BAMS Article

The February 2021 Cold Air Outbreak in the United States

A Subseasonal Forecast of Opportunity

John R. Albers, Matthew Newman, Andrew Hoell, Melissa L. Breeden, Yan Wang, and Jiale Lou

ABSTRACT: The sources of predictability for the February 2021 cold air outbreak (CAO) over the central United States, which led to power grid failures and water delivery shortages in Texas, are diagnosed using a machine learning–based prediction model called a linear inverse model (LIM). The flexibility and low computational cost of the LIM allows its forecasts to be used for identifying and assessing the predictability of key physical processes. The LIM may also be run as a climate model for sensitivity and risk analysis for the same reasons. The February 2021 CAO was a subseasonal forecast of opportunity, as the LIM confidently predicted the CAO's onset and duration four weeks in advance, up to two weeks earlier than other initialized numerical forecast models. The LIM shows that the February 2021 CAO was principally caused by unpredictable internal atmospheric variability and predictable La Niña teleconnections, with nominally predictable contributions from the previous month's sudden stratospheric warming and the Madden–Julian oscillation. When run as a climate model, the LIM estimates that the February 2021 CAO was in the top 1% of CAO severity and suggests that similarly extreme CAOs could be expected to occur approximately every 20–30 years.

KEYWORDS: ENSO; Extreme events; Madden-Julian oscillation; Stratosphere-troposphere coupling; Subseasonal variability; Machine learning

https://doi.org/10.1175/BAMS-D-21-0266.1
Corresponding author: John Albers, john.albers@noaa.gov
Supplemental material: https://doi.org/10.1175/BAMS-D-21-0266.2

In final form 22 September 2022

AFFILIATIONS: Albers, Newman, Breeden, Wang, and Lou—Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA/Physical Sciences Laboratory, Boulder, Colorado; **Hoell**—NOAA/Physical Sciences Laboratory, Boulder, Colorado

The February 2021 central U.S. cold air outbreak (CAO) was a record-breaking event of exceptional persistence, intensity, and expanse, which became the costliest U.S. winter storm on record (NCEI 2022). Starting in late January 2021, a potent blocking anticyclone amplified in the eastern North Pacific, while northerly flow on its eastern side ushered cold air southward into the interior of North America. Downstream of the ridge, low pressure aloft developed over the center of the continent, helping maintain frigid temperatures extending from western Alaska to southern Texas for almost three weeks (Figs. 1a,b). During the two weeks of 8–21 February 2021, average temperatures in the region were more than 30°F below normal, with more than 60 daily minimum temperature records broken at various locations

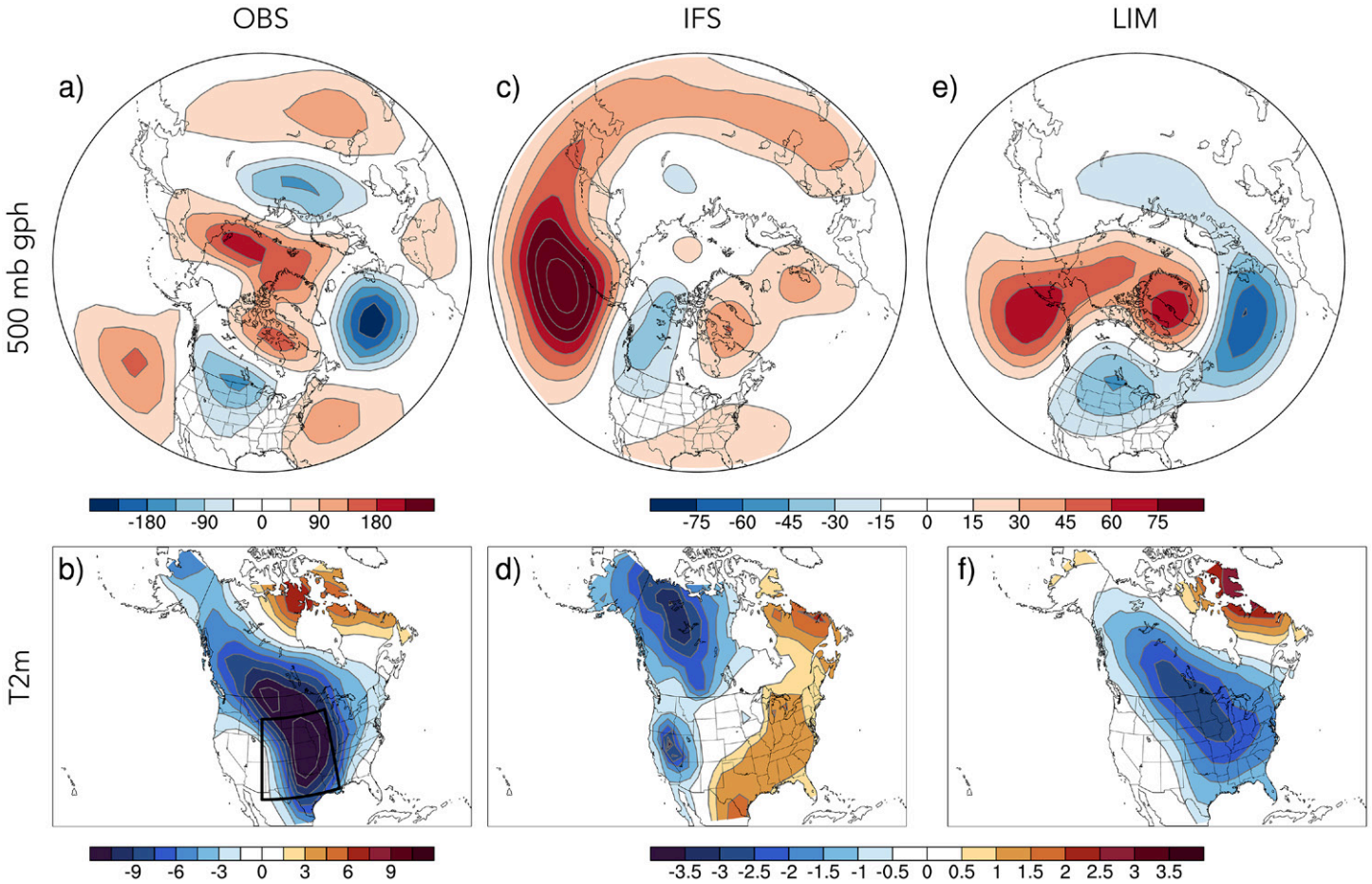

Fig. 1. (a),(b) JRA-55 verifications, (c),(d) IFS forecasts, and (e),(f) LIM reforecasts of (top) 500-mb geopotential height anomalies (gpm) and (bottom) 2-m temperature anomalies (°C). The JRA-55 verification is shown for 8–21 Feb 2021. The LIM reforecast is initialized on 24 Jan 2021 and verifies 8–21 Feb 2021 and is broadly similar to the real-time PSL/CPC LIM forecast initialized on the same date. The IFS forecast (model version CY47R1, operational 2021) is initialized on 25 Jan 2021 and verifies 9–22 Feb 2021. Note that the verification color scale is 3 times larger than the LIM and IFS color scales. The black box overlaying (b) is the region (250°–270°E, 30°–45°N) used for the calculations needed to create Figs. 3–7.

over the central CONUS (NCEI 2021a,b). Over the same period, over 1,000 power generating units failed in Texas alone (FERC 2021), leading to weeks of both power and water shortages for close to 10 million people.

In the aftermath of such a destructive event, it is natural to ask whether impacts resulting from future extreme events might be mitigated if advanced warning capabilities are available. Indeed, this is one of the primary goals of the emerging science of subseasonal-to-seasonal (S2S) forecasting, which seeks to provide forecast outlooks for a few weeks to a few months into the future (e.g., National Academies of Sciences, Engineering, and Medicine 2016a; Robertson and Vitart 2018; White et al. 2017, 2021; Domeisen and Butler 2020). Because this time scale is beyond the 2-week predictability limit of daily weather fluctuations (Lorenz 1969; Buizza and Leutbecher 2015; Zhang et al. 2019), S2S forecast skill is, on average, quite low. However, under some circumstances there may be climate signals whose amplitude and predictability allow for episodes of higher S2S forecast skill (e.g., Albers and Newman 2019 and references therein). This has motivated development of methods to identify "forecasts of opportunity" (Lang et al. 2020; Mariotti et al. 2020; and references therein), those occasions when forecasters could gain confidence in potentially higher S2S forecast skill *at the time of forecast*, such as from the presence of more predictable climate phenomena [e.g., El Niño–Southern Oscillation (ENSO), the Madden–Julian oscillation (MJO), and sudden stratospheric warmings (SSWs) known to drive potentially predictable extreme events].

To aid subseasonal forecasters, an empirical, machine learning–based forecast model called a linear inverse model (LIM; Penland and Sardeshmukh 1995; Newman et al. 2003; Albers and Newman 2019, 2021), developed at NOAA's Physical Sciences Laboratory (PSL), is currently being run experimentally at the NOAA Climate Prediction Center (CPC). This LIM was able to predict, with high confidence, the increased risk of the 2021 CAO 3–4 weeks in advance, which suggests that such extreme events may be similarly predictable in the future. Moreover, the LIM's empirical design allows climate variability to be separated into its component dynamical processes. Together, this suggests that it would be worthwhile to use the LIM to diagnose the 2021 CAO event and to assess which aspects made it predictable on S2S time scales.

In this study, we conduct a novel dynamical process attribution study of the 2021 CAO by using the LIM, rather than a traditional numerical weather model. Like a numerical model, the LIM is a dynamical model that predicts how different atmospheric and oceanic variables will change with time. The LIM reproduces the key ingredients of the 2021 CAO and can determine their relative importance to the event's evolution and predictability, allowing us to quantify two important aspects of extreme CAO risk. First, because of the LIM's dynamical simplicity and computational efficiency, we run several very large ensemble reforecasts incorporating dynamical process-based data denial (i.e., we "turn off" different dynamical processes in different reforecast datasets), to quantify how teleconnections related to La Niña sea surface temperature (SST) anomalies, the MJO, and the downward effects of an SSW each contributed to heightening the risk of a severe CAO during the winter of 2020/21. Second, we ask what the likelihood is of other such CAO events. Using a 3,000-yr LIM climate simulation, we find that while similarly severe CAOs only have about a 4% chance of occurring during a given year, such events are far from rare. Indeed, on average, the most severe events can be expected to occur every 20–30 years, with slightly less severe CAOs occurring at least once per decade.

