The Spring Minimum in Subseasonal 2-m Temperature Forecast Skill over North America

MELISSA L. BREEDEN,[a,b] JOHN R. ALBERS,[a,b] AMY H. BUTLER,[c] AND MATTHEW NEWMAN[a,b]

[a] *Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado*
[b] *NOAA/Physical Sciences Laboratory, Boulder, Colorado*
[c] *NOAA/Chemical Sciences Laboratory, Boulder, Colorado*

(Manuscript received 4 March 2022, in final form 14 June 2022)

ABSTRACT: On average, 2-m temperature forecasts over North America for lead times greater than two weeks have generally low skill in operational dynamical models, largely because of the chaotic, unpredictable nature of daily weather. However, for a small subset of forecasts, more slowly evolving climate processes yield some predictable signal that may be anticipated in advance, occasioning "forecasts of opportunity." Forecasts of opportunity evolve seasonally, since they are a function of the seasonally varying jet stream and various remote forcings such as tropical heating. Prior research has demonstrated that for boreal winter, an empirical dynamical modeling technique called a linear inverse model (LIM), whose forecast skill is typically comparable to operational forecast models, can successfully identify forecasts of opportunity both for itself and for other dynamical models. In this study, we use a set of LIMs to examine how subseasonal North American 2-m temperature potential predictability and forecasts of opportunity vary from boreal winter through summer. We show how LIM skill evolves during the three phases of the spring transition of the North Pacific jet—late winter, spring, and early summer—revealing clear differences in each phase and a distinct skill minimum in spring. We identify a subset of forecasts with markedly higher skill in all three phases, despite LIM temperature skill that is somewhat low on average. However, skill improvements are only statistically significant during winter and summer, again reflecting the spring subseasonal skill minimum. The spring skill minimum is consistent with the skill predicted from theory and arises due to a minimum in LIM forecast signal-to-noise ratio.

KEYWORDS: North America; Teleconnections; Surface temperature; Machine learning; Subseasonal variability

1. Introduction

Routinely producing skillful subseasonal (3–8-week lead) 2-m temperature (2mT) and precipitation forecasts remains difficult for the current generation of dynamical subseasonal forecast models in all seasons (Pegion et al. 2019; de Andrade 2019). As a result, much attention has been given to identifying the smaller subset of subseasonal forecasts that are useful, by considering phenomena that impart memory, and therefore predictability, to the system—so-called forecasts of opportunity (Albers and Newman 2019; Mariotti et al. 2020). These more skillful forecasts reflect periods of high signal-to-noise ratio, or when a predictable "signal" overwhelms unpredictable "noise" in the system evolution. Predictable signals in the extratropics can arise through many processes including tropical–extratropical teleconnections (Winkler et al. 2001), stratosphere-troposphere interactions (Baldwin et al. 2003; Butler et al. 2019a; Domeisen et al. 2020; Albers and Newman 2021), and long-lasting soil moisture anomalies (Koster et al. 2011). Which processes are most important for forecasts of opportunity is determined by forecast location, target variable, forecast lead time, and time of year.

A key source of subseasonal predictability in the Pacific–North American region is tropical diabatic heating (e.g., Newman et al. 2003). For example, elevated North American temperature skill follows certain phases of the Madden–Julian oscillation (MJO; Madden and Julian 1971; Johnson et al. 2014; Vigaud et al. 2018) and El Niño–Southern Oscillation (ENSO; Johnson et al. 2014; Wang and Robertson 2018). However, skill provided by the MJO and ENSO is not constant at all times of the year, in part because the nature of tropical convection evolves over the course of the annual cycle. Convection related to ENSO tends to weaken during spring, particularly following its mature winter phase, while the extratropical influence of tropical heating via teleconnections varies due to the annual cycle of the Pacific jet and waveguide (Newman and Sardeshkhmukh 1998). In addition to tropical heating, stratospheric variability can affect North American temperatures during winter and early spring before the final stratospheric warming, with some stratospheric states being associated with elevated subseasonal temperature forecast skill (Gerber et al. 2012; Butler et al. 2019b; Domeisen et al. 2020).

Perhaps as a result of these seasonally varying phenomena, spring is a particularly difficult forecast period. In a recent study, Albers et al. (2021) found that the ability of the European Centre for Medium-Range Weather Forecasts Integrated Forecast System (IFS) to predict variations in the North Pacific jet decreases markedly between early and late spring. This decrease in subseasonal skill could be related to the strong invigoration of the North Pacific storm track that occurs during this time of year (Breeden et al. 2021), representing an increase in

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/MWR-D-22-0062.s1.

Corresponding author: Melissa L. Breeden, melissa.breeden@noaa.gov

DOI: 10.1175/MWR-D-22-0062.1

unpredictable synoptic variability. Another possibility is that models have difficulty modeling the smaller spatial scales of tropical heating that tend to increasingly dominate tropical variability as spring progresses into summer (Newman and Sardeshmukh 1998). It is perhaps not surprising then, that spring temperature predictability has not been extensively studied compared to winter and summer. In this study, we hypothesize that seasonal changes in both the jet and tropical heating correspond to a reduction in subseasonal temperature predictability during spring.

Prior analysis has demonstrated the utility of an empirical-dynamical linear inverse model (LIM; Penland and Sardeshmukh 1995; Sardeshmukh et al. 2000) to generate subseasonal forecasts and objectively identify forecasts of opportunity. During winter, a LIM can produce 500-hPa geopotential height and mean sea level pressure forecasts with skill comparable to the National Centers for Environmental Prediction Climate Forecast System, version 2 (CFSv2), and IFS for lead times of 3–6 weeks, while also identifying periods of elevated skill in its own forecasts and those of the CFSv2 and IFS (Albers and Newman 2019). A similar LIM produced subseasonal North Atlantic Oscillation (NAO) index forecasts with skill comparable to the IFS. In the latter LIM, NAO forecasts of opportunity were found to be due to SST-related heating anomalies and downward propagating stratospheric circulation anomalies (Albers and Newman 2021). LIM forecasts during spring, however, have not been investigated in the present literature, nor has 2mT been selected as a target variable.

In this study, we employ a recently introduced method for tracking the winter-to-summer evolution of the North Pacific jet, which we use to define the spring state of the jet on a flow-dependent, rather than calendar day, basis (Breeden et al. 2021). This approach ensures that the substantial year-to-year variability in the seasonal cycle of the North Pacific jet is accounted for, and anomalies developing in similar mean states are correctly grouped together. Three LIMs, similar to those used in Winkler et al. (2001) and Newman et al. (2003), are subsequently developed to produce late winter, spring and early summer forecasts of North American 2mT. An optimal growth approach to identifying forecasts of opportunity is employed, which successfully identifies periods of elevated circulation and 2mT skill. Finally, we compare observed skill to that expected from theory and compare the evolution of predictable signal and unpredictable noise during the three jet phases to better understand the observed skill evolution.

2. Data and methods

This study uses gridded reanalysis data to construct separate LIMs for the winter, spring and summer phases of the spring transition.

a. Data

Daily mean 200-hPa zonal wind, 200-hPa (Ψ_{200}) and 850-hPa (Ψ_{850}) streamfunction, 2-m temperature (2mT), and outgoing longwave radiation (OLR) were accessed from the Japanese Meteorological Agency 55-year Reanalysis dataset (JRA-55; Kobayashi et al. 2015) for the years 1959–2018 and months of January–July. All variables are regridded to 2.5° × 2.5° horizontal resolution. Northern Hemisphere Ψ_{200} and Ψ_{850}, hemispheric OLR from 20°S to 20°N, and North American 2mT over land only, are used to create the LIMs described in section 2c. For tracking the spring transition of the North Pacific jet (described in section 2b), we use 200-hPa zonal wind values for January–July, with only the 60-yr January–July mean removed so that the seasonal cycle is retained, following the method of Breeden et al. (2021). Conversely, anomalies used in the LIM (section 2c) were calculated by removing the 60-yr daily climatology, and then applying a 7-day running mean to isolate weekly variability and minimize high-frequency, synoptic variations. Such averaging is consistent with the assumptions made by the modeling framework described below (Newman et al. 2003; Albers and Newman 2019). While it is possible a long-term trend related to climate change is present in the 2mT anomalies used, Wulff et al. (2022) recently showed that the fraction of 7-day averaged 2mT variance associated with such a trend was small over North America for the period 1998–2017. As such, we do not attempt to remove any such trend here, though we acknowledge that forecast skill can arise from many processes on a variety of time scales.

