strongly driven by the MJO forcing, while other extratropical modes show either a weak or no interaction with MJO forcing. In addition, some preference for a dipole MJO heating pattern (e.g., phase 6, convection over the western Pacific and subsidence over the eastern Indian Ocean) can be found in Fig. 3b, which is consistent with previous studies. Seo and Lee (2017) and Tseng et al. (2019) have documented that the convective heating on each side of the Maritime Continent can trigger similar PNA patterns but with opposite sign. The positive heating to the east (west) of the Maritime Continent can generate a negative (positive) PNA pattern based on a numerical experiment in a linear baroclinic model. Thus, a dipole heating pattern about the Maritime Continent can lead to a more robust PNA signal due to the constructive interference of similar wave signals generated by each heating center.

Our argument that only the PNA is strongly excited by tropical convection is based on our definition of the PNA as EOF2. We note that other definitions of the PNA may include contributions from other EOFs. For example, the definition of the PNA used by the NOAA Climate Prediction Center (CPC) includes a substantial contribution from EOF1 (not shown). However, even in that case, the EOF2 contribution to the PNA is the component that is primarily driven by MJO-related tropical convection. This can be demonstrated by leaving out the contributions of EOFs 1–10 to the CPC version of the PNA and repeating the calculations shown in Fig. 3. We find the connections between MJO and the PNA are greatly reduced when we leave out EOF2 but enhanced when we leave out EOF1, and all other EOFs have little contribution to the PNA (figure not shown). This indicates that only the EOF2 component of the CPC PNA is excited by the MJO.

One might be curious if we replace the RMMs with tropical OLR PCs, can we still reach the same conclusion? In Fig. 4, we show the results of the second LIM, which replaces the RMMs with the tropical OLR PCs. The columns denote different τ and rows correspond to different extratropical modes. The shading in Fig. 4 shows the initial optimal convection pattern in the tropics. The OLR patterns share similar features among many of the extratropical modes, which may reflect, in part, the truncation to 10 OLR PCs, a choice that removes much of the smaller-scale OLR features. However, the conclusion is not sensitive to this choice of truncation: similar to what we found in Fig. 3b, the PNA (EOF2) features an initial optimal tropical convective forcing pattern of substantial amplitude for lags beyond −5 days. In addition, the preferred spatial structure shows an MJO phase 6 pattern with a positive heating located in the western Pacific and a negative heating located in the eastern Indian Ocean. This again strengthens the conclusion acquired from Fig. 2, that the PNA is the leading mode of the first 10 EOFs over the extratropical Pacific strongly driven by tropical forcing, while other extratropical modes show a modest interaction with tropical convection.

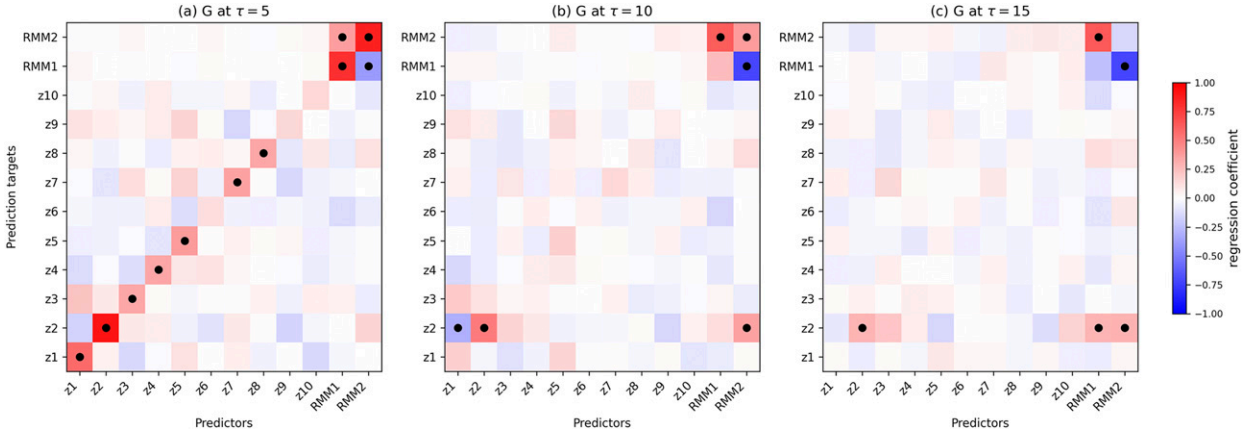

FIG. 5. The **G** matrix for $\tau = 5$, 10, and 15 days in the RMM-based LIM. Dotted regions indicate the values are significantly different from zero at the 1% level based on the t test. The x axis corresponds to the input variables, and the y axis corresponds to the prediction target.

Given that ENSO forcing of the PNA is a dominant driver of seasonal predictability, one may wonder why a clear ENSO signal does not emerge from the analysis above. For the subseasonal time scales considered here and for daily averaged data, the MJO forcing of the PNA dominates over the ENSO forcing. If we extend the LIM to focus on optimal initial OLR patterns for PNA forcing at lags beyond 30 days (not shown), then the typical ENSO OLR pattern emerges. However, the signal on daily time scales is weak, and so temporal averaging (e.g., monthly or seasonal) would allow the ENSO signal to emerge more clearly. Henderson et al. (2020) identified a similar result, whereby the tropical initial OLR signals are insignificant at a lead longer than 40 days (their Fig. 10b). They also found including SST in the state vectors helps identify the ENSO–PNA teleconnection, while the MJO signal still dominates at shorter time scales (their Fig. 3b).

c. *The initial optimal forcing in the extratropics*

Similar approaches can be used to identify the initial optimal forcing pattern over the extratropics for triggering these Z500 leading EOFs. Figure 5 shows the **G** matrix of RMM-based LIM at $\tau = 5$, 10, and 15 days, which represents the interaction between different modes. The amplitude of **G** denotes the strength of interaction. The x axis is the predictor, and the y axis is the prediction target. To examine the statistical significance (i.e., null hypothesis: $\mathbf{G} = 0$), we applied a t test to each component of **G**, where the effective degrees of freedom N^* is used. We define N^* as $N\Delta t/2\tau$, where $N\Delta t$ is the total length of data and 2τ is 2 times the e-folding time of the autocorrelation function. Dotted regions in Fig. 5 indicate the value is significantly different from zero at the 1% level.

From Fig. 5a, we can find that all of the significant values are concentrated along the diagonal element of **G**, indicating that the persistence of each leading mode dominates the linear predictable signals at short forecast leads. However, with the increase of forecast lead (i.e., Fig. 5c), the elements that originally have significant values vanish. Instead, most of the predictable signals are related to the growth of EOF2 (i.e., the row of z_2), which involves three processes: 1) transition from EOF1 to EOF2, 2) the persistence of EOF2, and 3) MJO-forced EOF2. While 3) represents the tropical–extratropical interactions, processes 1 and 2 are associated with the internal extratropical dynamics. Figure 1 indicates that EOF1 shows a similar wave propagating feature as with EOF2, which extends from Hawaii (20°N, 160°W) to Alaska and terminates around the southeastern United States. However, the MJO-forced extratropical response does not project onto EOF1 (i.e., no significant value in the bottom row of Fig. 5c) and its synoptic decorrelation time scale (~5.4 days) implies that the growth of EOF1 is dominated by extratropical synoptic eddies. Thus, the interaction between EOF1 and EOF2 implies that the growth of the PNA can also be triggered by internal extratropical dynamics, which is consistent with the conclusion in Black et al. (2017). The evolution of the anticyclonic anomaly over the extratropical Pacific into a PNA pattern is also documented by Franzke et al. (2011), which showed that some stochastic processes in the extratropics can also lead to the growth of PNA. If we focus on the explained variance by each leading mode, we can find that EOF1 and EOF2 are significantly separated from the rest of the leading modes (North et al. 1982), while they are not significantly separated from each other. This suggests the possibility that EOF1 and EOF2 are paired propagating waves. However, Fig. 5 also indicates that most of the extratropical modes show either weak or no interactions with each other at long forecast leads (>10 days, when most elements of **G** show insignificant values). This indicates that internal extratropical dynamics may provide limited information for the subseasonal prediction over the extratropical Pacific. We note that this conclusion is based on the linear assumptions of LIM and the use of 10 leading EOFs in this particular domain. Hu et al. (2019) found that the NPO–WP pattern shows a close relationship with the MJO convection by analyzing lagged composites of extratropical geopotential height in reanalysis data. In addition, the wave signals from upstream regions, such as Eurasia, potentially may provide additional predictability sources (Grazzini and Vitart 2015). The influence of domain