Forecasts

The CAO period overlapped with three NOAA CPC "week 3–4" temperature outlooks, which were issued on 15, 22, and 29 January to verify for 30 January–12 February, 6–19 February, and 13–26 February, respectively. For the 15 and 22 January outlooks (3–4 weeks prior to CAO onset), many operational S2S forecast models were predicting average to above average

temperatures for most of the CONUS, which was reflected in the official CPC forecast guidance suggesting average to above average temperatures for all of the CONUS except for the Pacific Northwest (www.cpc.ncep.noaa.gov/products/predictions/WK34/archives/). These models include the European Centre for Medium-Range Weather Forecasts Integrated Forecast System (IFS), the National Centers for Environmental Prediction Climate Forecast System version 2 (CFSv2), and the Japanese Meteorological Agency Ensemble Prediction System (JMA). It was not until about two weeks prior to CAO onset (the 29 January forecast) that both the IFS and CFSv2 began to pick up the impending cold surge for the central CONUS.

During this same period, the experimental LIM run at NOAA CPC was forecasting ensemble-mean cold anomalies with 2-week averaged peak amplitudes of −3°C, corresponding to 65%–80% probabilities of below average temperatures for most of the central to eastern CONUS, including as far south as the Texas–Mexico border. These significantly elevated LIM forecast probabilities suggested that the LIM-predicted cold surge was a high confidence forecast of opportunity. (For a more in-depth discussion concerning LIM identification of forecasts of opportunity, see "Identifying S2S forecasts of opportunity in the LIM" sidebar.) Thus, while the numerical forecast models did not predict the CAO until 2 weeks prior to its onset, the CPC LIM was forecasting the potential for a cold surge as far as 4 weeks in advance.

Identifying S2S forecasts of opportunity in the LIM

Climate anomalies are aggregates of daily weather fluctuations. The LIM assumes that the evolution of such aggregates may be approximated by a linearly predictable anomaly evolution, plus an additional "noise" component that represents most of the nonlinear effects that, while crucial to daily weather evolution, are unpredictable beyond daily time scales. This idea, first suggested by Nobel laureate Klaus Hasselmann in his pioneering work *Stochastic Climate Models* (Hasselmann 1976), hinges on the observation that climate is more Gaussian than weather, which allows the evolution of subseasonal-to-seasonal climate to be well-approximated as slowly varying climate anomalies forced by higher frequency daily weather fluctuations.

This simple picture of climate variability and predictability is also relevant for forecasts on S2S time scales, where the prediction of daily weather variations becomes largely or entirely probabilistic (Buizza and Leutbecher 2015). For such predictions, which have substantial uncertainty in the range of their potential outcomes and therefore, on average, generally low deterministic skill, it may be useful to identify when a forecast is expected to be especially skillful, that is, a "forecast of opportunity." On synoptic time scales, forecasts of opportunity typically occur when the forecast ensemble spread is much narrower than usual, sometimes called the "spread–skill relationship" (Whitaker and Loughe 1998). In contrast, on S2S time scales the forecast ensemble spread is wide, approaching the climatological spread, and forecasts of opportunity are largely associated with shifts of the ensemble mean (Pegion and Sardeshmukh 2011; Albers and Newman 2019). As a result, ensemble-mean S2S forecast anomalies, *even for confident and skillful forecasts*, are much smaller in magnitude than observed. For this reason, S2S forecasts are probabilistic and involve categorical predictions, such as binary (above or below normal) or tercile (above, near, or below normal) forecast guidance (e.g., Coelho et al. 2019; Vitart et al. 2019). Relatively small shifts in the ensemble mean can correspond to important shifts in the tails of the forecast ensemble distribution, however, and hence could still substantially increase the risk of extreme events.

Consequently, forecasts of opportunity may be identified using a forecast "signal-to-noise ratio" comparing the ensemble-mean forecast amplitude, which includes *all* potentially predictable climate phenomena (including ENSO, MJO, SSWs, etc.), to the ensemble spread (Sardeshmukh et al. 2000; Newman et al. 2003; Albers and Newman 2019, 2021). In the LIM, ensemble spread is a constant depending only on location and lead time, which represents the expected amplitude of unpredictable weather noise. Albers and Newman (2021), using this LIM approach, showed that periods of above average North Atlantic Oscillation forecast skill for both the LIM and numerical forecast models occur when the amplitude of anomalies related to ENSO and stratospheric Northern Annular Mode events (e.g., SSWs), and the corresponding signal-to-noise ratio, is relatively large. Since periods of higher signal-to-noise ratio also correspond with periods of high LIM forecast probabilities, the high probability of cold temperatures over North America forecasted by the CPC LIM indicated, in real time, that the February 2021 CAO was a forecast of opportunity.

We note that, while the pattern of the LIM ensemble-mean forecast anomalies largely matched observations, forecast amplitude was smaller. This is fairly typical of S2S forecasts (e.g., the IFS in Fig. 1): Even the most "skillful" S2S forecasts have forecast signals (ensemble mean) and forecast uncertainties (ensemble spread) with similar amplitudes. So, we do not generally expect the predicted and observed anomaly amplitudes to match. Instead, potential skill comes from relatively small shifts in the forecast ensemble mean (Albers and Newman 2019; also see "Identifying S2S forecasts of opportunity in the LIM" sidebar). Importantly, these small shifts in the forecast distribution also imply much larger relative changes in the risk of more extreme events (e.g., Coelho et al. 2019; Vitart et al. 2019), as we will see in our analysis below.

The CAO, which peaked in the 2-week period between 8 and 21 February 2021, was accompanied by several notable features in the observed [Japanese 55-year Reanalysis (JRA-55); Kobayashi et al. 2015] anomalous 500-mb (1 mb = 1 hPa) geopotential height field, including a strong high pressure ridge over the North Pacific, a deep trough stretching from the Yukon to Texas, and a negative North Atlantic Oscillation (NAO) pattern (Fig. 1a). These were largely not captured in the IFS S2S forecasts. For example, the IFS weeks 3–4 forecast initialized 25 January 2021 (verifying for 9–22 February 2021, closest to the CAO period) shifted both the Pacific ridge and the North American trough westward and northward, with no NAO pattern (Fig. 1c), which corresponds to the warm temperature anomalies over much of the CONUS predicted by the IFS (Fig. 1d). In contrast, the experimental CPC LIM S2S forecasts better captured the geopotential height pattern and hence the cold air anomalies. For this paper, we developed a modified version of the CPC LIM (see online supplemental material for details; https://doi.org/10.1175/BAMS-D-21-0266.2) that allows for additional diagnosis. Reforecasts using this LIM (Figs. 1e,f) compare well to the original CPC LIM forecasts. While weaker in amplitude and not perfectly in phase, the week 3–4 LIM reforecasted geopotential height anomaly includes all the important dynamical features seen in observations (cf. Figs. 1a and 1e). As a result, the LIM reforecast successfully reproduced below average temperatures for most of the CAO area (Fig. 1f).

Dynamical process attribution

The LIM's ability to anticipate the 2021 CAO prompts two closely related questions: 1) What dynamical processes contributed to the extreme cold temperatures, and 2) what aspects of these dynamical processes were predictable at 3–4-week lead times? Note that in asking the second question we are also asking which dynamical processes led to the elevated LIM forecast probabilities and forecast signal-to-noise ratio ("Identifying S2S forecasts of opportunity in the LIM" sidebar), and hence enabled the CPC LIM to identify the CAO as a forecast of opportunity in mid-January 2021.

Previous studies have suggested multiple phenomena could trigger a North American CAO, including La Niña (Kenyon and Hegerl 2008; Loikith and Broccoli 2014; Yu et al. 2015); downward propagating stratospheric anomalies, including those related to SSWs (Thompson et al. 2002; Kolstad et al. 2010; Cai et al. 2016; Kretschmer et al. 2018; Zhang et al. 2020; Zhang et al. 2021; Huang et al. 2021); the MJO (Johnson et al. 2014); and higher frequency daily weather variability (Cellitti et al. 2006; Westby and Black 2015; Winters et al. 2019). More specifically, both the January 2021 SSW and underlying 2021 La Niña conditions have been hypothesized as important to the development of the February 2021 CAO (Lee 2021; Lu et al. 2021; Zhang et al. 2021). Additionally, it is important to assess—in a probabilistic framework—how much of the CAO was due to these more predictable dynamical processes versus internal variability that is largely unpredictable beyond about a 2-week forecast lead.

The above studies motivate a phenomenological approach to the 2021 CAO, where we diagnose the effects of different dynamical processes and determine their relative importance

to the CAO's subseasonal prediction. However, isolating the potentially predictable processes from one another that are relevant to this and other CAOs can be difficult because they often evolve on similar time scales and can project onto similar spatial patterns. For example, extratropical teleconnections from La Niña (Johnson et al. 2014; Kim et al. 2021; Smith and Sheridan 2021), phase 7 of the MJO (Johnson et al. 2014; Tseng et al. 2018), and SSWs (Butler et al. 2017) tend to project onto similar anomaly patterns with low geopotential heights and cold surface temperatures extending from the eastern CONUS upward to northwestern Canada. Thus, as they evolve, the anomaly patterns associated with these different climate processes can both constructively and destructively interfere with one another, which means that these patterns are "nonorthogonal." Consequently, techniques that by construction yield orthogonal modes [e.g., empirical orthogonal function (EOF) analysis] are ill-suited to isolate actual *dynamical* modes of the climate system (Monahan et al. 2009). Likewise, any method that examines spatial and temporal scales separately, say by first identifying orthogonal spatial patterns and then temporal scales via spectral filtering, may convolve the effects of many different dynamical processes together (see "Dynamical climate modes and the LIM dynamical filter" sidebar).