b. Defining the spring transition

Following the approach in Breeden et al. (2021), we use the leading empirical orthogonal function (EOF1) and corresponding principal component (PC1) of 200-hPa zonal wind over the North Pacific domain, which is 100°–280°E, 10°–70°N (Fig. 1) to track the seasonality of the North Pacific jet. PC1 is positive during late winter and early spring, decreases throughout the spring transition, and becomes negative sometime in May or June. The substantial spread in PC1 early in the spring transition reflects the high variability in the jet and storm track at this time of year, and the limitation of using a single calendar date to define "spring." Therefore, we instead use each year's transition date when PC1 first falls below $+0.6\sigma$ to define the start of the spring phase. Similarly, when PC1 first falls below -0.6σ is used to define the beginning of the summer phase. The $\pm0.6\sigma$ threshold was chosen to have a roughly even number of samples in the three phases (N = 4986 days in winter, 4055 days in spring and 4103 days in summer). The threshold 0.6σ is used, instead of 0.5σ as in Breeden et al. (2021), to make the number of samples more even between the three phases (in this study we include the months of January and July to the winter and summer phases, respectively, increasing those phases' number of dates relative to spring). This difference in threshold does not affect the composite spring jet structure or LIM skill, although including July in the summer phase does weaken the jet magnitude in summer (cf. Figs. 1 and 2 of Breeden et al. 2021).

c. Linear inverse model

The LIM is based on the assumption that the dynamical evolution of the variables in the state vector **x** [Eq. (1)] can be

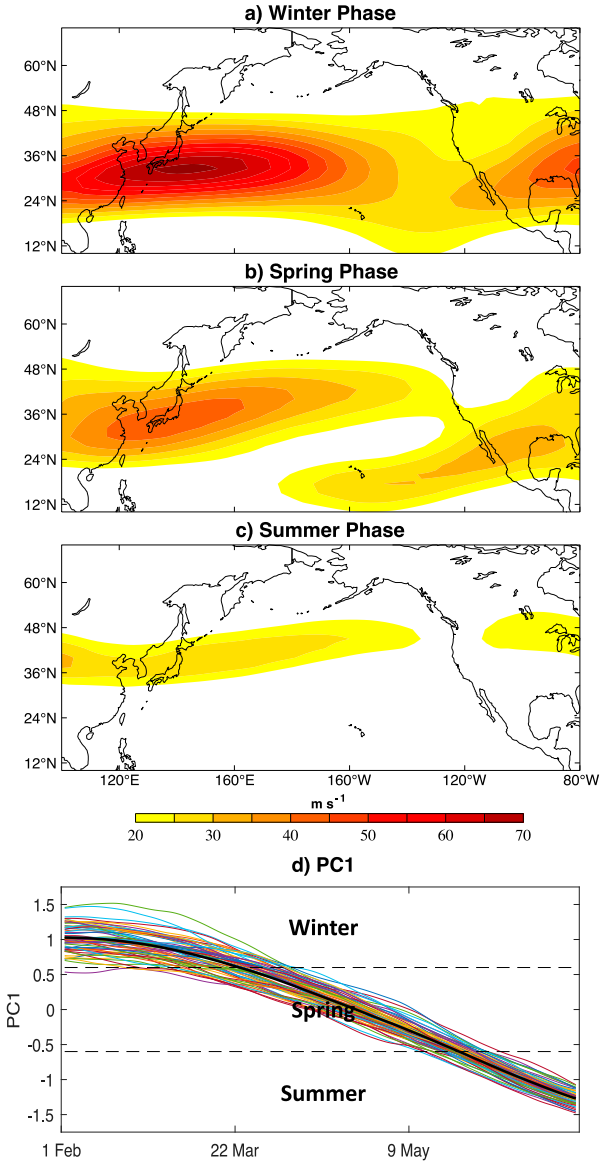

FIG. 1. (a) The composite 200-hPa zonal wind (m s^{-1}) based on periods when (a) PC1 $> 0.6\sigma$, the "winter phase," (b) PCI between $\pm 0.6\sigma$, the "spring phase," and (c) PCI $< -0.6\sigma$, the "summer phase." (d) The thin lines show PC1 for each year in the 60-yr JRA55 record, starting from 1 Feb through 27 Jun, and the thick black line is the 60-yr mean.

reasonably approximated as a linear system forced by stochastic white noise [Eq. (2)]:

$$\mathbf{x} = (\mathrm{OLR}, \Psi_{200}, \Psi_{850}, 2\mathrm{mT}), \tag{1}$$

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}\mathbf{x} + \mathbf{F}_s. \tag{2}$$

Equation (2) approximates the evolution of $\mathbf{x}$ by assuming that a time scale separation exists between the predictable, slowly evolving dynamics represented by $\mathbf{L}$ and the fast, rapidly decorrelating and therefore unpredictable variations represented by $\mathbf{F}_s$. Here, as in previous studies (Winkler et al. 2001; Newman et al. 2003; Albers and Newman 2019; Henderson et al. 2020; Breeden et al. 2020), slowly evolving refers to weekly varying anomalies (consistent with the 7-day running mean applied to the anomalies in $\mathbf{x}$), and fast refers to shorter synoptic and mesoscale variations. The matrix $\mathbf{L}$ can be determined from the time-averaged $\mathbf{C}_0$ and lagged covariance $\mathbf{C}_{\tau_o}$ statistics of the state vector [Eq. (3)]:

$$\mathbf{L} = \log[\mathbf{C}_0 * \mathrm{inv}(\mathbf{C}_0)]/\tau_o. \tag{3}$$

To reduce dimensionality, $\mathbf{x}$ is constructed from the principal components (PCs) resulting from empirical orthogonal function (EOF) analysis of each variable in the state vector, where enough PCs are retained to capture a majority of variance of each variable, while retaining a stable model (see Table S1 in the online supplemental material). A range of retained PCs was tested, and results are not sensitive to the exact number of PCs retained in $\mathbf{x}$ (not shown). Here a training lag τ_o of 5 days is used, consistent with prior studies that have shown this training lag falls within the range to which $\mathbf{L}$ is insensitive, within the constraints of sampling [see Winkler et al. (2001) for a detailed discussion; also Breeden et al.'s (2020) appendix A]. Note that the matrix $\mathbf{L}$, often referred to as the "dynamic operator," acts linearly on $\mathbf{x}$ but can include both linear relationships and linear approximations to nonlinearities, which may be included in $\mathbf{C}_0$ and $\mathbf{C}_{\tau_o}$. This is in contrast to, for instance, a model based upon linearized equations of motion.

1) LIM HINDCASTS

Given initial conditions $\mathbf{x}(0)$, a LIM forecast for any lead time, τ, can be generated using $\mathbf{L}$ by solving the homogeneous component of Eq. (2):

$$\hat{\mathbf{x}}(\tau) = \mathbf{x}(0)\exp(\mathbf{L}\tau) = \mathbf{x}(0)\mathbf{G}(\tau). \tag{4}$$

Cross-validated hindcasts are created by removing 10% of the data, recomputing $\mathbf{L}$ and creating forecasts using the portion of removed data as initial conditions (ICs). To evaluate the LIM hindcast skill, hindcasts are verified against the untruncated reanalysis anomalies using the anomaly correlation coefficient (ACC) at each grid point (e.g., Newman et al. 2003).

To determine whether the observed skill evolution during the three jet phases is consistent with the skill evolution predicted from the LIM's theoretical signal-to-noise ratio S [Eqs. (5)–(7)], the value "theoretical expected skill" $\rho_\infty(\tau)$ [Eq. (8); Sardeshmukh et al. 2000; Newman et al. 2003] is calculated and compared to actual forecast skill. The ACC version of S is a function of forecast lead time τ and is calculated at each grid point i, using the diagonal element of $\mathbf{F}(\tau)$ and $\mathbf{E}(\tau)$, ii (e.g., Newman et al. 2003):

$$\mathbf{F}(\tau) = \langle \hat{\mathbf{x}}(t+\tau)\hat{\mathbf{x}}(t+\tau)' \rangle, \tag{5}$$

$$\mathbf{E}(\tau) = \mathbf{C}_0 - \mathbf{G}(\tau)\mathbf{C}_0\mathbf{G}(\tau)', \tag{6}$$

$$S^2(\tau, i) = \frac{\mathbf{F}(\tau)_{ii}}{\mathbf{E}(\tau)_{ii}}, \tag{7}$$

where a prime indicates the transpose, $\mathbf{F}(\tau)$ is the forecast signal covariance matrix determined at a given lead time, indicating the strength of the predictable signal in the forecasts, and $\mathbf{E}(\tau)$ is the forecast error covariance matrix and represents lead-dependent, unpredictable "noise." Following Sardeshmukh et al. 2000, S^2 can be, in turn, used to calculate expected skill of a perfect infinite member ensemble forecast:

$$\rho_\infty(\tau, i) = \frac{S^2(\tau, i)}{\{[S^2(\tau, i) + 1]S^2(\tau, i)\}^{0.5}}. \quad (8)$$

To understand how the seasonal evolution in the signal and the noise—which are found to differ—act together to produce the theoretical expected skill evolution, the quantities $\mathbf{F}(\tau)_{ii}$, $\mathbf{E}(\tau)_{ii}$, and $\rho_\infty(\tau, i)$ will be compared during the three jet phases for $\tau = 21$ days for 2mT.