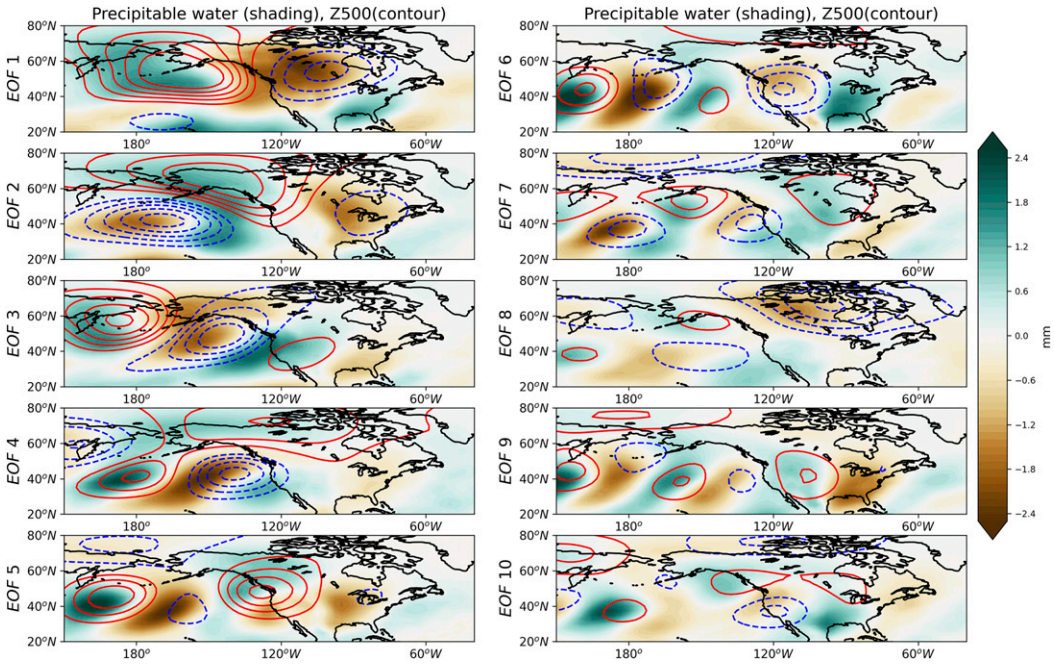

FIG. 6. Spatial pattern of a_i for precipitable water (shading, unit: mm) and the Z500 EOF (contours), which is identical to Fig. 1.

size on the current conclusions deserves further exploration in future study.

Given the results presented in Fig. 5, one may hypothesize that skillful subseasonal predictions over the PNA region, including for hydrological extremes, are dominated by the MJO teleconnection. Thus, the LIM-based subseasonal prediction of extratropical weather will be addressed in the next section.

4. Mapping red noise climate variability to hydrological extremes

a. Two-step linear regression

In this section, we examine the relationship between leading extratropical modes and the prediction of hydrological extremes over western North America. However, the LIM cannot be directly applied to the prediction of hydrological extremes because the statistics of hydrological fields are not Gaussian, indicating that the distribution of the model residuals would violate the inherent assumptions (see appendix B for further discussion). In addition, the hydrological fields show comparatively short memory (not shown), which implies the linear predictable components used for the prediction will be 0 if we apply a LIM to these variables. Thus, we propose a two-step linear regression, which enables us to map large-scale climate variability to the hydrological extreme without violating the τ test [i.e., a stable estimation of $\mathbf{B}$ in (5)]. Specifically, the PCs of the leading Z500 modes are first predicted by a LIM [i.e., Eq. (3)]. Then the anomalous fields, such as precipitable water or AR frequency, are predicted using a linear regression model with PCs of Z500 leading modes as predictors [Eq. (9)]:

$$y = \sum_{i=1}^{10} a_i z_i + b. \tag{9}$$

In (9), y is the target of prediction (e.g., Z500, precipitable water or AR frequency), z_i is the principal component of ith Z500 leading mode [i.e., $z_1, z_2, \ldots$ in (8)], a_i is the corresponding regression coefficient and b is the intercept (constant) of the regression model. For each grid point, we develop a single regression model. Physically, the sign of a_i represents the shift of the probability density function (PDF) of y with the change of z_i, which indicates the modulation on hydrological variables by large-scale climate variability. Here, we focus on deterministic predictions, given by y, but the model can be adapted for probabilistic predictions. For example, in the case of predicting AR frequency, if the predicted y is positive, the prediction is for an anomalously active AR state whereas negative y values indicate the prediction of a less active than normal AR state. With this model setup, we can leverage a LIM for the prediction of hydrological extremes without violating the τ test. We note that we use least squares regression because we are only focused on shifts in the distribution (i.e., binomial classification) rather than the prediction of the full distribution of these hydroclimate variables for which standard least squares regression would not be appropriate. The extensions to more informative probabilistic predictions would require other methods such as quantile regression. The discussion of "goodness of fit" can be found in the appendixes A and B.

Figures 6 and 7 show the spatial pattern of a_i for precipitable water and AR frequency, respectively. The shading represents the regression coefficient in (9) (i.e., a_i) for (precipitable water and AR) and the contour for Z500. In Figs. 6 and 7, one can

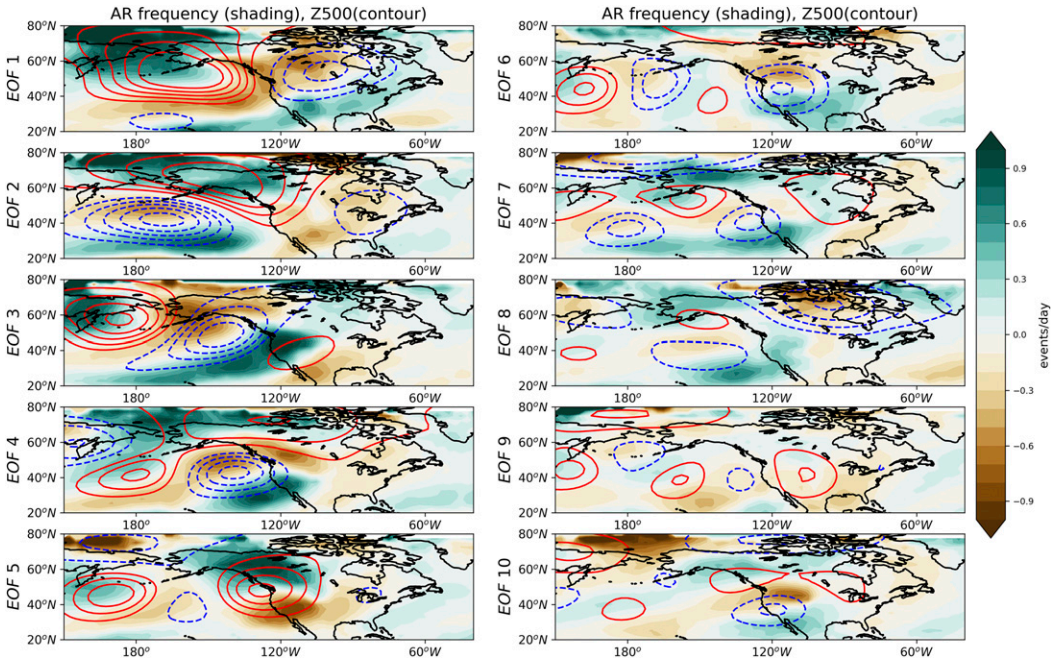

FIG. 7. As in Fig. 6, but for AR frequency (unit: day^{-1}).

observe the precipitable water and the AR frequency are strongly modulated by the large-scale circulations. Specifically, a dry anomaly of precipitable water and a negative moisture transport (less AR activity) typically are found to the southeast (northwest) of high (low) pressure anomalies. These spatial patterns can be simply explained by the steering flow pattern and the climatological distribution of moisture (i.e., meridional gradient and land–sea contrast). Despite that the regression patterns of AR frequency and precipitable water are alike in many places, some differences are still evident if one looks closely at Figs. 6 and 7. This can result from the difference in their definitions, as the detection of an AR takes the wind speed and geometry into account while precipitable water only incorporates moisture information. For example, near the region of 30°N, 180° of EOF2, the AR frequency regression on PC2 shows enhanced activity but the precipitable water regression has nearly zero amplitude. This region is characterized by strong meridional gradient of anomalous Z500 (contours in EOF2 of Figs. 6 and 7), which further leads to strong anomalous zonal wind and strong moisture flux in a confined region. A confined circulation with strong moisture flux implies an active AR state, which is not necessarily the case for the precipitable water. This difference also indicates why geometry plays a role in detecting ARs. When PC2 is elevated, we expect less frequent incursions of moisture into this region but more extreme rainfall.