To address these difficulties, then, requires a diagnostic approach that decomposes climate anomalies into their contributions from overlapping dynamical processes. Here, we use a

Dynamical climate modes and the LIM dynamical filter

Slowly varying climate anomalies are made up of generally "nonorthogonal" dynamical modes (e.g., Farrell and Ioannou 1996; Penland and Matrosova 2006; Coy and Reynolds 2014; Henderson et al. 2020; Albers and Newman 2021), which typically have similar spatial patterns yet evolve on different but overlapping time scales. This nonorthogonality is a consequence of fundamental asymmetries within the physical climate system (Farrell and Ioannou 1996). For example, winds directly force the ocean while oceanic currents only indirectly force the atmosphere, and contrasts in orography and diabatic heating combine to drive large spatial differences in how both the atmosphere and ocean transport heat and momentum (e.g., Borges and Sardeshmukh 1995; Moore and Kleeman 1999). The resulting nonorthogonal dynamical modes can interfere with one another either constructively or destructively, allowing rapid anomaly growth or decay while also making the drivers of the anomalies difficult to separate. That is, an evolving climate anomaly might change its shape from one that can extract energy from the basic state (say, by being tilted against sheared flow) to one that gives up energy to the basic state (say, by being rotated until it is tilted *with* the sheared flow). Such transient anomaly growth could give rise to a predictable climate signal (see "Identifying S2S forecasts of opportunity in the LIM" sidebar), and its dynamics may be captured by a LIM.

Moreover, phenomena such as ENSO and the MJO are not strictly periodic but rather evolve episodically and irregularly, making them "broadband" processes with substantial power lying well outside their spectral peaks: ENSO power not only maximizes over a broad range of periods (about 4–7 years) but also extends well into subseasonal and decadal time scales, and the MJO's 30–60-day spectral peak is broadened by irregular eastward propagation. ENSO also undergoes substantial structural evolution, both over its life cycle and across different events ("ENSO diversity"; Capotondi et al. 2015), so that it and its global impacts cannot be represented by a single pattern or index (Zhao et al. 2021). Thus, ENSO and MJO tropical and extratropical anomalies are convolved together, so that distinguishing between them requires a more sophisticated approach than regression on a few pattern-based index time series, even when temporally filtered [see Newman et al. (2009) and Henderson et al. (2020) for detailed discussion].

This motivates the LIM dynamical filter, constructed by decomposing the LIM operator into eigenmodes containing information about the *e*-folding decay time, period of oscillation, and relative contribution of each variable to each eigenmode. For example, ENSO-related modes have long periods and strong SST amplitudes, while modes with shorter periods, no SST amplitudes, and weak stratospheric amplitudes are associated with tropospheric internal variability. Any climate anomaly (including any LIM forecast or JRA-55 observation) can be projected onto the LIM state vector and subsequently filtered (see the supplemental material for technical details of filter creation). The filter used here decomposes modeled, predicted, and observed anomalies into contributions from atmospheric internal variability, tropical SSTs (including ENSO), a single MJO mode, and a single downward propagating stratospheric NAM that includes SSW effects.

LIM to construct such a "dynamical filter" (Penland and Matrosova 2006), which isolates individual dynamical modes driven by interactions between the tropics, extratropics, and stratosphere (Newman et al. 2009; Henderson et al. 2020; Albers and Newman 2021). The LIM is a good tool for this analysis since its forecast operator is deduced from climate statistics in a way that simultaneously accounts for *spatial and temporal* information, and hence yields dynamical modes that can be, and nearly always are, nonorthogonal. A dynamical filter can be applied to LIM forecasts and JRA-55 observations to conduct dynamical process attribution and to assess the predictability of the individual climate components. The LIM filter used here isolates tropical SSTs (including La Niña), the effects of downward propagating stratospheric Northern Annular Mode (NAM) anomalies (including SSWs), the MJO, and other aspects of internal variability. The SST, stratosphere, and MJO groups of modes capture dynamical processes that are potentially predictable and interpretable on subseasonal time scales. Higher frequency processes including extratropical synoptic variability (e.g., blocking; Breeden et al. 2020) and tropical waves (e.g., convectively coupled waves; Kiladis et al. 2009), on the other hand, are captured by the dynamical modes making up the interval variability subspace. While these modes substantially contribute to overall variability, they have short *e*-folding times and therefore provide little subseasonal predictability (e.g., Janiga et al. 2018; Breeden et al. 2020; Schreck et al. 2020; Albers and Newman 2021). See "Dynamical climate modes and the LIM dynamical filter" sidebar for further explanation of the above points and the supplemental material for more details of the filter construction.

Applying the LIM dynamical filter to the 2-m temperature forecast verification for 8–21 February 2021 reveals that the largest contributions to the CAO came from internal variability and La Niña teleconnections (Figs. 2a and 2b, respectively). An eastward-propagating MJO and the residual effects from the early January SSW also contributed (Figs. 2c and 2d, respectively), though the magnitudes of those contributions were notably smaller (note the different color scales). The latitude–time cross section of 2-m temperature (longitudinally averaged between 235° and 270°E) in Fig. 3 reveals how critical the timing of the four dynamical modes was to the total cold anomaly. For example, the cold anomalies due to internal variability (Fig. 3b) and the MJO (Fig. 3d) were highly transient and bookended by warm anomalies. And despite winterlong La Niña SST conditions (Fig. S1 in the online supplemental material),

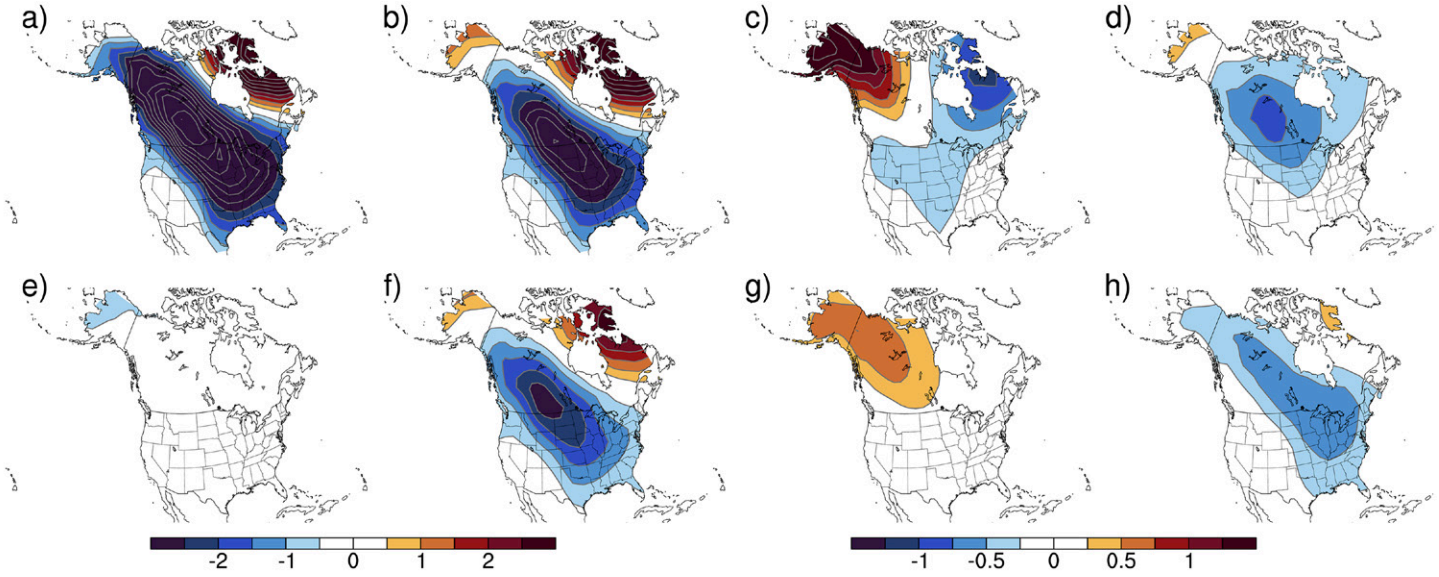

Fig. 2. (top) JRA-55 observations and (bottom) LIM forecast filtered into contributions from (a),(e) internal variability, (b),(f) tropical SSTs, (c),(g) the MJO, and (d),(h) downward propagating stratospheric NAM/SSW anomalies. Note the color scale for the internal variability and tropical SST anomalies is 2 times larger than that of the MJO and stratospheric anomaly color scales. Units are in degrees Celsius.