2) Forecasts of opportunity

Our approach to identifying forecasts of opportunity will focus on the signal component of signal-to-noise by anticipating periods of rapid, potentially predictable 2mT growth. The LIM is based on the concept that subseasonal anomaly growth can be modeled through the constructive interference of evolving non-orthogonal modes in the system that all have distinct spatial and temporal characteristics (Farrell 1988; Lacarra and Talagrand 1988; Farrell and Ioannou 1996). If the eigenmodes of $\mathbf{L}$ are non-orthogonal, then transient anomaly growth may occur via the constructive interference of the eigenmodes over a finite period of time [because (2) is asymptotically stable, the eigenvalues of $\mathbf{L}$ are all negative and the individual eigenmodes all decay in finite time]. Physically, the non-orthogonality of $\mathbf{L}$ can arise from the presence of asymmetric interactions in the system and patterns of variability that involve multiple physical processes evolving on different time scales (e.g., the NAO; Albers and Newman 2021). Particularly relevant here are asymmetries introduced by shear and zonal asymmetry in the mean state (Farrell 1982; Boyd 1983), as observed in the North Pacific jet exit region (Mak and Cai 1989; Breeden and Martin 2018). The most rapidly growing patterns that amplify via transient growth can be determined using the eigendecomposition of system growth [Eq. (9); Penland and Sardeshmukh (1995); Newman et al. (2003)]. System growth $\mu(\tau)$ can be constrained to occur under a particular initial or final state, which can be set using the initial and final "norm" kernels $\mathbf{D}$ and $\mathbf{N}$, respectively:

$$\mu(\tau) = \frac{\mathbf{x}(\tau)^{\mathrm{T}}\mathbf{N}\mathbf{x}(\tau)}{\mathbf{x}(0)\mathbf{D}\mathbf{x}(0)} = \frac{\mathbf{x}(0)\mathbf{G}(\tau)^{\mathrm{T}}\mathbf{N}\mathbf{G}(\tau)\mathbf{x}(0)}{\mathbf{x}(0)^{\mathrm{T}}\mathbf{D}\mathbf{x}(0)}, \quad (9)$$

$$\mathbf{G}(\tau)^{\mathrm{T}}\mathbf{N}\mathbf{G}(\tau) = \mu(\tau)\mathbf{v}(\tau). \quad (10)$$

The corresponding eigenvalues $\mu(\tau)$ determined from Eq. (10) represent the system growth rate associated with the evolution of the corresponding initial patterns contained in the eigenvectors $\mathbf{v}(\tau)$. The eigenmodes can be sorted from highest to lowest growth rate using the eigenvalues, and those with the strongest growth might be expected to be the most predictable. Growth is maximized for a prescribed period of τ days and without any constraint on the initial pattern, for which $\mathbf{D}$ is set to the identity matrix, as done in past studies (e.g., Sardeshmukh et al. 1997; Newman et al. 2003). Commonly used norms include the L2 or "energy" norm, in which $\mathbf{N}$ is also set to the identity matrix. For this study, our final norm $\mathbf{N}$ maximizes North American 2mT anomaly growth, which is specified by beginning with the identity matrix and setting the diagonals corresponding to the OLR, Ψ_{200}, Ψ_{850} PCs to zero, so that only 2mT amplitude is constrained to amplify. The resultant "optimal patterns" (OPs) we describe in section 3 are interpreted as the first and second patterns most conducive to 2mT growth, and will be referred to as OP1 and OP2, respectively. For each phase of the spring transition, we will examine the OP1 and OP2 maximizing 2mT growth over a 21-day period, which is the midpoint of the week 3–4 forecast period.

We identify forecasts of opportunity using the *optimal initial conditions,* determined from the eigenvectors of Eq. (10), which maximize 2mT growth of either OP1 or OP2. Specifically, we assume that forecasts initialized when the initial atmospheric state strongly resembles one of the optimal initial conditions associated with either OP1 and OP2 will be followed by periods of elevated forecast skill. The hypothesized correspondence between strong projections onto the optimal initial conditions, and higher ACC, will be tested for the three phases of the spring transition using OP1 and OP2 for each phase, and the corresponding optimal initial conditions OP1-IC and OP2-IC. To identify forecasts of opportunity, we use the OP-ICs corresponding to a τ at the midpoint of the forecast lead time, for example, for week 3–4 forecasts (days 15–28), we determine OP-IC using a lead time $\tau = 21$ days. Note that, while this lead time varies according to the forecast lead time of interest (weeks 3–4 or 5–6), the OPs themselves do not change markedly across these lead times, meaning the OPs that maximize 2mT growth at 21 and 35 days are not substantially different. For this reason, we show only the 21-day patterns Figs. 2–4).

The 20% of dates with the strongest projection—positive or negative—onto a given OP-IC is selected as forecasts of opportunity, corresponding to 997 dates in winter, 811 in spring, and 820 in summer. The statistical significance of skill changes for each subset of forecasts is compared against the skill of the remaining 80% of forecasts. Significance at the 95% confidence level is determined non-parametrically using bootstrapping, where the two groups of forecasts are resampled at the smaller sample size of the two groups, after it is reduced by a factor of 5 to account for autocorrelation in the 2mT field, which reduces the number of truly independent samples. Five is chosen because for 2mT, 5 days is the lag at which autocorrelation drops below 0.5 over North America. ACC is then recomputed for that subsample, and the process repeated 5000 times to establish confidence intervals for ACC at each grid point. Finally, we note that the forecasts that are all initialized during a particular jet phase are evaluated together, since they are generated using the same LIM, even if the forecast dates verify during a different phase.