Regardless of the difference in a_i between precipitable water and AR frequency, a strong modulation of large-scale climate variability in both fields indicates an existing potential of leveraging a LIM for the extended-range prediction of extreme events. However, there is still some restriction on applying a two-step linear regression to the prediction of hydrological extremes. First, the use of the leading Z500 EOFs ensures that most of the Z500 variability can be explained by the minimum amount of information, which may not be the case for hydrological variables since these are not the EOFs of the hydrological variables. Second, the probability density function of hydrological extremes can be highly skewed, which limits the application of a linear model to this problem, as discussed above and in appendix B. Despite these limitations, in the next section, we will examine the utility of applying a two-step linear regression to the prediction of hydrological extremes.

b. *The subseasonal prediction of hydrological extremes*

Here, we develop a two-class prediction scheme. Specifically, the OLR-based LIM is first used to predict the Z500 PCs over different forecast leads. Then, the predicted PCs are used to predict the shift in direction of the probability density function (i.e., higher or lower than the 50th percentile, see appendix A) for a given hydrological variable (e.g., precipitable water or AR frequency). The skill of the prediction scheme is evaluated using a leave-one-out cross validation and Heidke skill score (HSS). Specifically, the verification statistics for a given season are based on the climatological distribution excluding the one "left out" verification season, ensuring the independence of the verification subset. Thus, the 36-yr dataset yields 36 cross validations. The HSS for the 36 validation sets is then calculated as

$$\mathrm{HSS} = \frac{H - E}{T - E} \times 100, \tag{10}$$

where H is the number of correct forecasts (i.e., how many events correctly predict the sign of Z500, precipitable water or ARs), T is the total number of forecasts and E is the expected

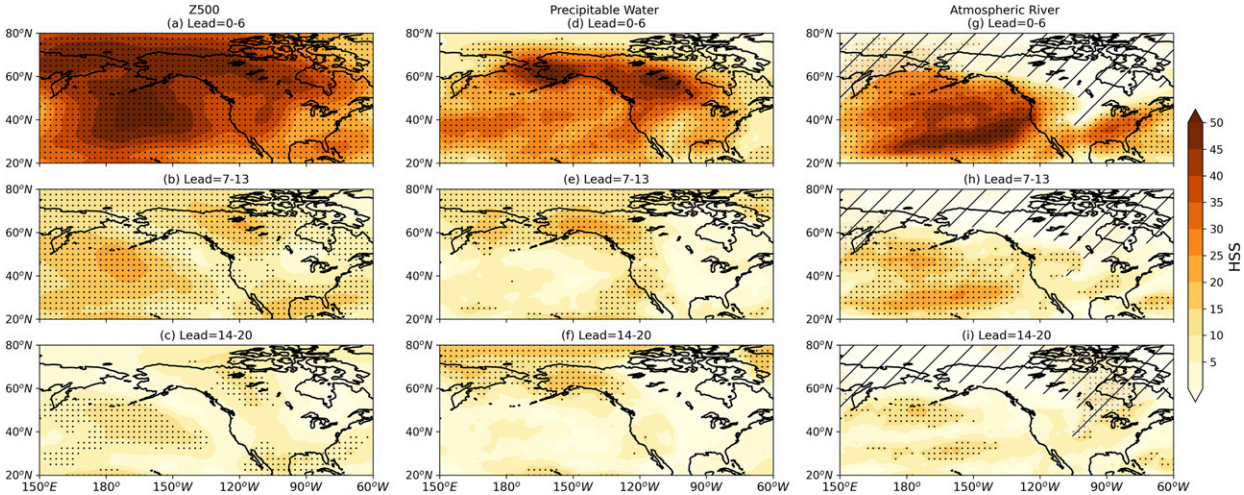

FIG. 8. The HSS for (left) Z500, (center) precipitable water, and (right) AR frequency for forecast days (top) 0–6, (middle) 7–13, and (bottom) 14–20. Dotted regions indicate the HSS value is significantly higher than that of a random forecast (50% correct chance) at the 5% level based on a binomial test. The hatching shows the regions where the frequency of AR climatology is less than 2%.

number of correct forecasts by random chance (binomial distribution, expected value is $T/2$ in the two-class scenario). The HSS can be interpreted as the prediction skill relative to the expected skill from a random forecast. For example, an HSS of 50 represents 3 times as many correct forecasts as incorrect forecasts and an HSS of 0 indicates the model prediction skill is equal to a climatological reference forecast, if a climatological forecast is defined as an equiprobable random draw.

For these calculations, we initialize the model every day from November to March and make daily forecasts for the following 5 weeks. Each forecast lead and grid point are evaluated independently. The HSS calculations are based on forecasts aggregated across all initialization dates (151 in total), years (36 in total), and weekly lead windows (7 per lead), and so H represents the number of correct forecasts among the total of 38 052 forecasts for the given grid point and lead window.

Figure 8 shows the 36-yr HSS for three variables: Z500, precipitable water, and atmospheric river frequency as a function of forecast lead time. Darker color indicates higher prediction skills and dotted regions indicate that the HSS is significantly higher than that of a random forecast at the 5% level by a binomial test. In Fig. 8, we can observe a few interesting features. First, although the prediction skills expectedly decrease with the increase of forecast lead, we can find limited regions where the Z500 skill scores remain elevated at all leads (i.e., first column of Fig. 8). Notably, these regions are spatially collocated with the PNA regions (Fig. 8c), which supports the finding in section 3 that the PNA is the extratropical mode showing the longest decorrelation time scale of any other mode. In addition, the analysis of initial optimal forcing also indicates that the PNA is one of a few leading modes that can be strongly driven by tropical forcing. These factors mainly explain why the regions characterized by better prediction skills at long forecast lead show a PNA-like pattern. Second, the skillful prediction of precipitable water and atmospheric river frequency only show up over specific regions even at short forecast leads. These regions are mostly concentrated around the extratropical Pacific and the west coast of North America.

As discussed earlier, one potential restriction on applying a two-step linear regression to the prediction of hydrological extremes is that the leading Z500 modes may explain most of the variance of daily Z500 but not necessarily of precipitable water or atmospheric river frequency. However, in support of the reliance on the leading Z500 modes for the basis of the hydroclimate predictions, only limited forecast skill improvements can be found when the additional Z500 modes are used (not shown). In addition, we see widespread regions over the extratropical Pacific showing skillful forecasts in days 0–6, demonstrating the importance of different extratropical modes in determining the prediction of hydrological extremes at short forecast leads. With the increase of forecast lead (i.e., days 7–13 and 14–20; Figs. 8e,f), only confined regions, such as the Gulf of Alaska and coastal California, show significant skill scores. Unsurprisingly, these are also the regions where the precipitable water is strongly modulated by the PNA circulation (i.e., the shading in Figs. 8a or 8f is similar to the shading of EOF2 in Fig. 6). Similar features can also be found in atmospheric river frequency, where the regions with high prediction skills are mostly concentrated around the Pacific basin and parts of the Atlantic storm track (Fig. 8g) at short forecast leads. At lead times of 7–20 days, the regions with high prediction skills of atmospheric river frequency shift to the subtropical Pacific, where the subtropical jet is characterized by strongest variability on intraseasonal time scales (figure not shown). As discussed previously, the detection of atmospheric rivers takes both intensity and geometry information into account, where the strong signal is typically found in the filament structure of extratropical storms. It is therefore intuitive that these regions show high prediction skills since the PNA can efficiently modulate the storm track variability over these regions. One should be aware that HSS values may be misleadingly high

