The February 2021 Cold Air Outbreak in the United States: a Subseasonal Forecast of Opportunity

John R. Albers\textsuperscript{a,b}, Matthew Newman\textsuperscript{a,b}, Andrew Hoell\textsuperscript{b}, Melissa L. Breeden \textsuperscript{a,b}, Yan Wang\textsuperscript{a,b}, Jiale Lou\textsuperscript{a,b}, and Samuel P. Lillo\textsuperscript{a,b}.

\textsuperscript{a} Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO
\textsuperscript{b} NOAA Physical Sciences Laboratory, Boulder, CO

Corresponding author: John Albers, john.albers@noaa.gov
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.

CAPSULE (BAMS ONLY)

The costliest winter weather event in United States history, the February 2021 cold air outbreak, was a subseasonal forecast of opportunity due to a confluence of multiple factors.

1. Motivation

The February 2021 central United States cold air outbreak (CAO) was a record-breaking event of exceptional persistence, intensity, and expanse, which became the costliest US winter storm on record (https://www.ncdc.noaa.gov/billions/events). 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 (Fig. 1a,b). During the two weeks of February 8-21, 2021, average temperatures in the region were more than 30°F below normal, with more than sixty daily minimum temperature records broken at various locations over the central CONUS (https://www.ncei.noaa.gov/news/national-climate-202102, https://www.ncei.noaa.gov/access/monitoring/monthly-report/synoptic/202102). 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 timescale is beyond the two-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 and 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 Nino-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 timescales.

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 (that is, 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 a 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 3000-year 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.

2. Forecasts

The CAO period overlapped with three NOAA Climate Prediction Center (CPC) “Week 3-4” temperature outlooks, which were issued on January 15, 22, and 29 to verify for January 30-February 12, February 6-19, and February 13-26, respectively. For the January 15 and 22 outlooks (three to four 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 (https://www.cpc.ncep.noaa.gov/products/predictions/WK34/archives/). These models include the European Centre for Medium-Range Weather Forecasting Integrated Forecast System (IFS), the National Center for Environmental Prediction Climate Forecast System version 2 (CFSv2), and Japanese Meteorological Agency Ensemble Prediction System (JMA). It was not until about two weeks prior to CAO onset (the January 29 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 two-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 Sidebar 1.] 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.
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 Sidebar 1). 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 two-week period between February 8-21, 2021, was accompanied by several notable features in the observed (Japanese 55-year Reanalysis, JRA-55) anomalous 500 mb 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 Supplement for details) that allows for additional diagnosis. Reforecasts using this LIM (Fig. 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. Fig. 1a and 1e). As a result, the LIM reforecast successfully reproduced below average temperatures for most of the CAO area (Fig. 1f).
Figure 1. JRA-55 verifications (left column), IFS forecasts (middle column), and LIM reforecasts (right column) of 500 mb geopotential height anomalies (top row) and 2m temperature anomalies (bottom row). The JRA-55 verification is shown for 8-21 February 2021. The LIM reforecast is initialized on 24 January 2021 and verifies 8-21 February 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 January 2021 and verifies 9-22 February 2021. Note that the verification color scale is three times larger than the LIM and IFS color scales. Units are geopotential meters for geopotential height and degrees Celsius for 2m temperature.

The black box overlaying panel (b) is the region (250°-270°E and 30°-45°N) used for the calculations needed to create Figs. 3-7.

3. 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 to 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 (Sidebar 1), 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

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, Qian Lu et al. 2020, 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 two-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 timescales 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 upwards 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 Sidebar 2).

To address these difficulties, then, requires a diagnostic approach that decomposes climate anomalies into their contributions from overlapping dynamical processes. Here, we use a 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 timescales. 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 Sidebar 2 for further explanation of the above points, and the Supplement for more details of the filter construction.

**Figure 2.** JRA-55 observations (top row) and LIM forecast (bottom row) filtered into contributions from (left to right): internal variability, tropical SSTs, the MJO, and downward propagating stratospheric NAM/SSW anomalies. Note the color scale for the internal variability and tropical SST anomalies is two times larger than that of the MJO and stratospheric anomaly color scales. Units are in degrees Celsius.

Applying the LIM dynamical filter to the 2m temperature forecast verification for February 8-21, 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 2m temperature (longitudinally averaged between 235°-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), 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).
Figure 3. Latitude-time cross-section of 2m temperature (longitudinally averaged between 235°-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. Units are in degrees Celsius.

206 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).

4. How likely was a CAO in 2021?

215 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 5000-member LIM forecast ensemble, where the first pair of experiments was initialized using data from December 1, 2020 and the second pair was initialized using data from January 24, 2021. All forecasts from both pairs of experiments verify during the same two-week period shown in Fig. 2 (i.e., February 8-21, 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 Supplement 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 February 8-21, 2021 verification period for each ensemble member in each respective reforecast experiment.