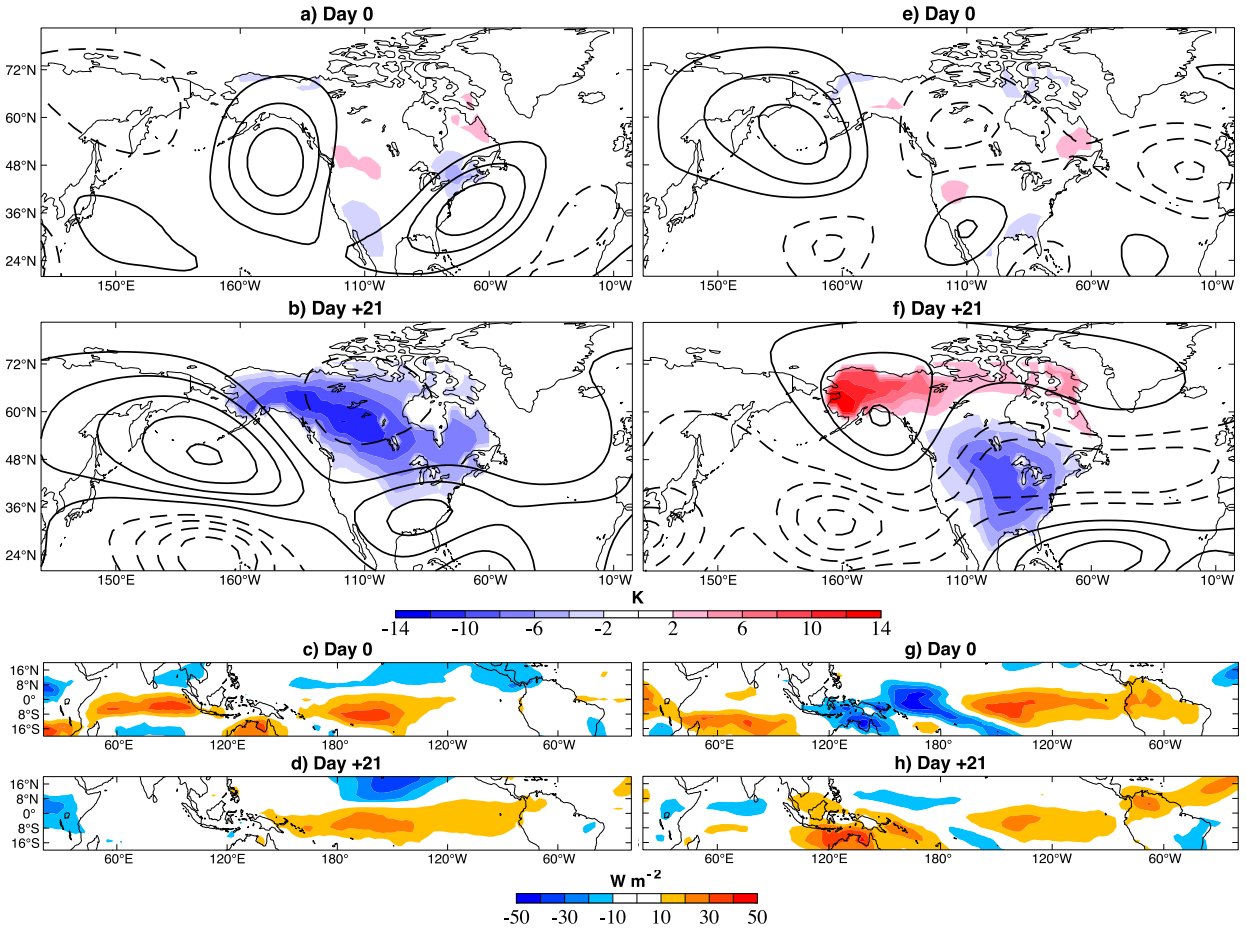

FIG. 2. (a)–(d) First and (e)–(h) second optimal patterns (OP1, OP2) maximizing North American 2-m temperature growth, over a 21-day period, during the winter jet phase. The color shading in (a), (b), (e), and (f) is anomalous 2mT (K), and the black contours are 200-hPa Ψ_{200} anomalies (positive in solid, negative in dashed). The contour interval for Ψ_{200} is $\pm 6 \times 10^6$ kg m^2 s^{-1} contoured at intervals of 6×10^6. The color shading in (c), (d), (g), and (h) is anomalous OLR (W m^{-2}).

3. Results

The three phases of the spring transition of the North Pacific jet show its seasonal evolution from the strong, wintertime jet to a weaker, split jet during spring, and finally, a very weak summer jet structure (Fig. 1). The onset of the spring phase coincides with an invigoration and northward shift of the North Pacific storm track, which is followed by relatively quiescent storm track conditions during summer (Breeden et al. 2021). The winter-to-spring transition is highly variable but typically occurs in late March or April, while the spring-to-summer transition is less variable and generally occurs in late May (Fig. 1d). In section 3a, we compare the OP1 and OP2 patterns during each jet phase, revealing the patterns are all notably different. Section 3b examines the corresponding evolution of LIM 2mT hindcast skill, and evaluates the success in using OP1-IC and OP2-IC to identify forecasts of opportunity. To better understand the observed forecast skill evolution, section 3c considers theoretical expected skill and the seasonal evolution of the signal [Eq. (5)] and the noise [Eq. (6)].

a. Optimal 2-m temperature structures

Under the constraint to maximize North American 2mT anomaly growth at a 21-day lead time [Eq. (10)], each phase of the jet is associated with different heating and circulation structures (Figs. 2–4). Note that due of the linearity of the LIM, equal and opposite patterns (e.g., warm anomalies instead of cold anomalies in Fig. 2b) correspond to equal and opposite signs of all variables.

During the winter phase (Fig. 2), OP1-IC involves an upper-level ridge in the east Pacific and negligible 2mT anomalies; 21 days later, widespread cold anomalies develop in conjunction with a blocking anticyclone (positive Ψ_{200} anomaly) upstream over the central Pacific and the associated cold air advection on the ridge's eastern flank. Note that while this pattern maximizes 2mT growth at a 21-day lead time, the 2mT anomalies are strongest in magnitude at day +15, indicating that the inherent time scale of this pattern is shorter than 21 days, but still produces substantial anomalies at day +21. A negative Ψ_{200} anomaly is located above the cold 2mT anomalies, reflecting an equivalent barotropic

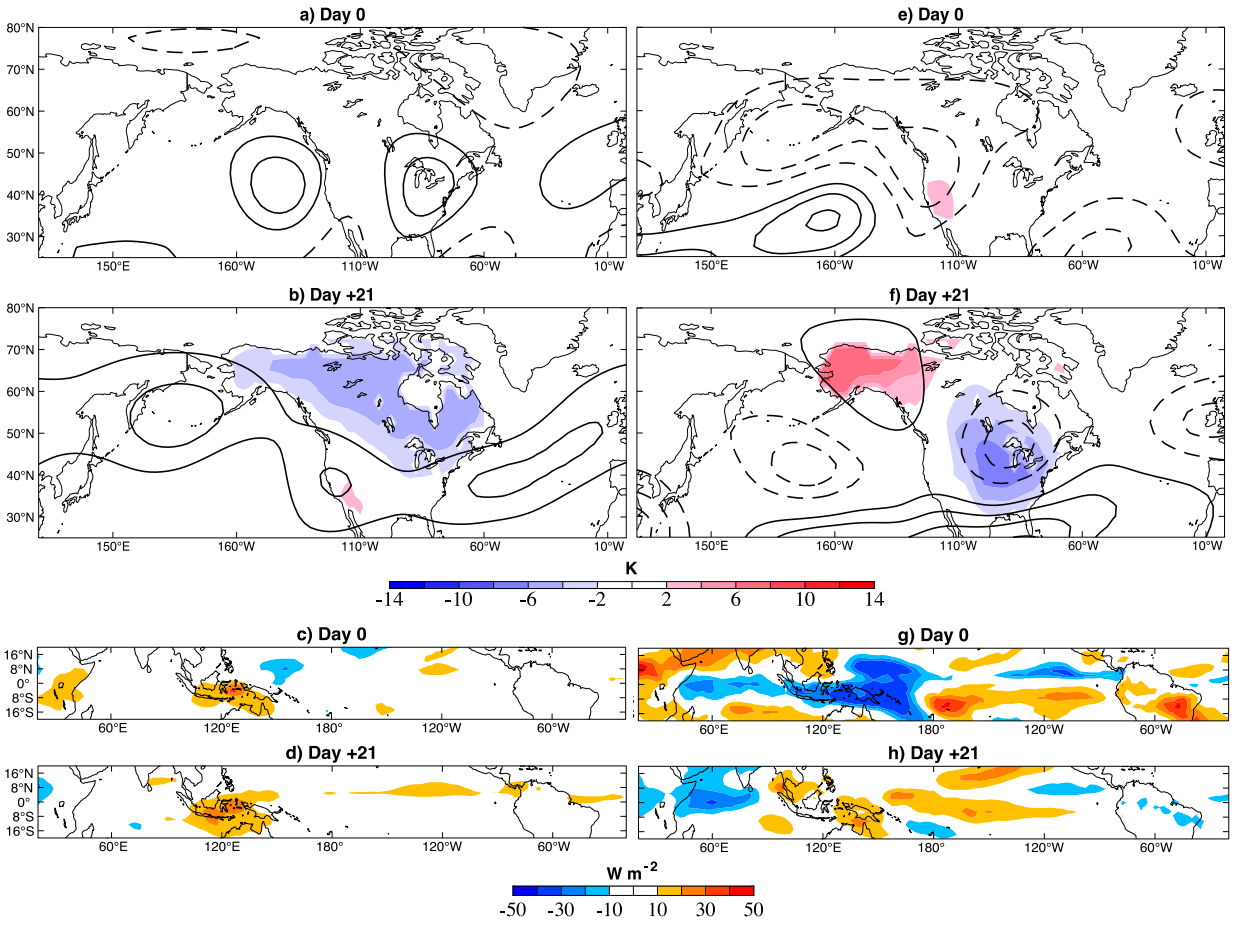

FIG. 3. As in Fig. 2, but for the first and second spring optimal growth structures.

structure, as will be observed over many evolved OPs. Meanwhile, positive OLR anomalies (representing suppressed convection) are located over the eastern Indian Ocean and central Pacific at day 0, with the former propagating eastward with time and the latter remaining stationary (Fig. S1b). Much of the evolution of OP1 resembles that of North Pacific blocking, whose subseasonal evolution can be well produced by a LIM (Breeden et al. 2020). The suppressed convection in the central Pacific is suggestive of La Niña conditions, which are known to increase North Pacific blocking (Renwick and Wallace 1996). OP2 is associated with a couplet of temperature anomalies with centers over Alaska and the central United States, with warm Alaskan temperatures located beneath an upper-level, high-latitude anticyclone and cold temperatures located beneath a broad region of negative Ψ_{200} anomalies (Fig. 2f). In contrast to OP1, OLR OP2-IC involves anomalous convection in the central Pacific, which decays by day +20, when it is replaced by positive OLR anomalies signifying suppressed convection (Figs. 2g,h; Fig. S1e). Both OP1 and OP2 during winter involve the coincidence of anomalously cold temperature developing over most of North America (excluding Alaska for OP2) with a stationary region of suppressed convection (positive OLR anomalies) over the central/eastern Pacific (Figs. S1b, S1e).