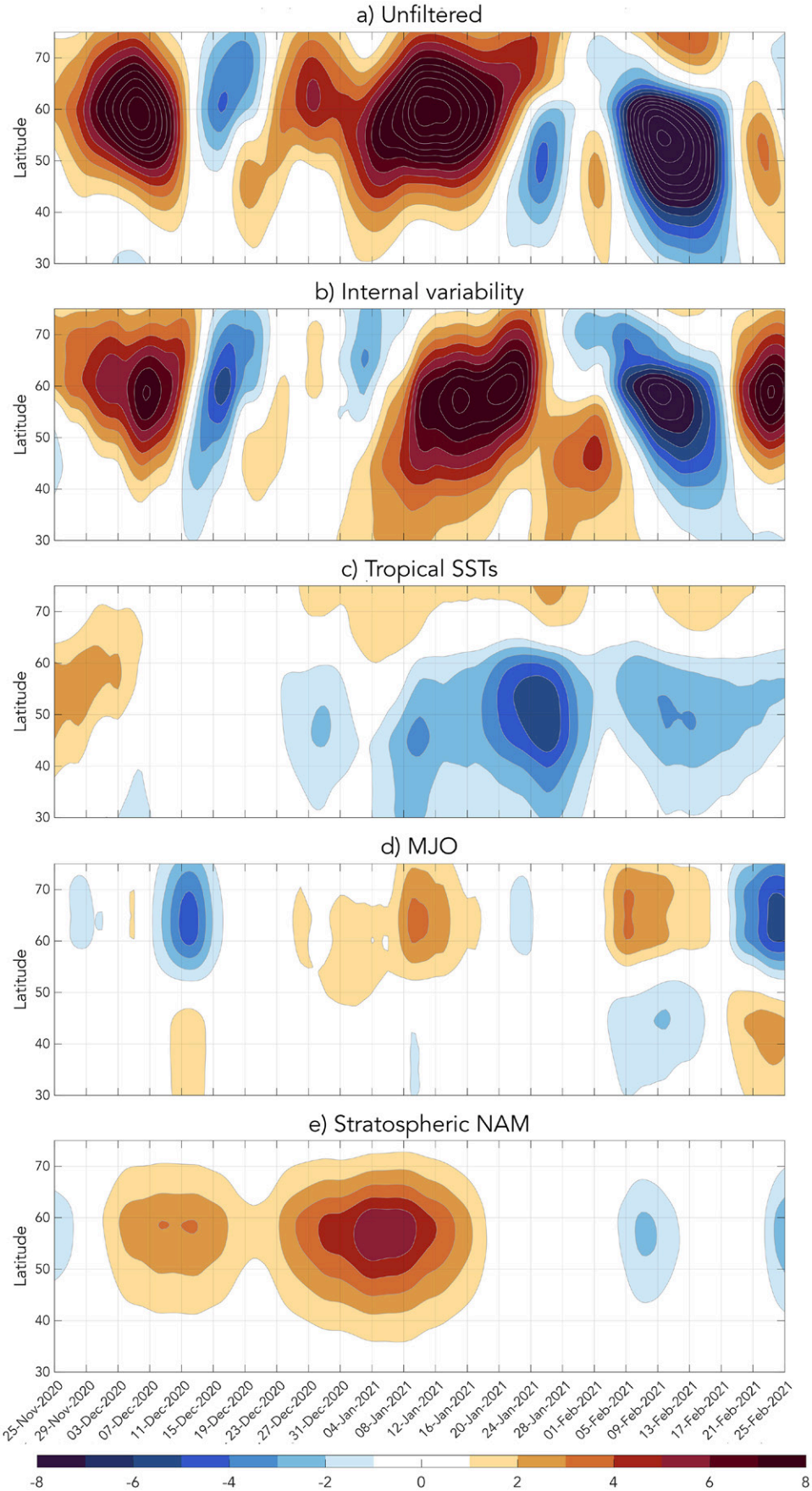

Fig. 3. Latitude–time cross section of 2-m temperature (°C; longitudinally averaged between 235° and 270°E) using JRA-55 verifications filtered into contributions from (a) no filter applied, (b) internal variability, (c) tropical SSTs, (d) the MJO, and (e) downward propagating stratospheric NAM/SSW anomalies.

SST-related cold anomalies over the CONUS did not appear prior to January (Fig. 3c). Likewise, the stratosphere actually contributed warm anomalies during most of January, which was likely related to anomalously strong December stratospheric polar vortex conditions (e.g., https://acd-ext.gsfc.nasa.gov/Data_services/met/metdata/annual/merra2/wind/u60n_10_2020_merra2.pdf), and the cold anomaly following the early January SSW only appeared during the first two weeks of February (Fig. 3e).

While Figs. 2a–d and Fig. 3 diagnose the relative dynamical contributions to the CAO, they do not show which processes were ultimately *predictable* 3–4 weeks in advance. From the filtered LIM forecasts, we find that the high CAO forecast probability (and high LIM signal-to-noise ratio) was almost completely due to the more predictable La Niña teleconnections (Fig. 2f) and downward impacts related to the SSW (Fig. 2h). The LIM predicted a portion of the weaker MJO signal (Fig. 2g) but limited to the MJO-induced warm anomaly over Alaska and far northwestern Canada. And as expected, the internal variability contribution to the CAO (Fig. 2e) was unpredictable on this weeks 3–4 lead time (Albers and Newman 2021).

How likely was a CAO in 2021?

Given the combination of processes involved in the February CAO, how likely was a CAO beginning on 1 December 2020, and how did the likelihood of a CAO change as the season progressed? For example, La Niña played a central role in the February CAO (Fig. 2b), yet La Niña conditions were present all winter long and there was no CAO in December or January. To quantify how La Niña and the early January SSW increased the probability and relative risk of a severe CAO, we conducted two pairs of data denial reforecast experiments (Table 1). Each reforecast experiment consists of a 5,000-member LIM forecast ensemble, where the first pair of experiments was initialized using data from 1 December 2020 and the second pair was initialized using data from 24 January 2021. All forecasts from both pairs of experiments verify during the same 2-week period shown in Fig. 2 (i.e., 8–21 February 2021); that is, the first experiment is a weeks 11–12 outlook, while the second experiment is a weeks 3–4 outlook. LIM forecast ensembles are generated by repeated forecast integrations forced by different, but observationally constrained, Gaussian white noise realizations, rather than by forecast initializations with different perturbed initial conditions. Thus, the white noise forcing in the LIM represents the chaotic effects of initial random uncertainties upon the evolution of the anomalous climate state (see supplemental material for additional technical details). Data denial is accomplished by applying the LIM dynamical filter to the initial conditions of the reforecasts. The probability of a CAO is quantified by calculating the average temperature inside the CAO region (outlined region in Fig. 1b) for the 8–21 February 2021 verification period for each ensemble member in each respective reforecast experiment.

For the first pair of reforecasts initialized on 1 December 2020, one ensemble is given observed initial conditions, which on that day includes moderate La Niña SSTs, while the

Table 1. LIM reforecast data denial experiments. Dynamical modes are suppressed by applying the LIM dynamical filter to the initial conditions (see "Dynamical climate modes and the LIM dynamical filter" sidebar and the supplemental material for details).

Exp. No.	Initialization date	Verification date	Initialization conditions
1	1 Dec 2020	8–21 Feb 2021	Observed initial conditions
2	1 Dec 2020	8–21 Feb 2021	Suppressed tropical SSTs (neutral ENSO)
3	24 Jan 2021	8–21 Feb 2021	Observed initial conditions
4	24 Jan 2021	8–21 Feb 2021	Suppressed SSW/NAM effects

second ensemble is given initial conditions where the effects of tropical SSTs are suppressed via the dynamical filter. Comparing these two reforecast ensembles (Fig. 4a) reveals that antecedent La Niña conditions during December increased the probability of cold conditions during mid-February, though the differences are only significant for CAOs of moderate severity even though the two forecast distributions are statistically different according to a two-sample Kolmogorov–Smirnov test. For example, the medians of the two December reforecast distributions are statistically distinct (i.e., the bootstrap confidence intervals do not overlap, not shown), but the distributions are not distinct for more severe CAOs (e.g., the bootstrap confidence intervals for the 1st percentiles of the two distributions overlap in Fig. 4a). The relatively moderate shift in the PDF that includes La Niña conditions is in part a result of the 11–12-week forecast lead time, which is sufficiently long that some ensemble members transition from La Niña conditions to neutral or even El Niño conditions. A related possibility is that the observed westward shift of the cold SST anomaly from 235°E to 180° in mid-January (supplemental material Fig. S1), was critical to the rapid onset of CONUS cold anomalies shown in Fig. 3c.

For the second pair of reforecasts initialized on 24 January 2021, one ensemble is given observed initial conditions, which again includes moderate La Niña SSTs, while the second ensemble is given initial conditions where the effects of the early January SSW are suppressed via the dynamical filter. Note that while La Niña SSTs on 24 January were similar in magnitude to the previous 1 December conditions, the location of the SST maximum was shifted notably westward, as mentioned above (Fig. S1). Initial conditions on 24 January were such that the probability of a CAO in February had now increased significantly. Interestingly, despite the SSW contributing a seemingly small cold anomaly to the total CAO pattern (Figs. 2d,h, and 3e), the probability of a severe CAO (1st percentile

