

Analyzing Stratospheric and Tropical Contributions to the Subseasonal Forecasts of the December 2017 and January 2004 North American Cold Air Outbreaks[✉]

OLIVER T. MILLIN^{✉a} AND JASON C. FURTADO^a

^a *School of Meteorology, University of Oklahoma, Norman, Oklahoma*

(Manuscript received 18 July 2025, in final form 4 February 2026, accepted 20 March 2026)

ABSTRACT: Cold air outbreaks (CAOs) in the central United States (U.S.) are high-impact events with widespread socioeconomic consequences, as seen in February 2021. While stratospheric and tropical variability have been shown to possibly influence CAO onset, their individual roles in driving subseasonal CAO forecast skill remain unclear. This study assesses their impact using targeted nudging experiments within the Community Earth System Model, version 2 (CESM2), with the Community Atmosphere Model, version 6 (CAM6), subseasonal forecasts. We compare three forecast setups: 1) nudging only the stratosphere above 100 hPa toward reanalysis (NUDGE-S), 2) nudging the full tropical atmospheric column toward reanalysis (NUDGE-T), and 3) a freely running control with no nudging (FREE). Forecasts are initialized at various dates preceding two CAO events: 1) December 2017/January 2018, associated with strong stratospheric wave reflection and Alaskan blocking, and 2) January 2004, following a major sudden stratospheric warming. In the 2017/18 case, NUDGE-S substantially improved forecasts of negative near-surface temperature anomalies across Canada and the eastern United States while enhancing the forecast of key tropospheric regimes. NUDGE-T modestly improves temperature forecasts while strengthening the Alaskan ridge. In contrast, both nudging approaches had limited near-surface temperature impact for the 2004 event, despite improving forecasts of winter weather regimes. These findings suggest that both tropical and stratospheric variability can influence CAO prediction skill, but their impacts are event dependent. Improved subseasonal CAO prediction will require better understanding of when and how these modes contribute to surface extremes.

SIGNIFICANCE STATEMENT: Cold air outbreaks (CAOs) in the central United States can strain utilities, damage infrastructure, and pose serious risks to public health, as seen during the February 2021 event. Improving forecasts of these events weeks in advance would give communities and decision-makers critical time to prepare. This study explores how large-scale patterns in the stratosphere and tropics influence the predictability of CAOs using targeted model experiments. We find that these influences vary by event and can either enhance or reduce forecast accuracy. Understanding when and how these factors improve prediction skill can help guide development of more reliable and longer-lead forecasts, ultimately supporting better planning and resilience in the face of extreme winter weather.

KEYWORDS: North America; Cold air surges; Stratospheric circulation; Forecasting; Subseasonal variability; Tropical variability

1. Introduction

Wintertime cold air outbreaks (CAOs) are spatially large and long-lasting intrusions of polar air into the midlatitudes (Cellitti et al. 2006; Wheeler et al. 2011; Smith and Sheridan 2018, 2020). CAOs affect many regions globally, including Europe, Asia, South America, and, notably, North America (Smith and Sheridan 2020), particularly the central United States (U.S.; e.g., Davis et al. 2022; Millin et al. 2022; Millin and Furtado 2022). Central U.S. CAOs receive heightened attention due to their major socioeconomic and infrastructure impacts, as exemplified by the February 2021 CAO (Doss-Gollin et al. 2021; Bolinger et al. 2022; Millin et al. 2022; Millin and Furtado 2022). That event led to widespread power outages, traffic accidents, cold-related deaths,

and \$27 billion (U.S. dollars) in consumer price index (CPI)-adjusted damages (Doss-Gollin et al. 2021; NCEI 2025). Texas was especially hard hit due to failures in power, natural gas, and water infrastructure (Doss-Gollin et al. 2021). These recent severe impacts motivate us to better understand sources of subseasonal-to-seasonal (S2S; i.e., 2–8 weeks) predictability for CAOs to improve warning capability and mitigate damages.

The S2S forecast horizon is challenging, falling between shorter-range forecasts driven by initial conditions and weather-scale features, and seasonal forecasts influenced by slowly evolving boundary conditions in the ocean and land surface (White et al. 2017; Millin et al. 2024). Recent work has shown that central U.S. CAOs are linked to several S2S modes of variability, offering potential windows of forecasting opportunity (i.e., intermittent periods of enhanced forecast skill; Mariotti et al. 2020; Bloomfield et al. 2021; Erner and Karpechko 2024). One such mode is variability in the Northern Hemisphere (NH) stratospheric polar vortex (SPV), which significantly influences surface weather (e.g., Baldwin and Dunkerton 2001; Butler et al. 2017; Lee et al. 2019; Afarghan-Gerstman and Domeisen 2020; Millin et al. 2022). Another is the Madden–Julian oscillation (MJO;

[✉] Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-25-0420.s1>.