During the spring jet phase (Fig. 3), the 2mT and Ψ_{200} 21-day evolved anomalies resemble the OP structures in winter (cf. Figs. 2b,3b and 2f,3f), but with overall weaker and northward-shifted temperature anomalies—coincident with the northward shift of the jet—and the development of weak warm anomalies over the southwestern United States. Interestingly, the OLR optimal initial conditions differ substantially, with the disappearance of stationary OLR anomalies in the central Pacific from winter to spring (Figs. 3c,d; Figs. S2a–c). This lack of persistent, ENSO-like convection in spring is consistent with the frequently observed decay of ENSO events at this time of year, meaning the associated teleconnection likely weakens at this as well. Instead, tropical OLR during spring OP1 involves a small-scale, stationary region of positive OLR anomalies (Fig. S2b) reflecting changes in the optimal heating conditions for 2mT growth from winter to spring with, it appears, significant changes to the predictable 21-day 2mT patterns. OP2 during spring also resembles OP2 in winter, although again with differences in the amplitude and precise location of the temperature and Ψ_{200} anomalies.

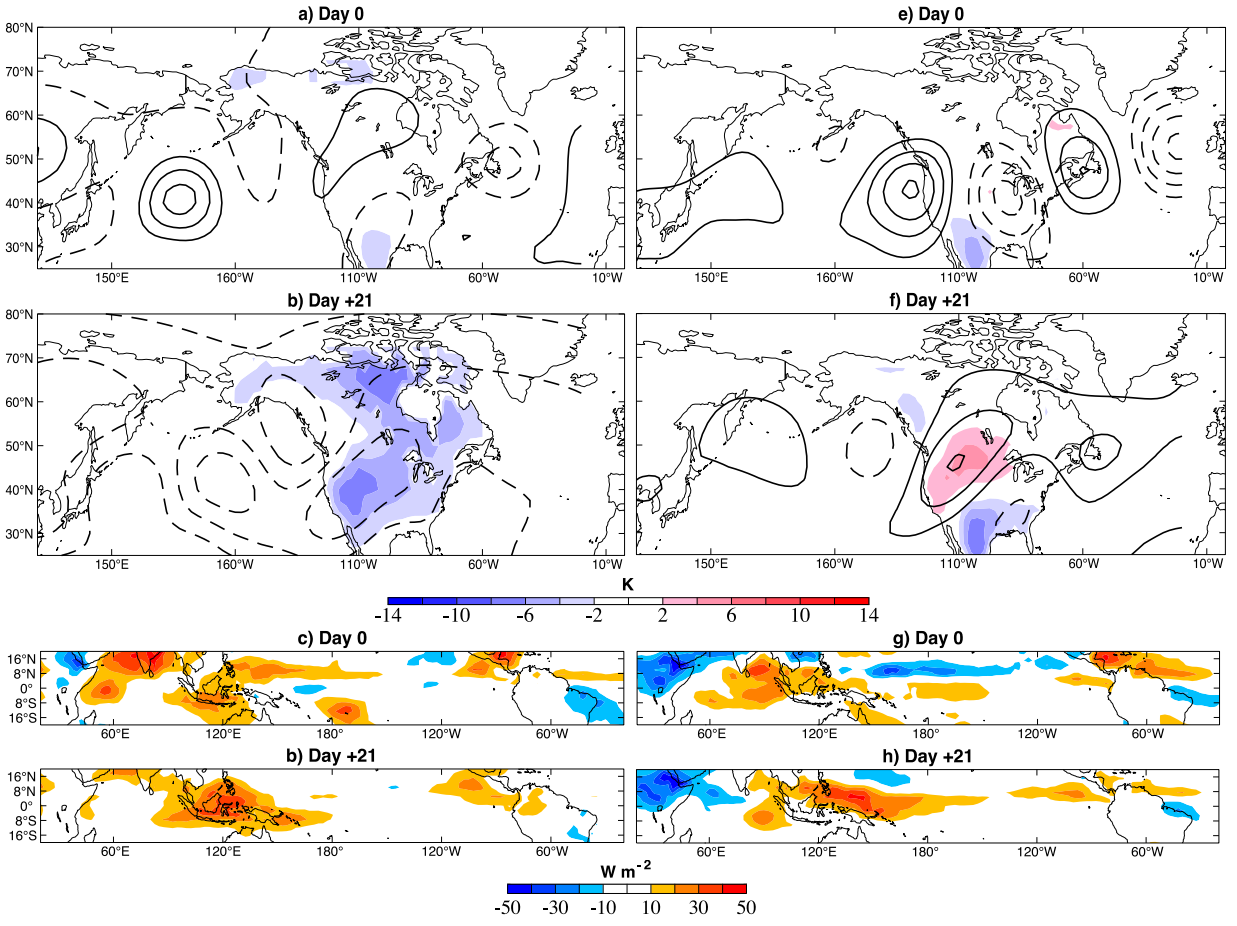

FIG. 4. As in Fig. 2, but for the first and second summer optimal growth structures.

The OPs that evolve in summer differ substantially from those in winter and spring (Fig. 4). In summer, OP1 involves cyclonic Ψ_{200} anomalies over the central Pacific, North America, and North Atlantic that develop along with cold temperature anomalies over central North America (Figs. 4a,b), in contrast to the mid and high-latitude ridging observed in the North Pacific during winter and spring. Positive OLR anomalies form over the Maritime Continent and are mainly stationary from initialization to past 40 days, peaking in strength from days +15 to +35 (Figs. 4c,b; Figs. S3a–c). While there is some consistency between cold temperatures and suppressed convection in the Indian Ocean and Maritime Continent between all three phases' OP1 patterns, the Ψ_{200} response differs in summer with generally cyclonic anomalies observed, suggesting the nature of the summer teleconnection differs from winter and spring, as found in Newman et al. (2003) for winter and summer. Finally, OP2 in summer involves development of a small-scale, zonally oriented wave train over the central Pacific and North America, while tropical OLR, similar to OP1, involves a persistent positive anomaly around 260°–300°W (Figs. 4e–h). Summer OP2 is one of the least persistent Ψ_{200} patterns and displays smaller-scale waves, with the strongest 2mT anomalies also developing at a shorter lead time, around day +10 (Fig. S3f). Overall, from winter to summer, the OP-ICs progressively shrink in zonal wavelength, consistent with the monthly evolution of North American height sensitivities to remote tropical heating found by Newman and Sardeshmukh (1998).

b. 2-m temperature skill evolution

Subseasonal temperature skill evolves during the three phases of the spring transition, but in all phases, periods of elevated forecast skill are identifiable ahead of time using OP1-IC and OP2-IC. On average, week 3–6 2mT skill from LIM forecasts is low (Figs. 5, 6), but similar to dynamical forecast models (Pegion et al. 2019; Wang and Robertson 2018). We find that temperature skill is greatest during winter, reaching a minimum in spring before increasing again in summer (Figs. 5a–c).

Nonetheless, even during the spring minimum, some skill is present in western North America from Alaska to Mexico (Fig. 5b). More encouraging, however, is that in all three phases, skill increases markedly for the forecasts initialized during the 20% of dates with the strongest projection—positive or negative—onto OP1-IC (Figs. 5d–f). For certain regions, skill is significantly different from the skill of the remaining 80% of forecasts during winter and summer, but not spring. Similarly,