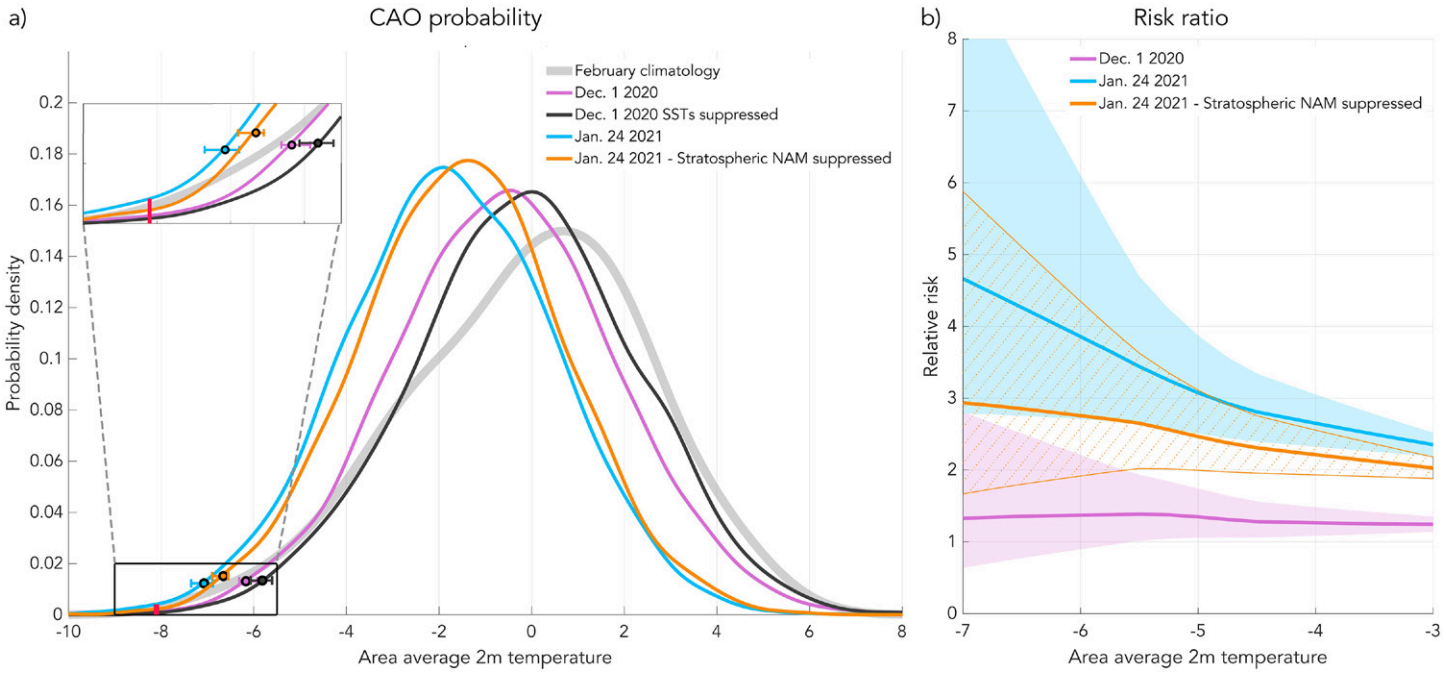

Fig. 4. (a) Area average 2-m temperature forecast probability density functions (PDFs) for the 2-week average verifications of the four reforecast experiments listed in Table 1. The thick gray curve denotes the climatological PDF for 5,000 two-week average periods randomly subsampled from all February 2-week periods between 1979 and 2017. The 95th percentile bootstrap confidence intervals (based on 10,000 bootstrap ensembles with replacement) are shown as whiskers. (b) The risk ratio for the reforecasts in Table 1, defined here as the probability CAOs of various magnitudes in each of the reforecast experiments 1, 3, or 4 relative to the probability of a CAO in reforecast experiment 2. The 95th percentile bootstrap confidence intervals (based on 10,000 bootstrap ensembles with replacement) are depicted via the shaded or hatched regions. Units in both panels are in degrees Celsius.

highlighted in Fig. 4a) is significantly increased when the effects of the SSW are included together with La Niña SSTs.

To better understand how La Niña and the early January SSW contributed to the risk of a wide range of CAO severity outcomes, we calculate the risk ratio (National Academies of Sciences, Engineering, and Medicine 2016b) between reforecasts that included La Niña and/or the SSW (experiments 1, 3, and 4 in Table 1) relative to the 1 December reforecast that was initialized with neutral ENSO conditions (experiment 2 in Table 1). Thus, the risk ratio is defined as the probability of a CAO of a particular severity occurring in experiments 1, 3, or 4 relative to the probability of a CAO of the same severity occurring in experiment 2. The CAO severity is defined for various thresholds of the area-average 2-m temperature in the central U.S. CAO region used in Fig. 4a (see Fig. 1b for area outline) averaged over the 2-week period of 8–21 February 2021.

Relative to ENSO neutral conditions, the La Niña SST conditions observed on 1 December nominally increased the risk of a mild to severe CAO during February (Fig. 4b). By 24 January, the risk due to La Niña *only* (experiment 4) made the risk of a CAO 2–3 times more likely. While the SSW only moderately increased the risk of a mild CAO relative to La Niña alone, the SSW made a severe CAO 3–5 times more likely, which is consistent with the significant shift in the tails of the distribution shown in Fig. 4a. It is worth noting that despite the fact that SSWs themselves are typically not predictable beyond 10 days in advance (Karpechko 2018; Stan and Straus 2009; Rao et al. 2021), once the SSW has occurred, it contributes to a sustained increase in the risk of a severe CAO (e.g., Huang et al. 2021). Indeed the 2021 SSW initiated on 5 January (Lee 2021), yet 3 weeks later, its aftereffects were still a source of S2S forecast skill (Fig. 2h).

How often could we expect a similar CAO to occur?

To determine whether the combination of dynamical processes involved in the 2021 CAO (La Niña, an SSW, and MJO) is typical of the most severe CAOs and to reveal how the different dynamical processes underlying such extreme events independently evolve, we use a 3,000-yr LIM climate simulation (see supplemental material for climate simulation details). Like the ensemble reforecasts, the LIM climate simulation is integrated using observationally constrained white noise forcing, which yields LIM climate statistics that are in excellent agreement with observations (e.g., Newman and Sardeshmukh 2008). The resulting simulation is analogous to a long fixed-climate simulation made by a numerical climate model, such as those used in climate change attribution studies (e.g., Hoell et al. 2021). From the LIM climate simulation, all 2-week average 2-m temperature anomalies for the central U.S. CAO region (see Fig. 1b) that are as or more severe than the 2021 event (threshold of −8.1°C, see Fig. 4a) are used to construct a time-lagged composite evolution of a "typical" extreme CAO (Figs. 5 and 6).

The individual contributions from each of the dynamical processes for the zero-lag 2-m temperature composite matches the 2021 CAO verification exceptionally well (cf. middle row of Fig. 5 and top row of Fig. 2). The temperature anomalies that were more predictable during the 2021 CAO—those associated with tropical SSTs and downward propagating stratospheric anomalies—evolve slowly with same signed anomalies over the course of 6 weeks (Figs. 5b and 5d, from lag −2 to lag 2). The stronger, tropically forced temperature anomalies are associated with an Arctic Oscillation–like teleconnection pattern that is excited by enhanced western Pacific tropical heating (Fig. 6a) due to La Niña SST conditions (not shown). On the other hand, the stratospheric 500-mb geopotential height pattern (Fig. 6b), which projects more strongly onto the NAO and has no tropical heating or SST signature (not shown), evolves on the 40–60-day time scale typical of strong NAM/polar-night jet oscillation events (Baldwin and Dunkerton 2001; Hitchcock et al. 2013). These geopotential height patterns (Figs. 6a,b) are consistent with Albers and Newman (2021), who showed that tropical SST-related

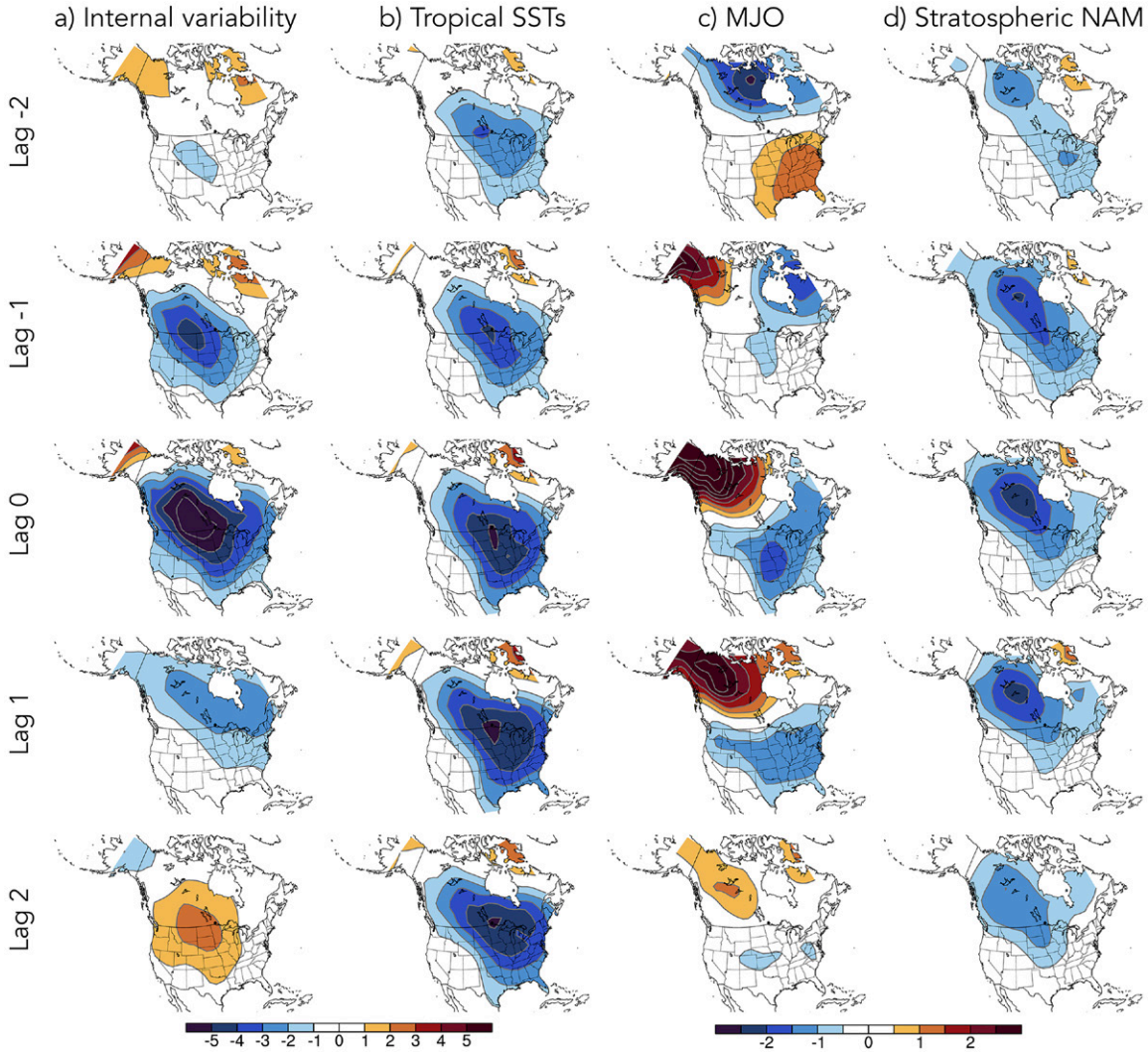

Fig. 5. Time-lagged 2-m temperature anomalies (°C) composited from the most severe CAOs in the 3,000-yr LIM climate simulation (2-week area average temperature < 8.1°C). The anomalies are filtered into contributions from (a) internal variability, (b) tropical SSTs, (c) the MJO, and (d) downward propagating stratospheric NAM/SSW anomalies. The lagged composites are created by identifying the events that meet the CAO criterion at lag 0, and then averaging the corresponding anomalies that are lagged relative to that time. Each lag represents a 2-week average that overlaps with the adjacent lag by one week (e.g., lag –2 is composited from days –14 to –1, lag –1 is composited from days –7 to 7, and lag 0 is composited from days 1 to 14). Note the color scale for the internal variability and tropical SST anomalies is 2 times larger than that of the MJO and stratospheric anomaly color scales.

teleconnections typically project more strongly onto the Arctic Oscillation (Thompson and Wallace 2001), while downward propagating stratospheric NAM anomalies project more strongly onto the NAO (see also Butler et al. 2017). In combination, the tropical SST and stratospheric geopotential height teleconnections roughly sum to the predicted LIM anomaly during the 2021 CAO (cf. Figs. 6a and 6b to Fig. 2b).