Corresponding author: Oliver T. Millin, omillin@ou.edu

Madden and Julian 1971, 1972), a key source of subseasonal prediction skill for the NH extratropics due to its modulation of the polar jet via Rossby wave trains (e.g., Johnson et al. 2014; Garfinkel et al. 2014; Henderson et al. 2016; Green and Furtado 2019; Quinting et al. 2024).

The connection between CAOs and the SPV is often discussed in the context of sudden stratospheric warmings (SSWs), when enhanced planetary-scale wave activity from the troposphere upwells into the stratosphere, breaks, and weakens the climatological westerlies. During major SSWs, the zonal-mean westerlies in the polar stratosphere reverse to easterly flow (e.g., Baldwin and Dunkerton 2001; Nakagawa and Yamazaki 2006; Charlton and Polvani 2007). These weakened winds may then descend into the troposphere, weakening and displacing the polar jet stream equatorward, increasing the likelihood of midlatitude CAOs (e.g., Thompson and Wallace 2000; Baldwin and Dunkerton 2001; Hitchcock and Simpson 2014; White et al. 2019; Millin et al. 2022). North American cold extremes following SPV weakening on subseasonal time scales have ties to specific North American winter weather regimes—quasi-persistent patterns in 500-hPa geopotential height (GPH) anomalies (Lee et al. 2019; Millin et al. 2022). For instance, Lee et al. (2019) found that the Arctic high (ArH) regime, marked by anomalous ridging over northern Canada and Greenland, is associated with a weak SPV and cold anomalies over central and northwestern North America. Millin et al. (2022) further showed that nearly half of central U.S. CAOs from 1950 to 2021 were ArH-type events. However, the SPV–CAO link in North America is weaker than in Europe (Lee et al. 2019; Millin et al. 2022). Instead, another mechanism, stratospheric wave reflection, plays a role in North American CAO development.

Stratospheric wave reflection occurs when planetary-scale waves encounter a negative shear layer in the stratosphere, causing them to reflect downward and amplify tropospheric wave patterns (e.g., Perlwitz and Harnik 2003, 2004; Kodera et al. 2008; Harnik 2009). Over North America, this involves upward-propagating waves from Siberia and downward-reflected waves over Canada (Kretschmer et al. 2018; Millin et al. 2022; Messori et al. 2022). Indeed, this mechanism contributed to major North American cold spells during winter 2017/18 (Matthias and Kretschmer 2020). Wave reflection is closely linked to the Alaskan ridge (AkR) regime, marked by anomalous ridging over Alaska, facilitating Arctic air intrusion into central North America (Millin et al. 2022; Messori et al. 2022; Schutte et al. 2025). AkR-type CAOs may also be influenced by tropical variability, as the ridge's development is associated with a Rossby wave train emanating from the tropical Pacific (Millin et al. 2022).

This connection between the tropics and the AkR is one example of tropical influence on North American weather. Active MJO phases, particularly phases 6–8, have also been linked to increased blocking over the North Pacific and Atlantic via Rossby wave excitation as convection moves into the tropical Pacific (Henderson et al. 2016; Henderson and Maloney 2018). Certain MJO phases also enhance wave driving into the stratosphere over the North Pacific, contributing to major SSWs and modulating the 10–14-day surface temperature response in North America (Garfinkel et al. 2012; Green and Furtado 2019). Thus, exploring the impact of both tropical convection patterns and extratropical

stratospheric circulation variability on CAOs in S2S forecasts is an important avenue of research for improved subseasonal predictions.

While Arctic stratospheric variability and MJO activity are known sources of subseasonal prediction skill for North American winter weather, their individual contributions, along with joint nonlinear interactions, remain unclear. Prior studies using reanalysis or S2S model output reflect the combined influence of these modes, making it difficult to isolate individual effects. This study addresses that gap by performing targeted nudging experiments with a subseasonal forecast model for two central U.S. CAO case studies. Nudging modifies model physics tendencies, usually for prognostic variables (i.e., zonal wind, meridional wind, and temperature), to steer the model toward a prescribed target state, often reanalysis, allowing the assessment of how the model responds to specific sources of variability.