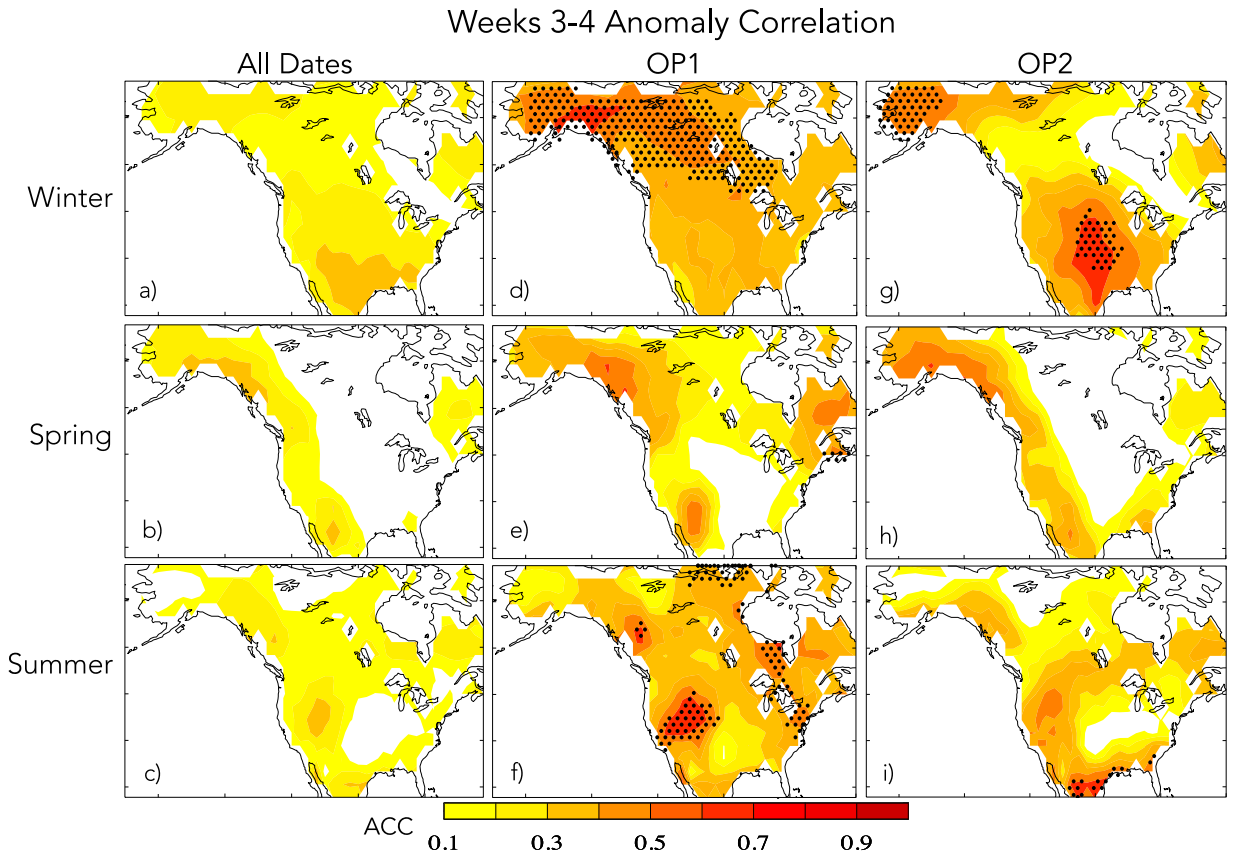

FIG. 5. Anomaly correlation coefficient (ACC) for weeks 3–4 temperature forecasts. ACC using all forecasts during (a) winter, (b) spring, and (c) summer. (d)–(f) ACC for the 20% of forecasts with the strongest projection onto the optimal initial conditions associated with OP1, and (g)–(i) ACC for the 20% of forecasts with the strongest projection onto the optimal initial conditions associated with OP2. Stippling in (d)–(i) indicates where skill changes are significant at the 95% confidence level. The mean ACC values in (a)–(c) are 0.21, 0.14, and 0.15, respectively.

skill increases following strong projections onto OP2-IC, albeit over a smaller area during spring compared to winter and summer (Figs. 5g–i). The regions of highest skill generally coincide with the areas of strong OP growth, further supporting the idea that predictable anomaly growth is associated with the enhanced skill (cf. Figs. 2b,5d; Fig. 4b, Fig. 5g). However, in spring, even for OP1 the spatial extent of ACC is confined to western North America, suggesting something is preventing strong optimal 2mT growth at this time.

For weeks 5–6 forecasts, skill is lower than weeks 3–4 (Figs. 6a–c), yet there are still regions with reasonably high skill (correlations between 0.4 and 0.6), particularly during summer for OP1 (Fig. 6f). Furthermore, spring skill is similar at weeks 5–6 and is statistically significantly different compared to the remaining forecasts for both the OP1 and 2 subsets in some areas, though these are still spatially confined to coastal areas where it is possible high-frequency fluctuations are weaker due to the nearby water, enhancing the impact of the predictable signal (Figs. 6e,h).

It thus appears that the predictable temperature patterns, and their optimal initial conditions, are indeed associated with an increase in the predictable "signal" in subseasonal forecasts, leading to an increase in skill. There are notable differences in when and where skill maximizes (Figs. 5, 6) during each of the three jet phases, but western North America is broadly the region with the most consistent skill. The spring phase displays the lowest overall skill and the weakest skill increase associated with forecasts of opportunity using OP1 and OP2, consistent with prior studies that have found a reduction in subseasonal forecast skill over the Pacific–North American region in spring (Wang and Robertson 2018; Albers et al. 2021).

c. *Theoretical expected skill*

Does the spring skill minimum reflect a change in predictable signal, unpredictable noise, or some combination of both? To address this question, we consider the theoretical expected skill, ρ_∞ [Eq. (8)], for the three jet phases using the 21-day forecast signal-to-noise ratio (Fig. 7).

The ρ_∞ evolution and spatial characteristics are quite similar to the observed ACC (Figs. 5a–c), though ρ_∞ is higher in all phases, as actual skill may be lower due to model imperfections or initial condition errors (Newman et al. 2003). During

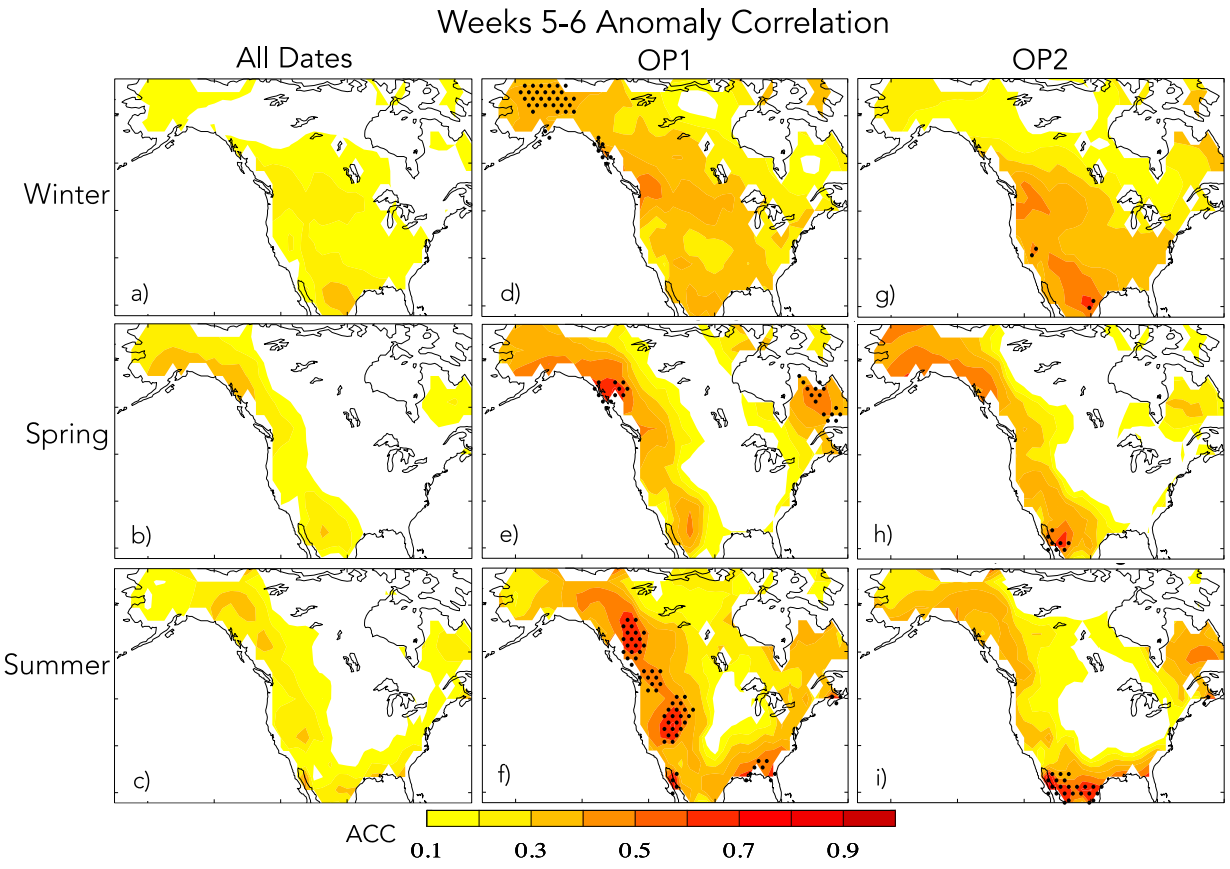

FIG. 6. As in Fig. 5, but for weeks 5–6 forecasts. Mean values for (a)–(c) are 0.13, 0.14 and 0.15, respectively.

the winter phase, both the signal and noise components have the greatest amplitude, with the signal extending farther southwestward and accounting for the skill maximum located at more southern latitudes (Figs. 7a,d,g). In spring, there is a minimum in ρ_∞ over most of North America, except in the far northern regions including Alaska, where the observed spring skill is also highest (cf. Fig. 7b, Fig. 5b). The signal component reaches its minimum amplitude at this time, while noise has diminished as well but to a lesser extent, overall reducing ρ_∞ in spring compared to winter and summer.

The relatively elevated noise in spring in the interior of the continent could account for the lack of ACC increase during forecasts of opportunity associated with OP1 and OP2 (Figs. 5e,h, 6e,h,). In the summer phase, signal increases in the central United States compared to spring, while noise continues to weaken, driving up the signal-to-noise ratio (S^2) and ρ_∞ to peak values in the interior continent (Fig. 7c). As such, it appears the observed skill minimum in spring is consistent with theoretical expected skill and the evolution of predictable signal and unpredictable noise components of the forecasts. It is notable that the signal and noise reach their minimum amplitudes at different times of the year, though the underlying mechanism for their different temporal evolutions is not well understood at present.