<table>
<thead>
<tr>
<th>Exp. #</th>
<th>Initialization date</th>
<th>Verification date</th>
<th>Initialization conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>December 1, 2020</td>
<td>February 8-21, 2021</td>
<td>Observed initial conditions</td>
</tr>
<tr>
<td>2</td>
<td>December 1, 2020</td>
<td>February 8-21, 2021</td>
<td>Suppressed tropical SSTs (neutral ENSO)</td>
</tr>
<tr>
<td>3</td>
<td>January 24, 2021</td>
<td>February 8-21, 2021</td>
<td>Observed initial conditions</td>
</tr>
<tr>
<td>4</td>
<td>January 24, 2021</td>
<td>February 8-21, 2021</td>
<td>Suppressed SSW/NAM effects</td>
</tr>
</tbody>
</table>

Table 1: LIM reforecast data denial experiments. Dynamical modes are suppressed by applying the LIM dynamical filter to the initial conditions (see Sidebar 2 and the Supplement for details).

For the first pair of reforecasts initialized on December 1, 2020, one ensemble is given observed initial conditions, which on that day includes moderate La Niña SSTs, while the 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° E in mid-January (Supplement Fig. S1), was critical to the rapid onset of CONUS cold anomalies shown in Fig. 3c.
For the second pair of reforecasts initialized on January 24, 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 January 24 were similar in magnitude to the previous December 1 conditions, the location of the SST maximum was shifted notably westward, as mentioned above (Fig. S1). Initial conditions on January 24 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, 2h, and 3e), the probability of a severe CAO (1st percentile 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 December 1 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 2m temperature in the central US CAO region used in Fig. 4a (see Fig. 1b for area outline) averaged over the two-week period of February 8-21, 2021.

Relative to ENSO neutral conditions, the La Niña SST conditions observed on December 24 nominally increased the risk of a mild to severe CAO during February (Fig. 4b). By January 24, 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 January 5 (Lee 2021), yet 3 weeks later, its aftereffects were still a source of S2S forecast skill (Fig. 2h).

5. 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, a 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 3000-year LIM climate simulation (see Supplement 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 two-week average 2m temperature anomalies for the central US CAO region (see Fig. 1b) that are as or more severe than the 2021 event (threshold of -8.1º C, see Fig. 4a) are used to construct a time-lagged composite evolution of a ‘typical’ extreme CAO (Figs. 5 and 6).
Figure 5. Time-lagged 2m temperature anomalies composited from the most severe CAOs in the 3000-year LIM climate simulation (two-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 two-week average that overlaps with the adjacent lag by one week (e.g., lag -2 is composited over days -14 to -1, lag -1 is composite over days -7 to 7, and lag 0 is composited over days 1 to 14). Note the color scale for the internal variability and tropical SST anomalies is two times larger than that of the MJO and stratospheric anomaly color scales. Units are in degrees Celsius.

The individual contributions from each of the dynamical processes for the zero-lag 2m 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

14
anomalies – evolve slowly with same signed anomalies over the course of 6 weeks (Fig. 5b and 5d, 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 timescale typical of strong NAM/polar-night jet oscillation events (Baldwin and Dunkerton 2001, Hitchcock et al. 2013). These geopotential height patterns (Fig. 6a and 6b) are consistent with Albers and Newman (2021), who showed that tropical SST-related 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. Fig. 6a and 6b to Fig. 2b).
Figure 6. 500 mb geopotential height and column integrated tropical irradiance composited from the most severe CAOs in the 3000-year LIM climate simulation (two-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. Units are geopotential meters for geopotential height and W/m² for tropical irradiance.

Unlike the SST and stratospheric contributions, the 2m 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 http://www.bom.gov.au/climate/mjo/).

Figure 7. Average return times (in years) for various severities of CAOs from the 3000-year 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. 2m temperature units are in degrees Celsius.

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 two-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, the last CONUS CAO of a similar magnitude occurred 32 years ago in late December 1989 (https://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 a 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 two-week period (e.g., Fig. 5 columns b and d), whereas the MJO anomaly switches sign (warm-to-cold) over a two to three week period (e.g., Fig. 5 column c), 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.

6. 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 DeSole 2021, Charlton-Perez et al. 2021, Martin et al. 2021, 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 timescales, 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-two week forecast lead timeframe, numerical weather forecasting models are currently, on average, more skillful than machine learning models (e.g., Albers and Newman 2019, Wyen et al. 2020). 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 would like to thank Joseph Barsugli 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 was originally accessed via the ECMWF data portal: https://apps.ecmwf.int/datasets/
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 timescales, 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 timescales 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.

Sidebar 2

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 timescales. 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 Sidebar 1), 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 timescales, and the MJO’s 30-60 day spectral peak is broadened by irregular eastward propagation. ENSO also undergoes substantial structural evolution, both over its lifecycle 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. 2022). 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 Supplement 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 Northern annular (NAM) mode that includes SSW effects.

REFERENCES


Hoell, A., M. Hoerling, J. Eischeid, J. Barsugli, 2021: Preconditions for Extreme Wet Winters over the Contiguous United States, Weather and Climate Extremes, 33, 100333