Nudging has been used to isolate stratospheric impacts in forecasts of European cold events following major SSWs, as well as anomalous cold responses across North America (e.g., Kautz et al. 2020; Huang et al. 2022; Ding et al. 2023). It has also been applied to tropical circulation (e.g., Huang et al. 2023), with studies showing that improved tropical dynamics reduce forecast errors in the stratosphere and MJO teleconnections on subseasonal time scales (e.g., Karpechko et al. 2024). Here, we conduct experiments that separately nudge the stratosphere and tropical atmosphere to assess their relative roles in shaping subseasonal predictions of two central U.S. CAO events. This study is structured as follows. Section 2 outlines the data and experimental design. Sections 3 and 4 present the results for both case studies. In section 5, we discuss implications for improving subseasonal CAO forecasts and broader applications to other extreme weather events.

2. Data and methods

a. Data

To analyze the evolution of the two central U.S. CAOs, we use daily averaged data from the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), reanalysis (Gelaro et al. 2017), spanning 1 January 1981–31 December 2022. MERRA-2 has a $0.5^\circ \times 0.625^\circ$ horizontal resolution and 42 pressure levels from 1000 to 0.1 hPa. We choose MERRA-2 as this is the reanalysis product used in the nudging input files of the subseasonal forecast model used in the study (see section 2b). We analyze key variables relevant to CAO evolution: 2-m temperature (T2M), GPH, meridional wind (v), zonal wind (u), and air temperature (T). For MERRA-2 anomalies, we use the 1999–2020 daily climatology where we remove the seasonal cycle from each day, aligning with the model hindcast period we use. However, for the regime definitions (described later), we use the full 1981–2022 period and its corresponding climatology to ensure a longer and more robust baseline for defining the regimes. We also compute the reflection index (RI; e.g., Matthias and Kretschmer 2020; Messori et al. 2022; Millin et al. 2022) to evaluate the role of stratospheric wave reflection in each event. Following Messori et al. (2022), the RI is defined as the difference between standardized anomalies of the area-averaged 100-hPa

meridional eddy heat flux ($v'T'$) over Siberia (45° – 75° N, 140° – 200° E) and Canada (45° – 75° N, 230° – 280° E):

$$RI = (v'T')_{\text{Sib}}^* - (v'T')_{\text{Can}}^*, \quad (1)$$

where asterisks indicate standardized quantities, primes denote deviations from the zonal mean, and the subscripts “Sib” and “Can” refer to Siberia and Canada, respectively. A positive RI suggests an anomalously higher meridional eddy heat flux over Siberia than Canada, suggestive of stratospheric wave reflection, especially when the Siberian component is positive (upward wave propagation) and the Canadian component is negative (downward wave propagation). We also use MJO phase and amplitude data from the Australian Bureau of Meteorology (BoM 2025) to assess tropical convective activity prior to the events.

b. Model description and experiments

Nudging in S2S forecasts plays a key role in evaluating the influence of modes of variability on surface predictability, extreme event attribution, and stratosphere–troposphere coupling—central themes in the Stratospheric Nudging And Predictable Surface Impacts (SNAPSI) protocol (Hitchcock et al. 2022). We use the Community Earth System Model, version 2 (CESM2), subseasonal prediction system (Richter et al. 2022), with the Community Atmosphere Model, version 6 (CAM6), as the atmospheric component, to conduct the targeted nudging experiments. CAM6 runs at approximately 1° horizontal resolution with 32 vertical levels from the surface to 2 hPa (~ 40 km). CESM2 CAM6 subseasonal forecasts extend to 45 days with 21 ensemble members and can only be initialized weekly on Mondays. The model is selected for its skill in forecasting variables such as T2M, the MJO, and the North Atlantic Oscillation (NAO) and for its ability to apply region-specific nudging (Richter et al. 2022). Additionally, CESM2 CAM6 represents elements of stratosphere–troposphere dynamical coupling well enough for our study as compared to our subseasonal forecasting systems, including the lagged downward propagation of polar cap (60° – 90° N) GPH anomalies following SSWs and their persistence in the troposphere (Richter et al. 2022; Garfinkel et al. 2025). To compute anomalies in our forecast experiments, we use CESM2 CAM6’s 11-member hindcasts (1999–2020) to build a lead time–dependent and initialization-dependent climatology. As the hindcasts are also initialized weekly on Mondays, not all calendar days are represented each year. To address this, we adapt the method of Pegion et al. (2019): We first average across ensemble members and initializations for each calendar day and then apply a 31-day rolling triangular smoother (± 15 days).