4. Discussion and conclusions

In this study, we document predictable North American 2mT patterns, their evolution during three phases of the seasonal cycle, and test the hypothesis that they can be used to anticipate subseasonal forecasts of opportunity. We find that notably different patterns of upper-level circulation and tropical OLR maximize 2mT growth during the winter, spring and summer phases of the North Pacific jet. Some patterns are familiar, or contain familiar elements, such as OP1 in winter and the PNA pattern, suggested by the quadrupole pattern of Ψ_{200} anomalies in Fig. 2b. The associated tropical heating evolution maximizing 2mT growth differs in each phase as well, but overall reveals a role for both stationary and propagating heating anomalies in the Indian Ocean, Maritime Continent and central Pacific. The coincidence of strong, large-scale and persistent tropical heating anomalies during winter and summer, and more skillful forecasts of opportunity for 2mT, is consistent with past research linking tropical–extratropical teleconnections to enhanced subseasonal predictability (e.g., Albers and Nemwan 2021). As such, ENSO, the MJO, and perhaps the Indian summer monsoon, may all be influencing subseasonal 2mT amplification, though further analysis is needed to confirm these potential links.

In all three phases, we are able to leverage knowledge of these predictable patterns to anticipate, at the time of forecast

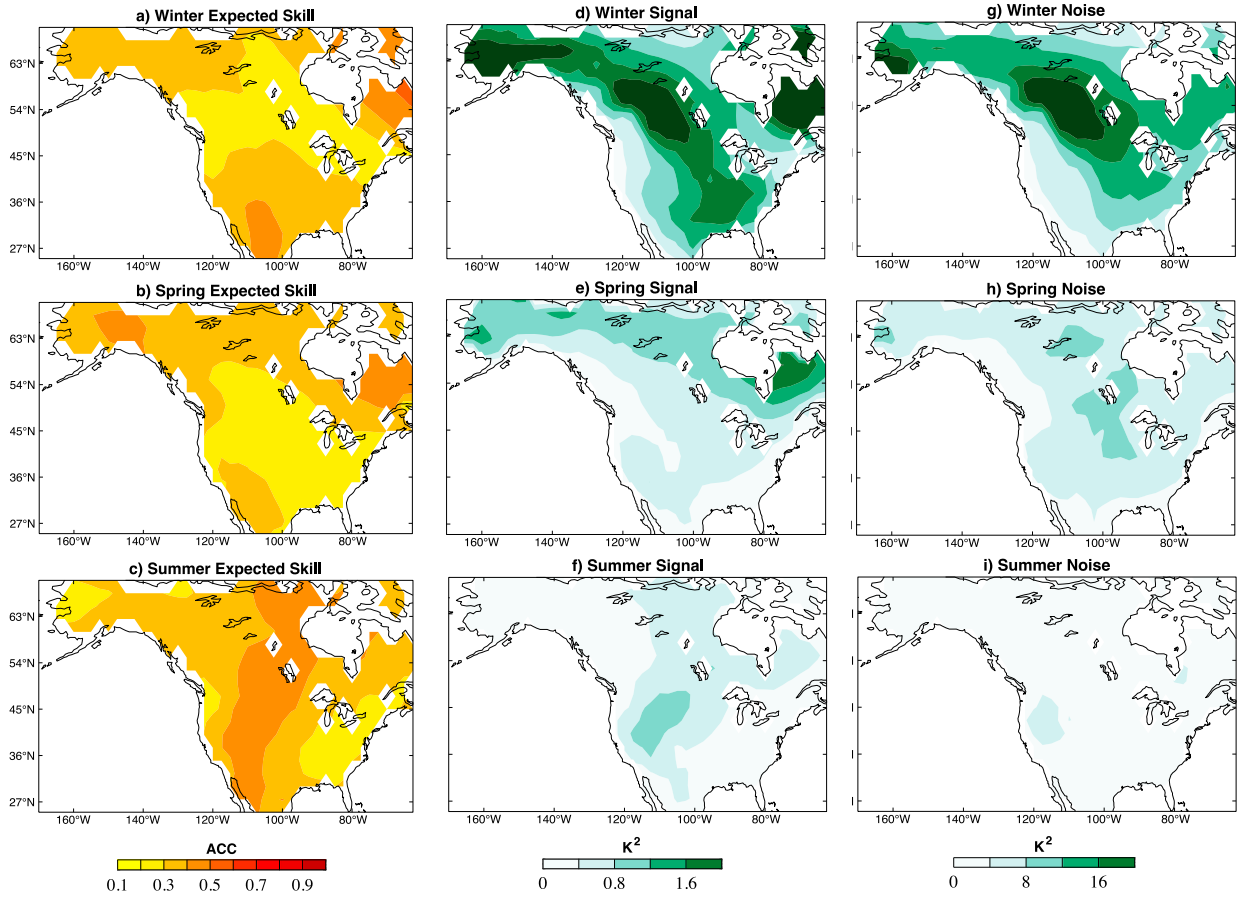

FIG. 7. (a)–(c) The 21-day expected skill [ACC; Eq. (8)] calculated for the three phases of the spring transition. (d)–(f) The 21-day forecast signal covariance component of expected skill [**F**; K^2; Eq. (5)] and (g)–(i) the 21-day error covariance component of expected skill [**E**; K^2; Eq. (6)].

initialization, forecasts of opportunity for 2mT, with the most success in winter and summer. Skill reaches a minimum in spring, and over much of North America, skill remains low, even during forecasts of opportunity, potentially reflecting higher variability in these forecasts that rendered skill changes mostly statistically insignificant. A spring skill minimum is also evident in theoretical expected skill (Fig. 7), indicating a reduced signal-to-noise ratio during spring that obscures predictable anomaly growth and renders subseasonal spring forecasts generally less skillful than those initialized during winter and summer.

Our results are consistent with the minimum in weeks 3–4 2mT skill found over CONUS in March–May (MAM) in the NCEP model by Wang and Robertson (2018). The same study showed the IFS showed a minimum in skill during MAM as well, but only over portions of North America, most notably the central and eastern United States. However, we note that since March includes at least a portion of the winter phase of the jet during most years (Fig. 1), a comparison to MAM skill and the spring skill evaluated in this study may not be the most direct. The spring minimum found in this study is also consistent with the skill reduction in the IFS subseasonal North Pacific jet forecasts during April and May, compared to March, that is shown by Albers et al. (2021), suggesting the spring minimum in skill is not limited to the LIM framework, but likely reflects a true limitation to subseasonal predictability in the PNA region.

Predictable winter and summer 2mT patterns are associated with stationary and persistent tropical convective heating anomalies, though they differ in heating location and the associated circulation and 2mT structures. The OLR evolution during spring OPs was found to be comparatively weak. The strong, persistent heating sources observed in winter and summer could be a key element to these phases' enhanced subseasonal 2mT skill, particularly during forecasts of opportunity (Mayer and Barnes 2021). Furthermore, the wintertime stratospheric polar vortex could contribute to the elevated winter signal observed in Fig. 7, given its known influence on the circulation (Albers and Newman 2021) and surface temperature (Butler et al. 2019a,b; Domeisen et al. 2020). To dynamically isolate these remote influences from tropical heating and the stratosphere, future work could employ the decoupling approach of Albers and Newman (2021) to reexamine subseasonal 2mT forecast skill and forecasts of opportunity. The same

decoupling approach could also be used to examine if any modes exhibit a long-term trend associated with climate change, and how that mode, if identified, contributes to 2mT skill.

Despite the presence of stationary tropical heating in both phases, the OPs that develop in winter and summer differ markedly, including their spatial structure, with winter OPs producing larger-scale waves than summer OPs. Still, both are associated with periods of elevated forecast skill, suggesting smaller-scale patterns are not necessarily less predictable than large-scale patterns. Summer also benefits from the lowest noise amplitude of the three phases, so while the overall amplitude of the signal in summer is far weaker than in winter, the patterns are still predictable, as supported by the success of the forecasts of opportunity in summer (Figs. 5, 6).

Changes in remote SST and stratospheric forcing are not the only potential underlying reasons for the spring minimum in skill. For example, Breeden et al. (2021) found an invigorated storm track during spring, which could mean that synoptic variability—approximated with the noise forcing term in Eq. (2)—is too high to realize the signal associated with OP1 or OP2, as suggested in Fig. 7. Alternatively, variables crucial for spring prediction could be missing from the state vector, although to have a big impact on model performance they would have to contain relevant information that is not implicitly captured in the original variables. Still, we note that the LIMs developed here are relatively simple, and the influence of additional variables can be easily tested. Finally, the LIMs are constructed using the fluctuation–dissipation relationship, assuming that the covariance of the system is constant with time (Penland and Sardeshmukh 1995). It is possible that, even defining spring on a flow-dependent basis as done in this study, that the mean state and variance are changing too rapidly during the transition season to be modeled effectively by the LIM under this assumption. Future work may consider if the spring minimum is evident in alternative models not based on such assumptions, to better discern what factors most strongly impacts subseasonal skill at this time of year.