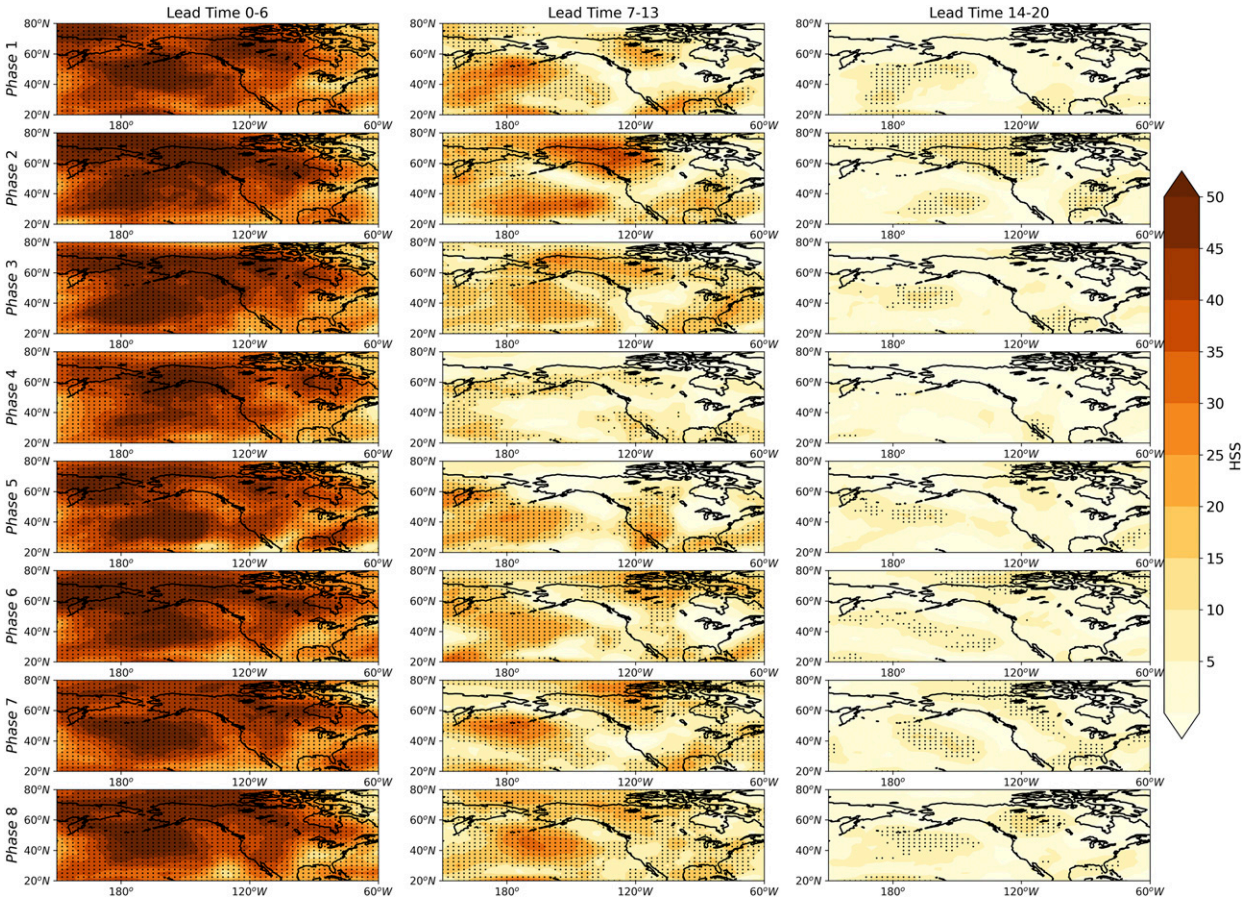

FIG. 9. The HSS of Z500 as a function of initial MJO phase (rows) and forecast lead time (columns). Dotted regions indicate the HSS value is significantly higher than that of a random forecast (50% correct chance) at the 5% level based on a binomial test.

over the regions with very rare AR occurrence, such as eastern Canada, since the highly skewed data allows a high proportion of hits for predictions of no AR occurrence. Thus, we mask out the regions where the frequency of AR climatology is less than 2%.

Previous studies have demonstrated that the prediction of extratropical circulations and associated hydrological extremes show a strong dependence on MJO phase (Mundhenk et al. 2018). Thus, the results shown in Fig. 8 do not account for the windows of forecast opportunity associated with preferred MJO phases since all winter days are used. In addition, the nonzero off-diagonal elements of **G** allow for nonmodal growth of different extratropical modes in short time scales, which may benefit the prediction of non-PNA patterns. Seo and Lee (2017) and Tseng et al. (2019) demonstrated that the MJO-induced PNA signal is only robust in specific phases due to the wave interference (see section 3b), while the prediction in other phases may be dominated by other extratropical modes. It is therefore worthwhile to show the HSS as a function of forecast lead and initial MJO phase. Figures 9–11 display the HSS of Z500, precipitable water and AR frequency, respectively. Specifically, we partition by MJO phase according to the initial day (i.e., $\tau = 0$) with MJO amplitude greater than 1 standard deviation. In Fig. 9, we can find the HSS patterns are quite similar to those shown in Figs. 8a–c. At short forecast leads (i.e., week 1), the LIM does well in predicting the sign of Z500 for all MJO phases (whole domain is characterized by significant prediction skills). With the increase of lead time (weeks 2 and 3), only the PNA regions show significant skill scores. Two additional features from this figure stand out. First, the significant prediction skills shown in Figs. 8b and 8c are contributed by certain MJO phases (i.e., 2, 3, 6, 7, and 8), which is consistent with previous research (e.g., Lin et al. 2010; Tseng et al. 2018). Second, the non-PNA extratropical modes provide limited information on extended-range forecasts (2–3 weeks, phase 4) even when the PNA signal is not robust. This feature is evident in days 14–20 forecasts initialized in MJO phase 4. According to previous studies, the MJO teleconnection signal is weak in phase 4 due to the destructive interference of Rossby wave signals from both sides of the Maritime Continent. However, although the MJO–PNA signal is weak, there are still only limited regions showing significant skill scores, indicating that the predictable signals from other non-PNA modes are weak as well.

Figure 10 shows the HSS of precipitable water. Similar to the feature found in Fig. 8, most regions are characterized by

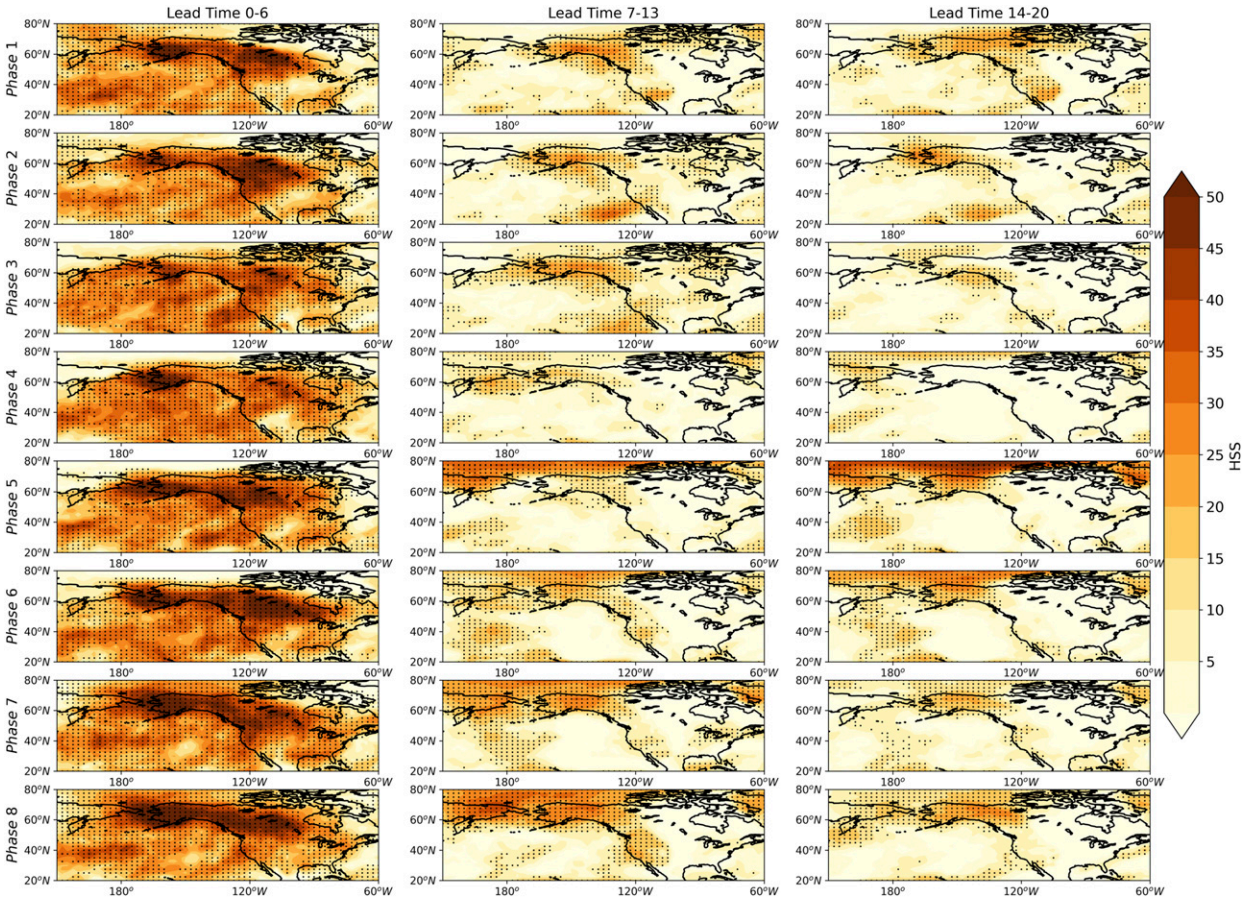

FIG. 10. As in Fig. 9, but for precipitable water (November–March 1979–2017).

significant skill scores at short forecast leads (i.e., the first week). In particular, Alaska and the west coast of North America (e.g., California and British Columbia) have the highest values over the domain. This feature remains evident in the second and the third weeks while the skill scores in other regions decrease dramatically. One might be curious why the regions that show high prediction skills in precipitable water and Z500 are not spatially collocated. The reason is that precipitable water prediction is more closely tied to the influence of atmospheric circulation than to the mass fields. According to geostrophic balance, the mass fields, as represented by Z500, and the atmospheric circulation are orthogonal in space, indicating that the regions with high AR prediction skills will tend to be orthogonal in space to the Z500 field as well. The MJO phase-dependent features and PNA-related precipitable water pattern are evident in the second and third forecast weeks, indicating that the MJO–PNA teleconnection is the dominant source of skill at longer lead times while other extratropical modes are important in week 1. The analysis of atmospheric river frequency forecast skill (Fig. 11) generally yields the same conclusion.

The other interesting feature shown in Figs. 9–11 is that the regions with high skill scores are not symmetric between the first half (i.e., phases 1–4) and the second half (i.e., phases 5–8) of the MJO life cycle. For example, the atmospheric river frequency shows highest skill scores in Alaska and the subtropical Pacific for MJO phases 1–4, weeks 2–3 (Fig. 11). However, the regions with high prediction skills shift to the Great Plains for MJO phases 5–8. Given that the patterns of MJO-related OLR anomalies are identical but with opposite sign between the first and second half of the MJO life cycle, one may hypothesize that such symmetry also would be observed in the extratropical response. However, Figs. 9–11 suggest that this hypothesis does not hold. One possible explanation is that the initial extratropical pattern is not symmetric between these two stages and thus leads to the growth of non-PNA modes. However, the asymmetric skill does not imply that the prediction skill of hydrological extreme is not a function of MJO phases. Instead, it suggests that some nonlinear processes cannot be captured by a simple linear model. This result is consistent with Mundhenk et al. (2018), which uses a binomial model with MJO phases as predictors. Since the asymmetry of extratropical internal dynamics is beyond the scope of this study, we will address this hypothesis in future study. Thus, we will focus on the role of tropical–extratropical teleconnection on extended (i.e., 2–4 weeks) prediction in the following analysis.