Unlike the SST and stratospheric contributions, the 2-m temperature and 500-mb geopotential height patterns associated with internal variability and the MJO are much more transient. The internal variability portion of the CAO follows the evolution of a retrogressing North Pacific block (Fig. 6c), with the peak cold anomaly lasting less than 3 weeks (Fig. 5a). This blocking pattern is in many respects quite canonical (e.g., Breeden et al. 2020, and references therein), with enhanced tropical heating over the Maritime Continent and suppressed heating over the central tropical Pacific, and anomalous warmth over Alaska and cold temperatures over the

interior of North America. As highlighted by Breeden et al. (2020), North Pacific blocks sustain their strongest growth over a 10–14-day period and appear largely unpredictable beyond 2 weeks, consistent with the inability of the LIM and IFS to predict the internal variability portion of the CAO 3–4 weeks in advance (Figs. 1 and 2).

The MJO composite shows evolving 500-mb geopotential height anomalies (Fig. 6d) consistent with RMM MJO phases 6–8 (e.g., Wheeler and Hendon 2004; Tseng et al. 2018), along with an initially warm CONUS anomaly progressing to a cold surge lasting a few weeks (Fig. 5c). Its evolution corresponds well with both the observed temperature anomaly (Fig. 2c) and, in the tropics, the observed MJO phase evolution during the 2021 CAO (observed phase propagation of the RMM MJO index is available at www.bom.gov.au/climate/mjo/).

To place the 2021 CAO into historical context and to estimate the future risk of such events, the LIM climate run was used to calculate the return time (recurrence interval) of CAOs of various severities, which here we calculate as the average time (National Academies of Sciences, Engineering, and Medicine 2016b) between CAOs of a specified severity, determined from specific ranges of CAO severity, not exceedances. For example, in Fig. 7, −8°C corresponds to the average return time for CAOs where the 2-week average temperatures over central North American (outlined region in Fig. 1b) reached between −7.5° and −8°C. Thus, for a CAO like the February 2021 event (−8.1°C), the LIM climate run estimates the average return time for an event between −8° and −8.25°C (specific bar not shown) as roughly 26 years. For comparison,

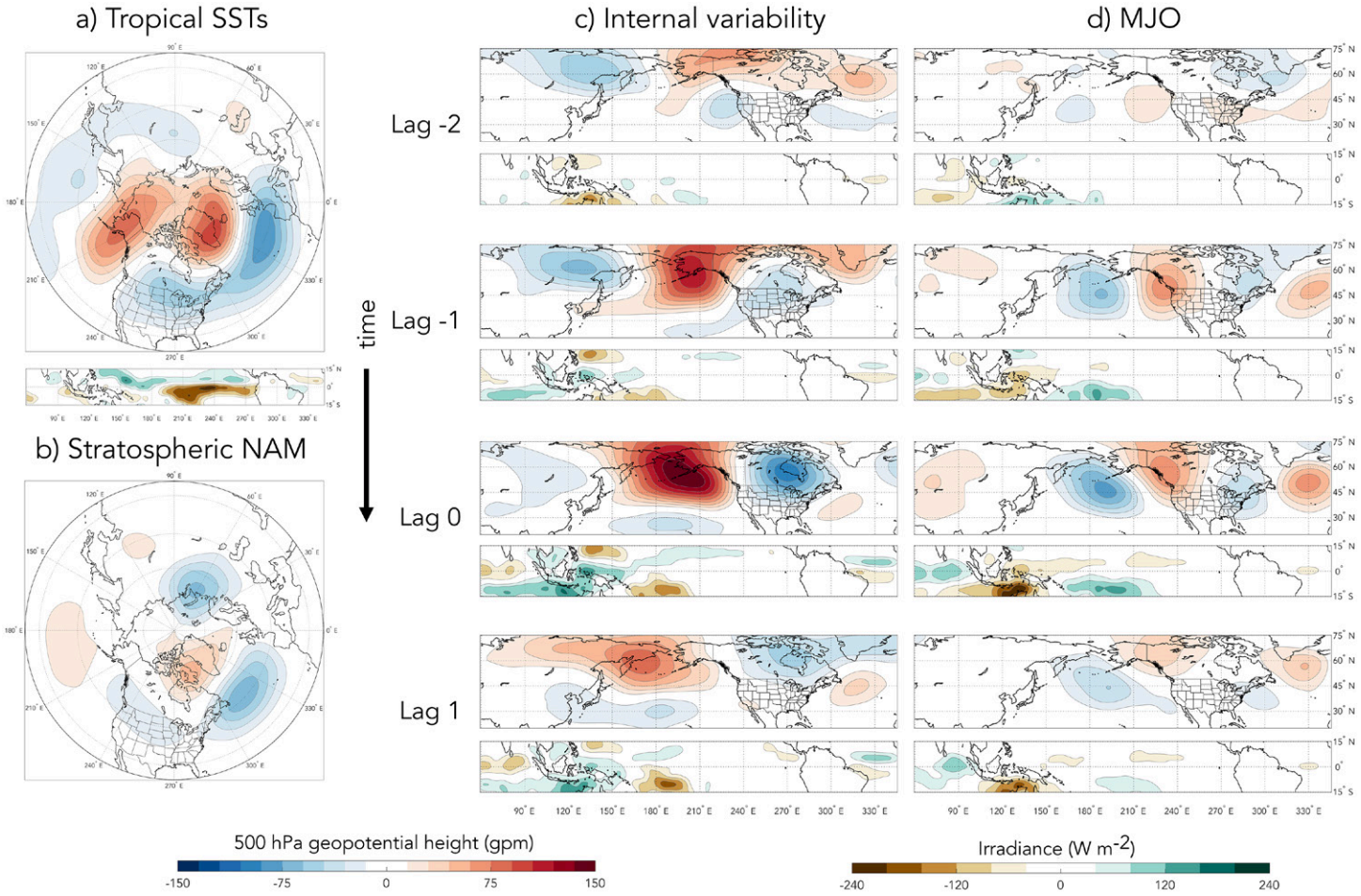

Fig. 6. The 500-mb geopotential height (gpm) and column integrated tropical irradiance (W m^{-2}) composited from the most severe CAOs in the 3,000-yr LIM climate simulation (2-week area average temperature < 8.1°C). The anomalies are filtered to show (a) the lag 0 composites of geopotential height and tropical irradiance related to tropical SSTs, (b) the lag 0 composite of geopotential height related to downward propagating stratospheric NAM anomalies, (c) the time-lagged composites of geopotential height and tropical irradiance related to internal variability, and (d) the time-lagged composites of geopotential height and tropical irradiance related to the MJO.

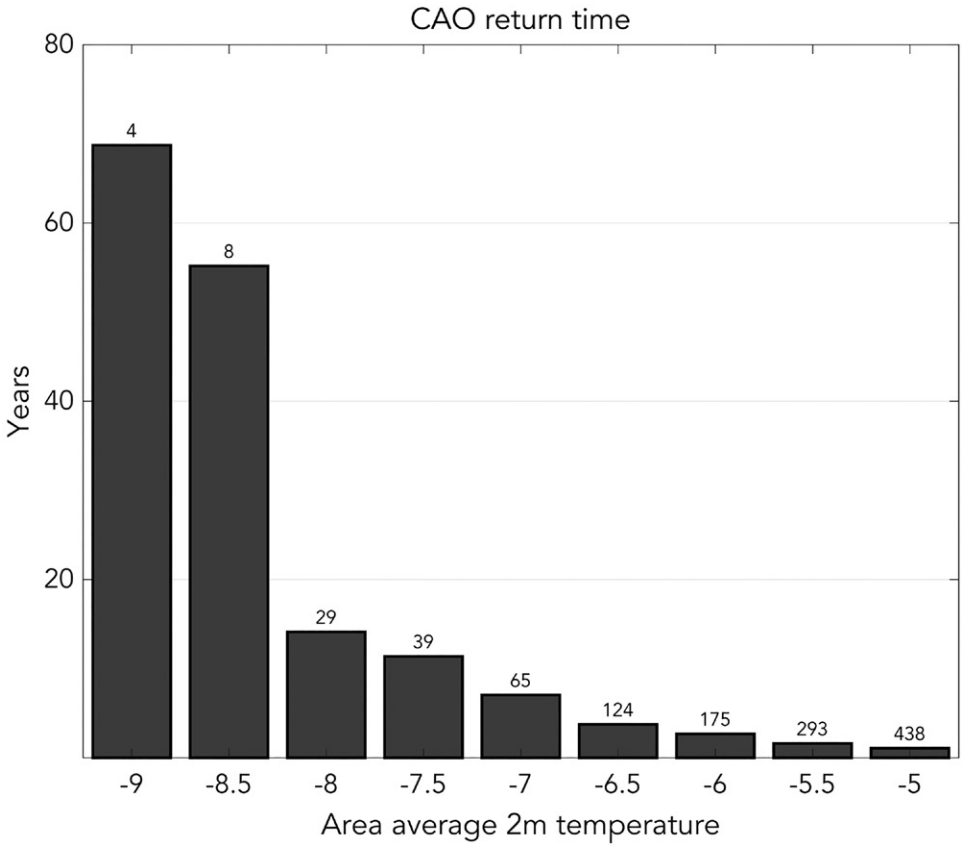

Fig. 7. Average return times (in years) for various severities of CAOs from the 3,000-yr LIM climate simulation. The numbers at the top of each bar denote the number of CAOs of the corresponding severity that occurred in the climate simulation. The 2-m temperature units are in degrees Celsius.

the last CONUS CAO of a similar magnitude occurred 32 years ago in late December 1989 (www.weather.gov/ilx/dec1989-cold; FERC 2021).