Acknowledgments. This research was supported by the NOAA/Climate and Global Change Postdoctoral Fellowship Program, administered by UCAR's Cooperative Programs for the Advancement of Earth System Science (CPAESS) under Award NA18NWS4620043B. Funding for this study was also provided by the University of Colorado Boulder, Cooperative Institute for Research in Environmental Sciences, Weather and Climate Dynamics Division.

Data availability statement. The reanalysis data used for this project, JRA-55 (Kobayashi et al. 2015), are freely available from the National Center for Atmospheric Research/University Consortium for Atmospheric Research (NCAR/UCAR) Research Data Archive: https://rda.ucar.edu/datasets/ds628.8/index.html#!access.

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.

——, and ——, 2021: Subseasonal predictability of the North Atlantic Oscillation. *Environ. Res. Lett.*, **16**, 044024, https://doi.org/10.1088/1748-9326/abe781.

——, A. H. Butler, M. L. Breeden, A. O. Langford, and G. N. Kiladis, 2021: Subseasonal prediction of springtime Pacific–North American transport using upper-level wind forecasts. *Wea. Climate Dyn.*, **2**, 433–452, https://doi.org/10.5194/wcd-2-433-2021.

Baldwin, M. P., D. B. Stephenson, D. W. Thompson, T. J. Dunkerton, A. J. Charlton, and A. O'Neill, 2003: Stratospheric memory and skill of extended-range weather forecasts. *Science*, **301**, 636–640, https://doi.org/10.1126/science.1087143.

Boyd, J. P., 1983: The continuous spectrum of linear Couette flow with the beta effect. *J. Atmos. Sci.*, **40**, 2304–2308, https://doi.org/10.1175/1520-0469(1983)040<2304:TCSOLC>2.0.CO;2.

Breeden, M. L., and J. E. Martin, 2018: Analysis of the onset of an extreme North Pacific jet retraction using piecewise tendency diagnosis. *Quart. J. Roy. Meteor. Soc.*, **144**, 1895–1913, https://doi.org/10.1002/qj.3388.

——, B. T. Hoover, M. Newman, and D. J. Vimont, 2020: Optimal North Pacific blocking precursors and their deterministic subseasonal evolution during boreal winter. *Mon. Wea. Rev.*, **148**, 739–761, https://doi.org/10.1175/MWR-D-19-0273.1.

——, A. H. Butler, J. R. Albers, M. Sprenger, and A. O. Langford, 2021: The spring transition of the North Pacific jet and its relation to deep stratosphere-to-troposphere mass transport over western North America. *Atmos. Chem. Phys.*, **21**, 2781–2794, https://doi.org/10.5194/acp-21-2781-2021.

Butler, A., and Coauthors, 2019a: Sub-seasonal predictability and the stratosphere. *Sub-Seasonal to Seasonal Prediction*, A. W. Robertson and F. Vitart, Eds., Vol. 129, Elsevier, 223–241, https://doi.org/10.1016/B978-0-12-811714-9.00011-5.

——, A. Charlton-Perez, D. I. V. Domeisen, I. Simpson, and J. Sjoberg, 2019b: Predictability of Northern Hemisphere final stratospheric warmings and their surface impacts. *Geophys. Res. Lett.*, **46**, 10 578–10 588, https://doi.org/10.1029/2019GL083346.

de Andrade, F. M., C. A. S. Coelho, and I. F. A. Cavalcanti, 2019: Global precipitation hindcast quality assessment of the subseasonal to seasonal (S2S) prediction project models. *Climate Dyn.*, **52**, 5451–5475, https://doi.org/10.1007/s00382-018-4457-z.

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.

Farrell, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.*, **39**, 1663–1686, https://doi.org/10.1175/1520-0469(1982)039<1663:TIGODI>2.0.CO;2.

——, 1988: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.*, **45**, 163–172, https://doi.org/10.1175/1520-0469(1988)045<0163:OEONRW>2.0.CO;2.

——, and P. J. Ioannou, 1996: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.*, **53**, 2025–2040, https://doi.org/10.1175/1520-0469(1996)053<2025:GSTPIA>2.0.CO;2.

Gerber, E. P., and Coauthors, 2012: Assessing and understanding the impact of stratospheric dynamics and variability on the Earth system. *Bull. Amer. Meteor. Soc.*, **93**, 845–859, https://doi.org/10.1175/BAMS-D-11-00145.1.

Henderson, S. A., D. J. Vimont, and M. Newman, 2020: The critical role of non-normality in partitioning tropical and

extratropical contributions to PNA growth. *J. Climate*, **33**, 6273–6295, https://doi.org/10.1175/JCLI-D-19-0555.1.

Johnson, N. C., 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.

Kobayashi, S., and Coauthors, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, https://doi.org/10.2151/jmsj.2015-001.

Koster, R. D., and Coauthors, 2011: The second phase of the global land–atmosphere coupling experiment: Soil moisture contributions to subseasonal forecast skill. *J. Hydrometeor.*, **12**, 805–822, https://doi.org/10.1175/2011JHM1365.1.

Lacarra, J.-F., and O. Talagrand, 1988: Short range evolution of small perturbations in a barotropic model. *Tellus*, **40A**, 81–95, https://doi.org/10.1111/j.1600-0870.1988.tb00408.x.

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.

Mak, M., and M. Cai, 1989: Local barotropic instability. *J. Atmos. Sci.*, **46**, 3289–3311, https://doi.org/10.1175/1520-0469(1989)046<3289:LBI>2.0.CO;2.

Mariotti, A., and Coauthors, 2020: Windows of opportunity for skillful forecasts subseasonal to seasonal and beyond. *Bull. Amer. Meteor. Soc.*, **101**, E608–E625, https://doi.org/10.1175/BAMS-D-18-0326.1.

Mayer, K. J., and E. A. Barnes, 2021: Subseasonal forecasts of opportunity identified by an explainable neural network. *Geophys. Res. Lett.*, **48**, e2020GL092092, https://doi.org/10.1029/2020GL092092.

Newman, M., and P. D. Sardeshmukh, 1998: The impact of the annual cycle on the North Pacific/North American response to remote low-frequency forcing. *J. Atmos. Sci.*, **55**, 1336–1353, https://doi.org/10.1175/1520-0469(1998)055<1336:TIOTAC>2.0.CO;2.

——, ——, C. R. Winkler, and J. S. Whitaker, 2003: A study of subseasonal predictability. *Mon. Wea. Rev.*, **131**, 1715–1732, https://doi.org/10.1175//2558.1.

Pegion, K., and Coauthors, 2019: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.*, **100**, 2043–2060, https://doi.org/10.1175/BAMS-D-18-0270.1.

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.

Renwick, J. A., and J. M. Wallace, 1996: Relationships between North Pacific wintertime blocking, El Niño, and the PNA pattern. *Mon. Wea. Rev.*, **124**, 2071–2076, https://doi.org/10.1175/1520-0493(1996)124<2071:RBNPWB>2.0.CO;2.

Sardeshmukh, P. D., M. Newman, and M. D. Borges, 1997: Free barotropic Rossby wave dynamics of the wintertime low-frequency flow. *J. Atmos. Sci.*, **54**, 5–23, https://doi.org/10.1175/1520-0469(1997)054<0005:FBRWDO>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.

Vigaud, N., A. W. Robertson, and M. K. Tippett, 2018: Predictability of recurrent weather regimes over North America during winter from submonthly reforecasts. *Mon. Wea. Rev.*, **146**, 2559–2577, https://doi.org/10.1175/MWR-D-18-0058.1.

Wang, L., and A. W. Robertson, 2018: Week 3-4 predictability over the United States assessed from two operational ensemble prediction systems. *Climate Dyn.*, **52**, 5861–5875, https://doi.org/10.1007/s00382-018-4484-9.

Winkler, C. R., M. Newman, and P. D. Sardeshmukh, 2001: A linear model of wintertime low-frequency variability. Part I: Formulation and forecast skill. *J. Climate*, **14**, 4474–4494, https://doi.org/10.1175/1520-0442(2001)014<4474:ALMOWL>2.0.CO;2.

Wulff, C. O., F. VItart, and D. I. V. Domeisen, 2022: Influence of trends on subseasonal temperature prediction skill. *Quart. J. Roy. Meteor. Soc.*, **148**, 1280–1299, https://doi.org/10.1002/qj.4259.