To support the conclusion that the MJO-induced teleconnection is the dominant predictability source at long forecast

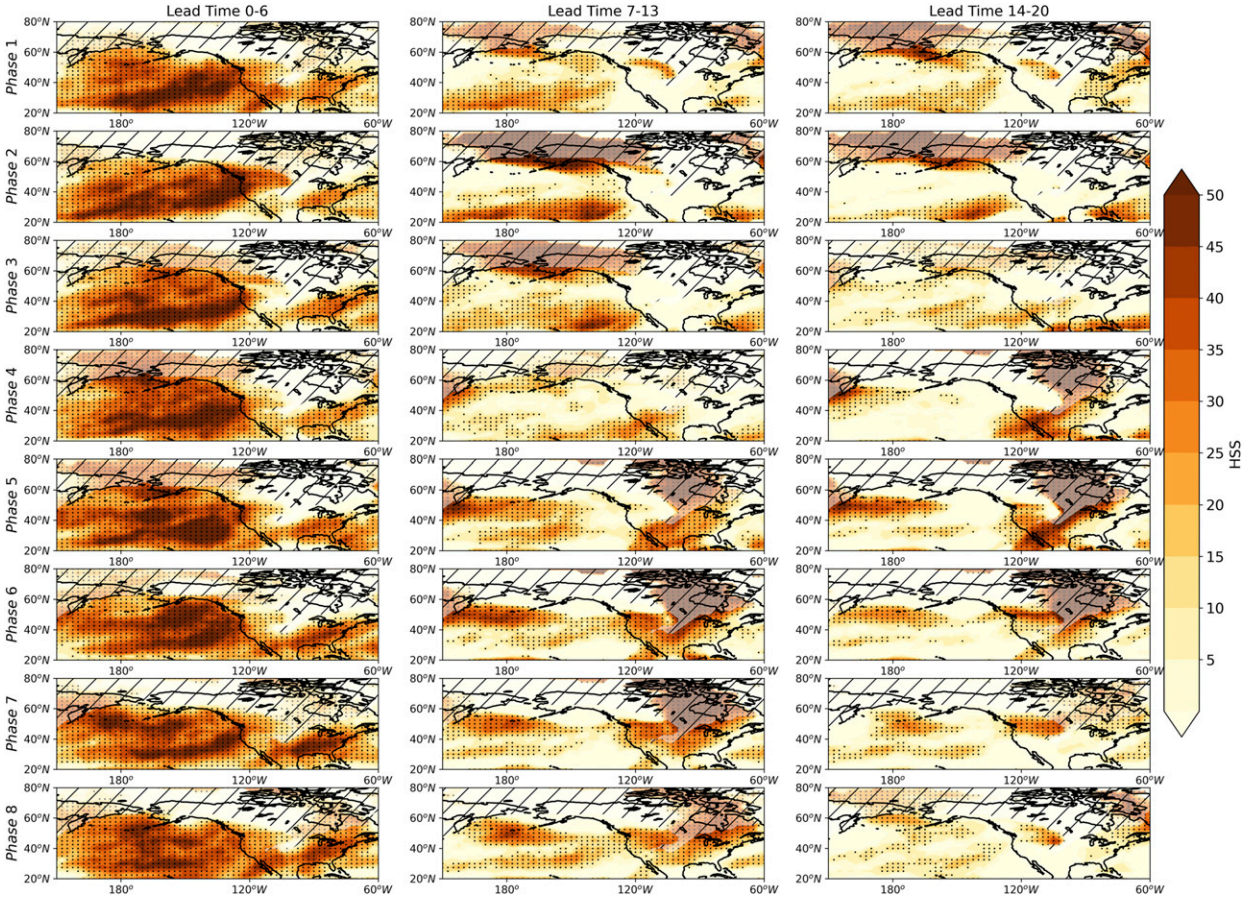

FIG. 11. As in Fig. 9, but for atmospheric river frequency (November–March 1979–2017).

leads, we carried out an additional experiment in which the PCs of every extratropical leading mode at $\tau = 0$ are set to zero (i.e., $[z_1, z_2, \ldots, z_{10}] = 0$ in Eq. (8)). This model setup inhibits the growth of any disturbance by the internal extratropical dynamics and only allows the tropical forcing to influence the extratropics. The result is demonstrated in Fig. 12, which shows the fractional area with significant skill scores in both the original LIM (OLR-based LIM) and in the version for which the initial conditions of the extratropical modes have been set to zero. From this analysis, we find that the original LIM skillfully predicts the sign of Z500, precipitable water and atmospheric river frequency in most regions (>60%) for the first 5 days, with the greatest skill for Z500 forecasts. The tropical–extratropical teleconnection only accounts for about 20%–30% of the significant fractional area. However, with the increase of lead time (>15 days), the MJO teleconnection emerges as the dominant signal for all variables. It is worth mentioning that the extratropics-removed initial condition experiment has even better skill than the original LIM at some long forecast leads, particularly for AR frequency although these differences generally are modest. This is because the extratropical modes introduced at the initial state may be a greater source of noise than skill at the longest lead. This experiment supports the finding in section 3c that the internal extratropical dynamics are only beneficial for week 1 forecasts due to their short decorrelation time scales, while the MJO teleconnection is the dominant signal at longer lead times.

5. Concluding remarks

The role of the MJO–PNA teleconnection in modulating the extratropical circulation and associated hydrological extremes has been extensively examined in the past decade. However, other extratropical modes, which potentially can influence the hydrological extremes, have received comparatively little attention. By using a linear inverse model, this study explores the role of internal extratropical dynamics and tropical–extratropical teleconnection in subseasonal prediction. In general, two processes jointly determine the importance of different modes in modulating extratropical circulation over different time scales: 1) the necessary time scales for the development of the tropical–extratropical teleconnection and 2) the *e*-folding time of each extratropical mode. These two processes are illustrated in Fig. 13. An initial optimal forcing analysis indicates that the PNA pattern is one of a few leading modes (the first 10 EOFs) over the extratropical Pacific that can be strongly driven by the MJO forcing, while other modes show either weak or no interaction with tropical forcing. In

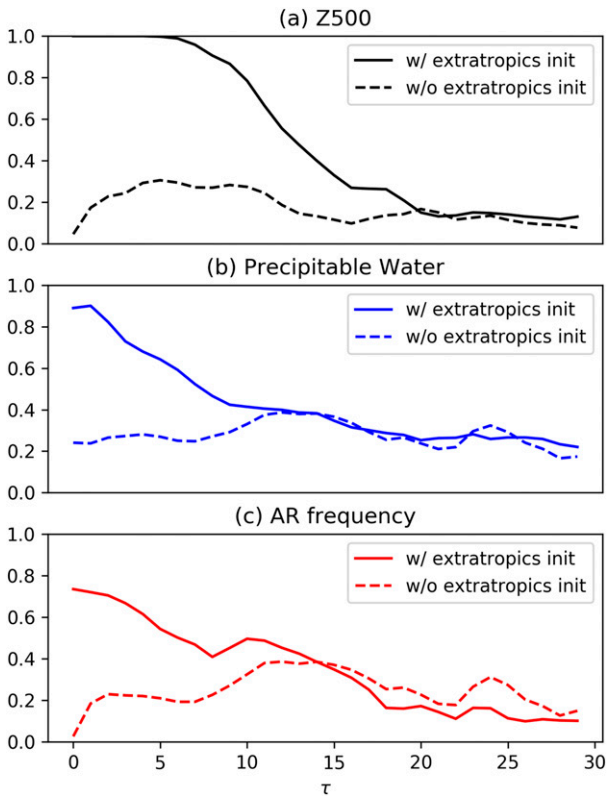

FIG. 12. The fractional area of statistically significant (5% level) skill scores over the domain 20°–80°N, 120°E–90°W in the original LIM (solid lines) and in the version for which the initial conditions of the extratropical modes have been set to zero (dashed lines).

addition, an autoregression analysis shows that the NPO–WP (i.e., EOF1) and PNA pattern (i.e., EOF2) are characterized by the longest decorrelation time scales than any other modes. Thus, for time scales shorter than a week (process 2 in Fig. 13), every extratropical mode contributes a certain amount of information to the extratropical prediction. However, with increasing lead time, due to the short *e*-folding time of non-PNA modes (<7 days) and the necessary time scales (~14 days, process 1 in Fig. 13) for the development of the MJO–PNA teleconnection, the PNA is the only extratropical mode beneficial for subseasonal prediction for lead times beyond about 2 weeks, at least for western North America.