While the most extreme CAOs *typically* involve contributions for all four groupings of dynamical modes (Fig. 5), severe CAOs can result from any combination of the modes. However, the return time for a severe CAO can notably change depending on which dynamical modes are involved. For example, additional return time calculations conducted by applying our dynamical filter reveal that when only ENSO or an SSW-like event occurs in combination with internal variability, severe CAO return times increase only modestly; however, when only internal variability and the MJO co-occur, return times are significantly longer. The notably longer return time for MJO-based CAOs can be understood by considering the anomaly evolution differences shown in Fig. 5. La Niña and SSWs cause a slowly evolving, single-signed cold anomaly that spans more than a 2-week period (e.g., Figs. 5b,d), whereas the MJO anomaly switches sign (warm to cold) over a 2–3-week period (e.g., Fig. 5c), which means that the phasing between internal variability and any MJO cold signal must perfectly align in order to not destructively interfere and weaken the total cold anomaly.

Looking to the future

Subseasonal-to-seasonal forecast outlooks are typically based on some combination of guidance from an ensemble of forecast models, methods for identifying forecasts of opportunity, and knowledge of current climate mode conditions (e.g., ENSO, the MJO, or SSWs). Thus, we suggest that to be useful to operational subseasonal-to-seasonal forecasters, machine learning models need to meet one (ideally all) of the following requirements: the model must have forecast skill comparable to operational numerical forecasting models,

the model should help identify forecasts of opportunity *at time of forecast*, and the model should relate forecasts of opportunity to known dynamical climate modes. The LIM does all of these, and several other machine learning approaches currently under development may also be aiming to meet these requirements to varying degrees (e.g., Ham et al. 2019; Scheuerer et al. 2020; Qian et al. 2020; Buchmann and DelSole 2021; Charlton-Perez et al. 2021; Martin et al. 2022; Mayer and Barnes 2021; Silini et al. 2021; Toms et al. 2021; van Straaten et al. 2022).

The LIM presented here was able to predict some portion of the 2021 CAO up to four weeks in advance, identify the 2021 CAO as a forecast of opportunity, and identify the dynamical processes responsible for the predictable and unpredictable portions of the event. However, the extreme nature of the CAO was not completely predictable on subseasonal time scales, due to the large cold contribution from internal variability that was not predicted by the LIM or operational numerical forecast models until 10–14 days prior to the onset of the event. Moreover, within the shorter, sub-2-week forecast lead time frame, numerical weather forecasting models are currently, on average, more skillful than machine learning models (e.g., Albers and Newman 2019; Weyn et al. 2021). Thus, in the future, the LIM, and other machine learning techniques like it, will likely add the most value by serving as complementary forecast guidance tools used in conjunction with numerical forecast models.

Acknowledgments. The authors thank Joseph Barsugli, Samuel Lillo, and NOAA CPC week 3–4 personnel for their helpful comments and discussions that helped to improve this manuscript. The CPC LIM was developed via funding from a National Weather Service Level Agreement. Funding from the NOAA–CIRES cooperative agreement (NA17OAR4320101) helped support this work.

Data availability statement. The data that support the findings of this study are openly available at the following URL/DOI: ftp://ftp2.esrl.noaa.gov/Projects/LIM/Weekly/Albers_etal_BAMS2022/. IFS data were originally accessed via the ECMWF data portal: https://apps.ecmwf.int/datasets/.

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.

Baldwin, M. P., and T. J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science*, **294**, 581–584, https://doi.org/10.1126/science.1063315.

Borges, M. D., and P. D. Sardeshmukh, 1995: Barotropic Rossby wave dynamics of zonally varying upper-level flows during northern winter. *J. Atmos. Sci.*, **52**, 3779–3796, https://doi.org/10.1175/1520-0469(1995)052<3779:BRWDOZ>2.0.CO;2.

Breeden, M. L., 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.

Buchmann, P., and T. DelSole, 2021: Week 3–4 prediction of wintertime CONUS temperature using machine learning techniques. *Front. Climate*, **3**, 81, https://doi.org/10.3389/fclim.2021.697423.

Buizza, R., and M. Leutbecher, 2015: The forecast skill horizon. *Quart. J. Roy. Meteor. Soc.*, **141**, 3366–3382, https://doi.org/10.1002/qj.2619.

Butler, A. H., J. P. Sjoberg, D. J. Seidel, and K. H. Rosenlof, 2017: A sudden stratospheric warming compendium. *Earth Syst. Sci. Data*, **9**, 63–76, https://doi.org/10.5194/essd-9-63-2017.

Cai, M., Y. Yu, Y. Deng, H. M. van den Dool, R. Ren, S. Saha, X. Wu, and J. Huang, 2016: Feeling the pulse of the stratosphere: An emerging opportunity for predicting continental-scale cold-air outbreaks 1 month in advance. *Bull. Amer. Meteor. Soc.*, **97**, 1475–1489, https://doi.org/10.1175/BAMS-D-14-00287.1.

Capotondi, A., and Coauthors, 2015: Understanding ENSO diversity. *Bull. Amer. Meteor. Soc.*, **96**, 921–938, https://doi.org/10.1175/BAMS-D-13-00117.1.

Cellitti, M. P., J. E. Walsh, R. M. Rauber, and D. H. Portis, 2006: Extreme cold air outbreaks over the United States, the polar vortex, and the large-scale circulation. *J. Geophys. Res.*, **111**, D02114, https://doi.org/10.1029/2005JD006273.

Charlton-Perez, A. J., J. Bröcker, A. Y. Karpechko, S. H. Lee, M. Sigmond, and I. R. Simpson, 2021: A minimal model to diagnose the contribution of the stratosphere to tropospheric forecast skill. *J. Geophys. Res. Atmos.*, **126**, e2021JD035504, https://doi.org/10.1029/2021JD035504.

Coelho, C. A., B. Brown, L. Wilson, M. Mittermaier, and B. Casati, 2019: Forecast verification for s2s timescales. *Sub-Seasonal to Seasonal Prediction*, F. Vitart and A. W. Robertson, Eds., Elsevier, 337–361.

Coy, L., and C. A. Reynolds, 2014: Singular vectors and their nonlinear evolution during the January 2009 stratospheric sudden warming. *Quart. J. Roy. Meteor. Soc.*, **140**, 1013–1024, https://doi.org/10.1002/qj.2181.

Domeisen, D. I., and A. H. Butler, 2020: Stratospheric drivers of extreme events at the Earth's surface. *Commun. Earth Environ.*, **1**, 59, https://doi.org/10.1038/s43247-020-00060-z.

Farrell, B. F., 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.

FERC, 2021: The February 2021 cold weather outages in Texas and the south central United States. FERC Rep., 316 pp., www.ferc.gov/media/february-2021-cold-weather-outages-texas-and-south-central-united-states-ferc-nerc-and.

Ham, Y.-G., J.-H. Kim, and J.-J. Luo, 2019: Deep learning for multi-year ENSO forecasts. *Nature*, **573**, 568–572, https://doi.org/10.1038/s41586-019-1559-7.

Hasselmann, K., 1976: Stochastic climate models: Part I. Theory. *Tellus*, **28**, 473–485, https://doi.org/10.3402/tellusa.v28i6.11316.

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.

Hitchcock, P., T. G. Shepherd, and G. L. Manney, 2013: Statistical characterization of Arctic polar-night jet oscillation events. *J. Climate*, **26**, 2096–2116, https://doi.org/10.1175/JCLI-D-12-00202.1.

Hoell, A., M. Hoerling, J. Eischeid, and J. Barsugli, 2021: Preconditions for extreme wet winters over the contiguous United States. *Wea. Climate Extremes*, **33**, 100333, https://doi.org/10.1016/j.wace.2021.100333.

Huang, J., P. Hitchcock, A. C. Maycock, C. M. McKenna, and W. Tian, 2021: Northern Hemisphere cold air outbreaks are more likely to be severe during weak polar vortex conditions. *Commun. Earth. Environ.*, **2**, 147, https://doi.org/10.1038/s43247-021-00215-6.

Janiga, M. A., C. J. Schreck, J. A. Ridout, M. Flatau, N. P. Barton, E. J. Metzger, and C. A. Reynolds, 2018: Subseasonal forecasts of convectively coupled equatorial waves and the MJO: Activity and predictive skill. *Mon. Wea. Rev.*, **146**, 2337–2360, https://doi.org/10.1175/MWR-D-17-0261.1.

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.

Karpechko, A. Y., 2018: Predictability of sudden stratospheric warmings in the ECMWF extended-range forecast system. *Mon. Wea. Rev.*, **146**, 1063–1075, https://doi.org/10.1175/MWR-D-17-0317.1.

Kenyon, J., and G. C. Hegerl, 2008: Influence of modes of climate variability on global temperature extremes. *J. Climate*, **21**, 3872–3889, https://doi.org/10.1175/2008JCLI2125.1.

Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Rev. Geophys.*, **47**, RG2003, https://doi.org/10.1029/2008RG000266.

Kim, S. T., Y. Y. Lee, J. H. Oh, and A. Y. Lim, 2021: Errors in the winter temperature response to ENSO over North America in seasonal forecast models. *J. Climate*, **34**, 8257–8271, https://doi.org/10.1175/JCLI-D-21-0094.1.

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.

Kolstad, E. W., T. Breiteig, and A. A. Scaife, 2010: The association between stratospheric weak polar vortex events and cold air outbreaks in the Northern Hemisphere. *Quart. J. Roy. Meteor. Soc.*, **136**, 886–893, https://doi.org/10.1002/qj.620.