In the second part of this study, we proposed a two-step linear regression that maps large-scale climate variability to hydrological extremes. This setup enables us to use a LIM to predict hydrological extremes without violating the τ test. The predictions of Z500, precipitable water, and atmospheric river frequency are consistent with the earlier results in that the predictable signals are dominated by the MJO–PNA teleconnection for time scales longer than 2 weeks.

This study provides a likely explanation for why the PNA pattern is a dominant source of skill for time scales longer than 2 weeks when other modes have modest impact at this forecast lead time. However, there are a few questions that remain unanswered. First, the assumption that nonlinear processes decorrelate much faster than the deterministic linear components may be badly violated if the nonlinear feedback is slow. Most of the long-term signals (i.e., >3 months) in the atmosphere are determined by the slow nonlinear processes such as the asymmetry between El Niño and La Niña (An and Jin 2004) or troposphere–stratosphere interactions (Domeisen et al. 2020), which are excluded in the LIMs of this study. Henderson et al. (2020) has demonstrated the ENSO dynamics plays a nonnegligible role in determining the growth of the PNA. In addition, other low-frequency variability such as quasi-biennial oscillation (QBO) can modulate the MJO teleconnection as well (Feng and Lin 2019). Thus, the dependence of the **G** matrix change on climate state and the incorporation of slow nonlinear processes are areas of ongoing research. Second, the linear regression in (9) might not be the most ideal approach of mapping climate variability to the hydrological extremes or other variables since the underlying distribution can be highly skewed. The test of Gaussianity in appendix B shows that extensions to more informative probabilistic predictions would require other methods such as quantile regression. Sardeshmukh et al. (2000) and Tseng et al. (2020) demonstrated the spread of daily Z500 PDF varies in different ENSO states, suggesting the importance of the nonlinear dynamics for probabilistic forecasts. Thus, taking the underlying distribution into account is a possible direction for future research. Third, some previous studies such as Xiang et al. (2020) demonstrated the NPO–WP can be skillfully predicted to lead times up to 3 weeks in a state-of-the-art dynamical forecast model, while our study suggests that MJO–PNA is the only predictability source at this time scales. A few reasons might lead to this difference. The domain that we used is limited to the extratropical Pacific and part of North America, while some signals from upstream regions (e.g., Eurasia) can be additional predictability sources (Grazzini and Vitart 2015). In addition, the selection of state vectors also influences the predictable signals. Henderson et al. (2020) demonstrated that including SST in the state vectors helps identify the predictable signals in tropical convection at the longer leads (≥40 days). These processes deserve additional analysis in future work.

The insight we have gained about tropical–extratropical interaction through the use of LIM raises some interesting questions for future work. First, what other processes contribute to subseasonal predictability in this region, and how well can extensions of this LIM capture these processes? For example, would a LIM that incorporates troposphere–stratosphere interaction yield better skill, and if so, what modes would be responsible for the increase in skill? Second, why is the PNA the only mode that is strongly modulated by the MJO? Third, in dynamical forecast and global climate models, does the bias in tropical convection preferentially impact the PNA-related circulation while having little impact on all other extratropical modes? All of these questions deserve further exploration in future studies.

Acknowledgments. This research has been conducted as part of the NOAA MAPP S2S Prediction Task Force and supported by NOAA Grant NA16OAR4310064 and NA18OAR4310296 and by the Climate and Large-Scale Dynamics Program of the National Science Foundation under grant AGS-1841754. K.-C. Tseng was supported by Award NA18OAR4320123 from the NOAA, U.S. Department of Commerce and by the NOAA Office of Water and Air Quality FACETs program. K.-C. Tseng was also supported by the NOAA OAR Weather

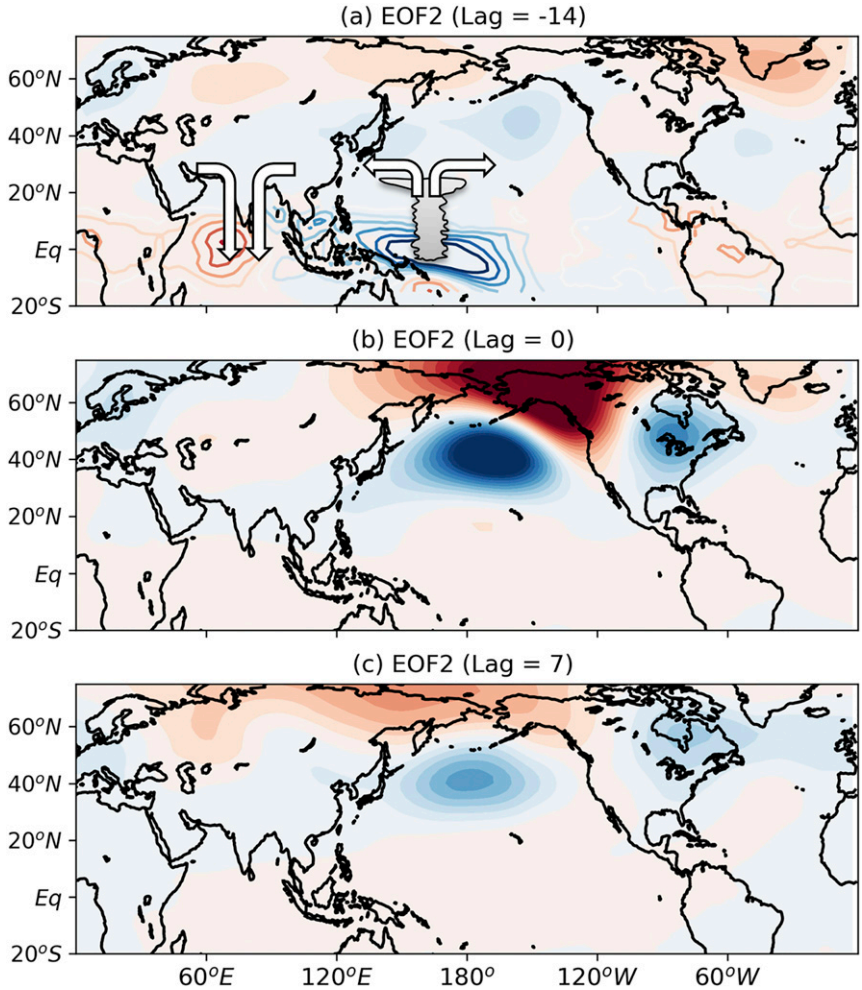

FIG. 13. Two processes in determining the predictable signals over different time scales. From (a) to (b) is process 1, which shows the tropical convection forcing the PNA pattern and the necessary time for the development of the tropical–extratropical teleconnection, which is about 14 days (only the PNA pattern is strongly excited by process 1). From (b) to (c) is process 2, which is the memory of the extratropical response, which depends on the *e*-folding time of each mode. The *e*-folding times range from 1.4 days (EOF9) to 7.2 days (EOF2). Shading shows the lagged Z500 composite anomalies for times when the PCs of extratropical leading modes are equal or greater than one standard deviation at lag 0. Blue shading indicates negative values and red shading indicates positive values. The values of Z500 ranges from −100 to 100 m with an interval 10 m. Contours are the lagged OLR composite anomalies. The values of contours range from $-10\,\mathrm{W\,m^{-2}}$ (blue) to $10\,\mathrm{W\,m^{-2}}$ (red) with an interval of $1\,\mathrm{W\,m^{-2}}$.

Portfolio grant. We also thank Ji-Jen Sun for using his cloud cartoon in Fig. 13.

APPENDIX A

Examining the Goodness of Fit by a Binomial Distribution and Eq. (9) in a Two-Class Classification Problem

a. Testing goodness of fit by a binomial distribution

The linear assumption of (9) can be badly violated if the underlying distribution of the predictand is highly skewed (see appendix B). However, this violation may not invalidate the analysis if we only focus on the directional shift of the PDF since we only consider the sign of the predicted anomaly rather than the amplitude or distribution of the predictand. Thus, in this appendix, we examine the suitability of using (9) and a binomial distribution for modeling the sign of the precipitable water and atmospheric river frequency anomaly. Specifically, we will first test the following null hypothesis and then evaluate the goodness of fit by (9):

H_0: The shift direction of the PDF (i.e., greater/smaller than 50th percentile) for a given hydrological variable follows a binomial distribution.

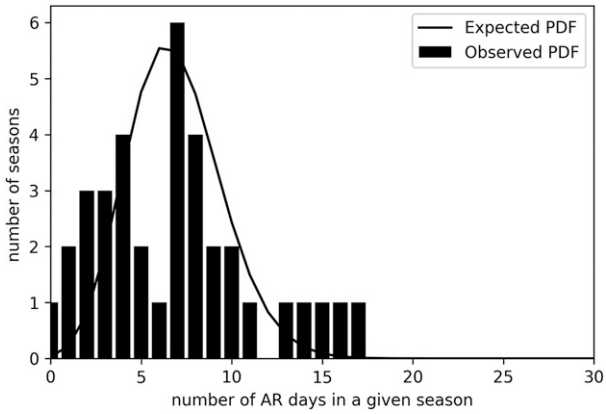

FIG. A1. The histogram of atmospheric river occurrence (bars) and the expected PDF (line) given by Eq. (A1) at 27.5°N, 110°W.

To test the hypothesis, we divide the winter season data into different subsets according to year (i.e., 1979/80 NDJFM, 1980/81 NDJFM, etc.); 36 seasons yield 36 subsets with each subset having a length about 151 days (November–March). Then, each day in a given season is marked with 1 if AR or precipitable water anomaly is positive; days are marked with 0 otherwise. Following standard binomial classification terminology, we call all days marked with 1 "success trials." Then, the sum of success trials for each season gives us a total indicating the number of days with positive AR or precipitable water anomalies. The bar graph in Fig. A1 shows the PDF according to H_0 from these 36 subsets over an example location of 27.5°N, 110°W. The x axis represents the number of days in a season with positive AR frequency anomaly and the y axis represents the number of seasons corresponding to each bin. For example, the peak around $x = 7$ and $y = 16\%$ indicates that a total of 16% of the seasons (~6 seasons) had 7 active AR days. To compare this histogram with the corresponding binomial distribution, we use the following formula to derive the theoretical binomial PDF:

$$\Pr(x = k) = \binom{n}{k} p^k (1-p)^{n-k}, \tag{A1}$$

where $\Pr(x = k)$ is the probability density when $x = k$, n is the number of total trials (i.e., ~151 days in this case), and k is the number of success trials (e.g., number of active AR days in the case of Fig. A1, which ranges from 0 to 151) and p is the climatological frequency (e.g., total AR days in 36 seasons divided by total number of days). The solid curve in Fig. A1 shows the theoretical curve based on Eq. (A1). The next step is using the chi-square to test if the observed histogram and the theoretical PDF are significantly different from each other. The chi-square statistic is given by

$$\chi^2 = \sum_{i=1}^{i=n} \frac{(O_i - E_i)^2}{E_i}, \tag{A2}$$

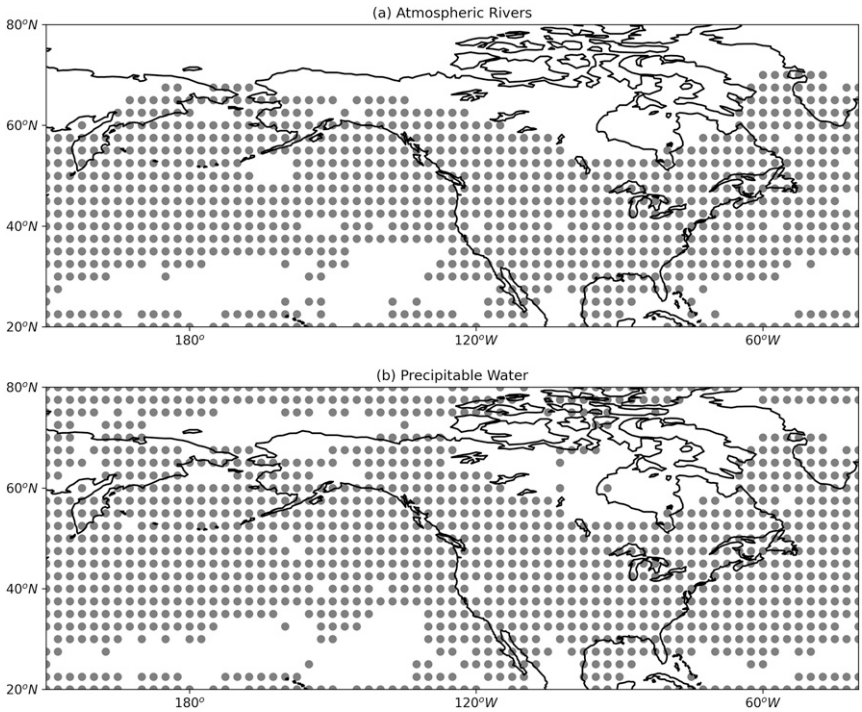

FIG. A2. Stippling indicates grid points where the chi-square statistic is smaller than 185, which is the 95% confidence level for (a) atmospheric rivers and (b) precipitable water. Mathematically, this indicates the PDFs over these regions are not significantly different from that of a binomial distribution.

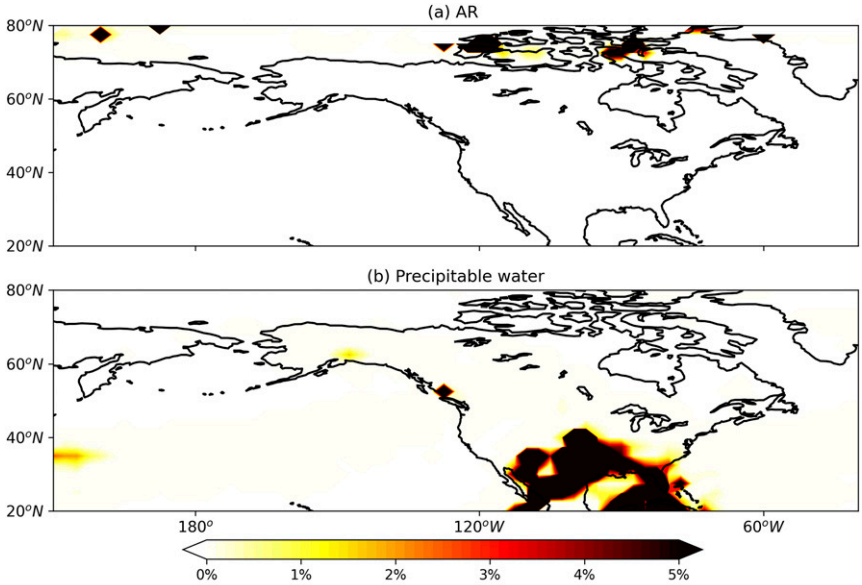

FIG. B1. The p value of ξ based on Jarque–Bera test for (a) atmospheric river and (b) precipitable water prediction.

where O_i is observed values (i.e., bar in Fig. A1) and E_i is the expected values (i.e., curve in Fig. A1). In the case that the null hypothesis is correct, the chi-square statistic follows a chi-square distribution. For sufficiently large values of the statistic, we can reject the null hypothesis. In our case, the 95% confidence interval for chi-square statistics is about 185 with a given degrees of freedom = 149 [151 bins (days) − 2 degrees of freedom due to the preprocessing]. As long as the estimated chi-square statistic is greater than 185, we reject the null hypothesis at the 5% level. In Fig. A2, we show the regions where we cannot reject the null hypothesis (i.e., binomial distribution is a reasonable approximation). We find that the binomial distribution null hypothesis cannot be rejected over most parts of the domain. The regions where the null hypothesis cannot be rejected correspond with those where AR or precipitable water variability is most active (e.g., Pacific or Atlantic storm track).

b. *Comparing goodness of fit by (9) with random binomial forecast*

In this part of appendix, we examine if Eq. (9) is a better model for forecasting the shift direction of the PDF than the random binomial forecast [i.e., $p = 0.5$ in (A1)]. Similarly, we first generate the PDF of the successful forecast by (9) as what we did in the bar plot of Fig. A1. The only difference is that the x axis represents the number of successful forecasts in a given season rather than active AR days shown in Fig. A1. We also derived the theoretical PDF with $p = 0.5$ by using Eq. (A1). To test the statistical significance, we can check if the observed PDF falls outside the 95% confidence level of the theoretical PDF. The result is shown in Fig. 8, where scattered regions indicate the prediction made by Eq. (9) is superior to random binomial forecast. Similar approaches can be applied to different subsets such as dividing the data into subsets according to the MJO phases (i.e., Figs. 9–11).

APPENDIX B

Examining the Gaussianity of Eq. (9)

In this appendix, we examine the suitability of using (9) for modeling the amplitude of precipitable water and atmospheric river frequency anomalies. First, we derive the residual term of regression models:

$$\xi = y_{\text{obs}} - y = y_{\text{obs}} - \left(\sum_{i=1}^{10} a_i z_i + b\right), \tag{B1}$$

where y_{obs} is the observed value of precipitable water or atmospheric river anomaly, y is the predictive value, and ξ is the residual term. We then estimate skewness γ and kurtosis K of ξ based on the following formulas:

$$\gamma = \frac{\frac{1}{n}\sum_{j=1}^{j=n}(\xi_j - \bar{\xi})^3}{\sqrt{\frac{1}{n}\sum_j(\xi_j - \bar{\xi})^2}^{\,3}} \quad \text{and} \tag{B2}$$

$$K = \frac{\frac{1}{n}\sum_{j=1}^{j=n}(\xi_j - \bar{\xi})^4}{\left[\frac{1}{n}\sum_{j=1}^{j=n}(\xi_j - \bar{\xi})^2\right]^2} - 3, \tag{B3}$$

where n is the total number of events, ξ_j is the residual of jth events, and $\bar{\xi}$ is the averaged residual for all events [which is 0 since b is

included in (9)]. With these definitions, a Jarque–Bera (JB) test is applied to (B1) to examine the Gaussianity of residual term:

$$\mathrm{JB} = \frac{n-k}{6}\left(\gamma^2 + \frac{1}{4}K^2\right), \tag{B4}$$

where k is the number of regressors used in (B1). If residual term is well-behaved, JB will approach a chi-squared distribution. Hence, once γ and K are estimated p values can be calculated to test the assumption of non-Gaussianity: if the p value of each of these is less than 5% the null hypothesis that the data is normally distributed can be rejected at the 5% level. Figure B1 shows the p value for the precipitable water and AR frequency predictions based on (B4). As we can see, the p value is less than 5% over most regions, indicating that we can reject the null hypothesis of Gaussian residuals. The most notable exception is for precipitable water around part of the Great Plains and the Gulf of Mexico.