Kretschmer, M., J. Cohen, V. Matthias, J. Runge, and D. Coumou, 2018: The different stratospheric influence on cold-extremes in Eurasia and North America. *npj Climate Atmos. Sci.*, **1**, 44, https://doi.org/10.1038/s41612-018-0054-4.

Lang, A. L., K. Pegion, and E. A. Barnes, 2020: Introduction to special collection: "Bridging weather and climate: Subseasonal-to-seasonal (S2S) prediction." *J. Geophys. Res. Atmos.*, **125**, e2019JD031833, https://doi.org/10.1029/2019JD031833.

Lee, S. H., 2021: The January 2021 sudden stratospheric warming. *Weather*, **76**, 135–136, https://doi.org/10.1002/wea.3966.

Loikith, P. C., and A. J. Broccoli, 2014: The influence of recurrent modes of climate variability on the occurrence of winter and summer extreme temperatures over North America. *J. Climate*, **27**, 1600–1618, https://doi.org/10.1175/JCLI-D-13-00068.1.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, https://doi.org/10.3402/tellusa.v21i3.10086.

Lu, Q., J. Rao, Z. Liang, D. Guo, J. Luo, S. Liu, C. Wang, and T. Wang, 2021: The sudden stratospheric warming in January 2021. *Environ. Res. Lett.*, **16**, 084029, https://doi.org/10.1088/1748-9326/ac12f4.

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.

Martin, Z. K., E. A. Barnes, and E. D. Maloney, 2022: Using simple, explainable neural networks to predict the Madden-Julian oscillation. *Earth Space Sci.*, **14**, e2021MS002774, https://doi.org/10.1029/2021MS002774.

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

Monahan, A. H., J. C. Fyfe, M. H. Ambaum, D. B. Stephenson, and G. R. North, 2009: Empirical orthogonal functions: The medium is the message. *J. Climate*, **22**, 6501–6514, https://doi.org/10.1175/2009JCLI3062.1.

Moore, A. M., and R. Kleeman, 1999: The nonnormal nature of El Niño and intraseasonal variability. *J. Climate*, **12**, 2965–2982, https://doi.org/10.1175/1520-0442(1999)012<2965:TNNOEN>2.0.CO;2.

National Academies of Sciences, Engineering, and Medicine, 2016a: *Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts.* National Academies Press, 350 pp., https://doi.org/10.17226/21873.

——, 2016b: *Attribution of Extreme Weather Events in the Context of Climate Change.* National Academies Press, 186 pp., https://doi.org/10.17226/21852.

NCEI, 2021a: Assessing the U.S. climate in February 2021. NOAA, www.ncei.noaa.gov/news/national-climate-202102.

——, 2021b: State of the climate: Synoptic discussion for February 2021. NOAA, accessed 15 April 2022, www.ncdc.noaa.gov/sotc/synoptic/202102.

——, 2022: U.S. billion-dollar weather and climate disasters. NOAA, www.ncei.noaa.gov/access/billions/.

Newman, M., and P. D. Sardeshmukh, 2008: Tropical and stratospheric influences on extratropical short-term climate variability. *J. Climate*, **21**, 4326–4347, https://doi.org/10.1175/2008JCLI2118.1.

——, ——, 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.

——, ——, and C. Penland, 2009: How important is air–sea coupling in ENSO and MJO evolution? *J. Climate*, **22**, 2958–2977, https://doi.org/10.1175/2008JCLI2659.1.

Pegion, K., and P. D. Sardeshmukh, 2011: Prospects for improving subseasonal predictions. *Mon. Wea. Rev.*, **139**, 3648–3666, https://doi.org/10.1175/MWR-D-11-00004.1.

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.

——, and L. Matrosova, 2006: Studies of El Niño and interdecadal variability in tropical sea surface temperatures using a nonnormal filter. *J. Climate*, **19**, 5796–5815, https://doi.org/10.1175/JCLI3951.1.

Qian, Q. F., X. J. Jia, and H. Lin, 2020: Machine learning models for the seasonal forecast of winter surface air temperature in North America. *Earth Space Sci.*, **7**, e2020EA001140, https://doi.org/10.1029/2020EA001140.

Rao, J., C. I. Garfinkel, T. Wu, Y. Lu, Q. Lu, and Z. Liang, 2021: The January 2021 sudden stratospheric warming and its prediction in subseasonal to seasonal models. *J. Geophys. Res. Atmos.*, **126**, e2021JD035057, https://doi.org/10.1029/2021JD035057.

Robertson, A., and F. Vitart, Eds., 2018: *Sub-Seasonal to Seasonal Prediction: The Gap between Weather and Climate Forecasting.* Elsevier, 588 pp.

Sardeshmukh, P. D., 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.

Scheuerer, M., M. B. Switanek, R. P. Worsnop, and T. M. Hamill, 2020: Using artificial neural networks for generating probabilistic subseasonal precipitation forecasts over California. *Mon. Wea. Rev.*, **148**, 3489–3506, https://doi.org/10.1175/MWR-D-20-0096.1.

Schreck, C. J., M. A. Janiga, and S. Baxter, 2020: Sources of tropical subseasonal skill in the CFSv2. *Mon. Wea. Rev.*, **148**, 1553–1565, https://doi.org/10.1175/MWR-D-19-0289.1.

Silini, R., M. Barreiro, and C. Masoller, 2021: Machine learning prediction of the Madden-Julian oscillation. *npj Climate Atmos. Sci.*, **4**, 57, https://doi.org/10.1038/s41612-021-00214-6.

Smith, E. T., and S. C. Sheridan, 2021: The relationship between teleconnections, surface temperature, and cold air outbreaks. *Int. J. Climatol.*, **42**, 1531–1543, https://doi.org/10.1002/joc.7318.

Stan, C., and D. M. Straus, 2009: Stratospheric predictability and sudden stratospheric warming events. *J. Geophys. Res.*, **114**, D12103, https://doi.org/10.1029/2008JD011277.

Thompson, D. W., and J. M. Wallace, 2001: Regional climate impacts of the Northern Hemisphere annular mode. *Science*, **293**, 85–89, https://doi.org/10.1126/science.1058958.

——, M. P. Baldwin, and J. M. Wallace, 2002: Stratospheric connection to Northern Hemisphere wintertime weather: Implications for prediction. *J. Climate*, **15**, 1421–1428, https://doi.org/10.1175/1520-0442(2002)015<1421:SCTNHW>2.0.CO;2.

Toms, B. A., K. Kashinath, Prabhat, and D. Yang, 2021: Testing the reliability of interpretable neural networks in geoscience using the Madden–Julian oscillation. *Geosci. Model Dev.*, **14**, 4495–4508, https://doi.org/10.5194/gmd-14-4495-2021.

Tseng, K.-C., E. Barnes, and E. 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.

van Straaten, C., K. Whan, D. Coumou, B. van den Hurk, and M. Schmeits, 2022: Using explainable machine learning forecasts to discover subseasonal drivers of high summer temperatures in western and central Europe. *Mon. Wea. Rev.*, **150**, 1115–1134, https://doi.org/10.1175/MWR-D-21-0201.1.

Vitart, F., and Coauthors, 2019: Sub-seasonal to seasonal prediction of weather extremes. *Sub-Seasonal to Seasonal Prediction*, F. Vitart and A. W. Robertson, Eds., Elsevier, 365–386.

Westby, R. M., and R. X. Black, 2015: Development of anomalous temperature regimes over the southeastern United States: Synoptic behavior and role of low-frequency modes. *Wea. Forecasting*, **30**, 553–570, https://doi.org/10.1175/WAF-D-14-00093.1.

Weyn, J. A., D. R. Durran, R. Caruana, and N. Cresswell-Clay, 2021: Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models. *J. Adv. Model. Earth Syst.*, **13**, e2021MS002502, https://doi.org/10.1029/2021MS002502.

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 A. F. Loughe, 1998: The relationship between ensemble spread and ensemble mean skill. *Mon. Wea. Rev.*, **126**, 3292–3302, https://doi.org/10.1175/1520-0493(1998)126<3292:TRBESA>2.0.CO;2.

White, C. J., and Coauthors, 2017: Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteor. Appl.*, **24**, 315–325, https://doi.org/10.1002/met.1654.

——, and Coauthors, 2021: Advances in the application and utility of subseasonal-to-seasonal predictions. *Bull. Amer. Meteor. Soc.*, **103**, E1448–E1472, https://doi.org/10.1175/BAMS-D-20-0224.1.

Winters, A. C., L. F. Bosart, and D. Keyser, 2019: Antecedent North Pacific jet regimes conducive to the development of continental U.S. extreme temperature events during the cool season. *Wea. Forecasting*, **34**, 393–414, https://doi.org/10.1175/WAF-D-18-0168.1.

Yu, B., X. Zhang, H. Lin, and J.-Y. Yu, 2015: Comparison of wintertime North American climate impacts associated with multiple ENSO indices. *Atmos.–Ocean*, **53**, 426–445, https://doi.org/10.1080/07055900.2015.1079697.

Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S. J. Lin, J. H. Chen, and K. Emanuel, 2019: What is the predictability limit of midlatitude weather? *J. Atmos. Sci.*, **76**, 1077–1091, https://doi.org/10.1175/JAS-D-18-0269.1.

Zhang, P., Y. Wu, G. Chen, and Y. Yu, 2020: North American cold events following sudden stratospheric warming in the presence of low Barents-Kara Sea sea ice. *Environ. Res. Lett.*, **15**, 124017, https://doi.org/10.1088/1748-9326/abc215.

Zhang, X., Y. Fu, Z. Han, J. E. Overland, A. Rinke, H. Tang, T. Vihma, and M. Wang, 2021: Extreme cold events from East Asia to North America in winter 2020/21: Comparisons, causes, and future implications. *Adv. Atmos. Sci.*, **39**, 553–565, https://doi.org/10.1007/s00376-021-1229-1.

Zhao, Y., M. Newman, A. Capotondi, E. D. Lorenzo, and D. Sun, 2021: Removing the effects of tropical dynamics from North Pacific climate variability. *J. Climate*, **34**, 9249–9265, https://doi.org/10.1175/JCLI-D-21-0344.1.