REFERENCES

Albers, J. R., and M. Newman, 2019: A priori identification of skillful extratropical subseasonal forecasts. *Geophys. Res. Lett.*, **46**, 12 527–12 536, https://doi.org/10.1029/2019GL085270.

An, S.-I., and F.-F. Jin, 2004: Nonlinearity and asymmetry of ENSO. *J. Climate*, **17**, 2399–2412, https://doi.org/10.1175/1520-0442(2004)017<2399:NAAOE>2.0.CO;2.

Bjerknes, J., 1969: Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev.*, **97**, 163–172, https://doi.org/10.1175/1520-0493(1969)097<0163:ATFTEP>2.3.CO;2.

Black, J., N. C. Johnson, S. Baxter, S. B. Feldstein, D. S. Harnos, and M. L. L'Heureux, 2017: The predictors and forecast skill of Northern Hemisphere teleconnection patterns for lead times of 3–4 weeks. *Mon. Wea. Rev.*, **145**, 2855–2877, https://doi.org/10.1175/MWR-D-16-0394.1.

Cash, B. A., and S. Lee, 2001: Observed nonmodal growth of the Pacific–North American teleconnection pattern. *J. Climate*, **14**, 1017–1028, https://doi.org/10.1175/1520-0442(2001)014<1017:ONGOTP>2.0.CO;2.

Charney, J. G., 1963: A note on large-scale motions in the tropics. *J. Atmos. Sci.*, **20**, 607–609, https://doi.org/10.1175/1520-0469(1963)020<0607:ANOLSM>2.0.CO;2.

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.

Domeisen, D. I. V., and Coauthors, 2020: The role of the stratosphere in subseasonal to seasonal prediction: 2. Predictability arising from stratosphere–troposphere coupling. *J. Geophys. Res. Atmos.*, **125**, e2019JD030923, https://doi.org/10.1029/2019JD030923.

Feng, P.-N., and H. Lin, 2019: Modulation of the MJO-related teleconnections by the QBO. *J. Geophys. Res. Atmos.*, **124**, 12 022–12 033, https://doi.org/10.1029/2019JD030878.

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, https://doi.org/10.1002/qj.768.

Grazzini, F., and F. Vitart, 2015: Atmospheric predictability and Rossby wave packets. *Quart. J. Roy. Meteor. Soc.*, **141**, 2793–2802, https://doi.org/10.1002/qj.2564.

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.

Henderson, S. A., E. D. Maloney, and E. A. Barnes, 2016: The influence of the Madden–Julian oscillation on Northern Hemisphere winter blocking. *J. Climate*, **29**, 4597–4616, https://doi.org/10.1175/JCLI-D-15-0502.1.

——, D. J. Vimont, and M. Newman, 2020: The critical role of nonnormality in partitioning tropical and extratropical contributions to PNA growth. *J. Climate*, **33**, 6273–6295, https://doi.org/10.1175/JCLI-D-19-0555.1.

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.

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.

——, and T. Ambrizzi, 1993: Rossby wave propagation on a realistic longitudinally varying flow. *J. Atmos. Sci.*, **50**, 1661–1671, https://doi.org/10.1175/1520-0469(1993)050<1661:RWPOAR>2.0.CO;2.

Hsu, H.-H., 1996: Global view of the intraseasonal oscillation during northern winter. *J. Climate*, **9**, 2386–2406, https://doi.org/10.1175/1520-0442(1996)009<2386:GVOTIO>2.0.CO;2.

Hu, W., P. Liu, and Q. Zhang, 2019: Dominant patterns of wintertime intraseasonal surface air temperature over the CONUS in response to MJO convections. *Climate Dyn.*, **53**, 3917–3936, https://doi.org/10.1007/s00382-019-04760-x.

Lee, Y., and R. Grotjahn, 2019: Evidence of specific MJO phase occurrence with summertime California Central Valley extreme hot weather. *Adv. Atmos. Sci.*, **36**, 589–602, https://doi.org/10.1007/s00376-019-8167-1.

Liebmann, B., and C. A. Smith, 1996: Description of a complete (interpolated) outgoing longwave radiation dataset. *Bull. Amer. Meteor. Soc.*, **77**, 1275–1277, https://doi.org/10.1175/1520-0477-77.6.1274.

Lin, H., 2018: Predicting the dominant patterns of subseasonal variability of wintertime surface air temperature in extratropical Northern Hemisphere. *Geophys. Res. Lett.*, **45**, 4381–4389, https://doi.org/10.1029/2018GL077509.

——, G. Brunet, and R. Mo, 2010: Impact of the Madden–Julian oscillation on wintertime precipitation in Canada. *Mon. Wea. Rev.*, **138**, 3822–3839, https://doi.org/10.1175/2010MWR3363.1.

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, https://doi.org/10.2151/jmsj.86.213.

Mundhenk, B. D., E. A. Barnes, and E. D. Maloney, 2016: All-season climatology and variability of atmospheric river frequencies over the North Pacific. *J. Climate*, **29**, 4885–4903, https://doi.org/10.1175/JCLI-D-15-0655.1.

——, ——, ——, and C. F. Baggett, 2018: Skillful empirical subseasonal prediction of landfalling atmospheric river activity using the Madden–Julian oscillation and quasi-biennial oscillation. *npj Climate Atmos. Sci.*, **1**, 20177, https://doi.org/10.1038/s41612-017-0008-2.

Newman, M., and P. D. Sardeshmukh, 2017: Are we near the predictability limit of tropical Indo-Pacific sea surface temperatures? *Geophys. Res. Lett.*, **44**, 8520–8529, https://doi.org/10.1002/2017GL074088.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, **110**, 699–706, https://doi.org/10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2.

Penland, C., and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**,

1999–2024, https://doi.org/10.1175/1520-0442(1995)008<1999:TOGOTS>2.0.CO;2.

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.

——, G. P. Compo, and C. Penland, 2000: Changes of probability associated with El Niño. *J. Climate*, **13**, 4268–4286, https://doi.org/10.1175/1520-0442(2000)013<4268:COPAWE>2.0.CO;2.

Seo, K.-H., and H.-J. Lee, 2017: Mechanisms for a PNA-like teleconnection pattern in response to the MJO. *J. Atmos. Sci.*, **74**, 1767–1781, https://doi.org/10.1175/JAS-D-16-0343.1.

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**, 14 291–14 324, https://doi.org/10.1029/97JC01444.

Tseng, K.-C., E. A. Barnes, and E. D. Maloney, 2018: Prediction of the midlatitude response to strong Madden–Julian oscillation events on S2S time scales. *Geophys. Res. Lett.*, **45**, 463–470, https://doi.org/10.1002/2017GL075734.

——, E. Maloney, and E. A. Barnes, 2019: The consistency of MJO teleconnection patterns: An explanation using linear Rossby wave theory. *J. Climate*, **32**, 531–548, https://doi.org/10.1175/JCLI-D-18-0211.1.

——, ——, and ——, 2020: The consistency of MJO teleconnection patterns on interannual time scales. *J. Climate*, **33**, 3471–3486, https://doi.org/10.1175/JCLI-D-19-0510.1.

Tziperman, E., L. Zanna, and C. Penland, 2008: Nonnormal thermohaline circulation dynamics in a coupled ocean–atmosphere GCM. *J. Phys. Oceanogr.*, **38**, 588–604, https://doi.org/10.1175/2007JPO3769.1.

Vimont, D. J., M. A. Alexander, and M. Newman, 2014: Optimal growth of central and east Pacific ENSO events. *Geophys. Res. Lett.*, **41**, 4027–4034, https://doi.org/10.1002/2014GL059997.

Vitart, F., 2017: Madden—Julian oscillation prediction and teleconnections in the S2S database. *Quart. J. Roy. Meteor. Soc.*, **143**, 2210–2220, https://doi.org/10.1002/qj.3079.

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

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.

Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932, https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

Whitaker, J. S., and P. D. Sardeshmukh, 1998: A linear theory of extratropical synoptic eddy statistics. *J. Atmos. Sci.*, **55**, 237–258, https://doi.org/10.1175/1520-0469(1998)055<0237:ALTOES>2.0.CO;2.

Xiang, B., Y. Q. Sun, J.-H. Chen, N. C. Johnson, and X. Jiang, 2020: Subseasonal prediction of land cold extremes in boreal wintertime. *J. Geophys. Res. Atmos.*, **125**, e2020JD032670, https://doi.org/10.1029/2020JD032670.