To isolate the roles of stratospheric versus tropical atmospheric variability in central U.S. CAOs, we conduct three experiments that nudge T , u , and v toward MERRA-2 reanalysis over specific domains. The first, NUDGE-S, applies full nudging at and above ~ 100 hPa across all longitudes and latitudes, tapering to zero near 200 hPa using a hyperbolic transition centered at ~ 150 hPa. This setup, modified slightly from previous studies, ensures complete lower-stratospheric nudging to capture mechanisms like wave reflection and downward

propagation from SSWs. The second experiment, NUDGE-T, nudges the full atmospheric column between 10° S and 10° N across all longitudes, with a hyperbolic transition tapering to zero at around 30° S/N. No nudging occurs poleward of these latitudes. These bounds are similar to those used in previous studies (e.g., Jung et al. 2010; Dias et al. 2021; Karpechko et al. 2024; Vitart and Balmaseda 2024). Our third experiment, FREE, is a control simulation with no nudging, allowing the atmosphere to evolve freely for baseline comparison.

Each nudging experiment uses 3-hourly MERRA-2 input files (UCAR 2024) for the 32-level CAM6 configuration, with nudging applied on a 12-h time scale (nudging coefficient of 0.25). The model linearly interpolates target values to the current model time, with a constant time scale based on the target interval. By comparing these experiments, we assess how strongly the stratosphere or tropical circulation influences subseasonal forecasts of central U.S. CAOs. We evaluate nudging impacts through ensemble-mean differences in T2M and 500-hPa GPH relative to the FREE experiment. Statistical significance is assessed using 5000-iteration bootstrapping. Additionally, we interpret ensemble-mean 500-hPa GPH responses in terms of North American winter weather regimes.

c. North American winter weather regimes

To assess circulation differences during CAOs across forecast experiments, we examine the evolution of the North American winter weather regimes (Millin et al. 2022). This framework is first applied to MERRA-2 reanalysis from 1981 to 2022. We use linearly detrended 500-hPa GPH anomalies over 20° – 80° N and 180° – 330° E during November–March, removing the seasonal cycle for each day relative to the 1981–2022 climatology. Anomalies are weighted by the square root of the cosine of latitude, and empirical orthogonal function (EOF) analysis is performed. The 12 leading principal components (PCs), capturing approximately 80% of the variance, are retained. We then apply k -means clustering to these PCs to identify five winter weather regimes, assigning each day to the closest centroid using Euclidean distance via Python’s “scikit-learn” package (Pedragosa et al. 2011). To evaluate the regime strength in MERRA-2 and CESM2 CAM6 forecasts, we compute the weather regime index (WRI) by projecting daily 500-hPa GPH anomalies onto the cluster-mean patterns (e.g., Michel and Rivière 2011) derived from MERRA-2, then standardizing the projections using the respective climatologies (1981–2022 for the MERRA-2 regime definition, 1999–2020 for CESM2 CAM6). We choose this approach rather than deriving regime patterns directly from the model for the model WRIs to avoid potential model biases in regime identification and ensure that we are comparing the same characteristic 500-hPa GPH patterns for the regimes between the verification (MERRA-2) and the forecast in which we are interested.

d. Event choices

To investigate the impact of stratospheric and tropical influences on the evolution of central U.S. CAOs and their associated circulation patterns, we select two events from the Millin et al. (2022) Great Plains CAO database. Our first event is the central

U.S. CAO from 26 December 2017 to 2 January 2018, characterized by stratospheric wave reflection (Matthias and Kretschmer 2020) and an active MJO phase, occurring during an AkR weather regime (Millin et al. 2022). The second event is a central U.S. ArH-type CAO from 27 to 31 January 2004, which was preceded by a major SSW on 5 January 2004 (Butler et al. 2017), and similarly active MJO conditions. These two cases offer a unique opportunity to explore the role of stratospheric representation in subseasonal forecasting of central U.S. CAOs, each with distinct stratospheric mechanisms but influenced by concurrent tropical convection. For both of our chosen CAOs, we perform a FREE, a NUDGE-S, and a NUDGE-T experiment weekly for initializations approximately 4–1 week prior to the CAO start date (i.e., 12 total experiments per CAO). For the 2004 CAO, the three experiments are performed on 29 December 2003, 5 January 2004, 12 January 2004, and 19 January 2004. For the 2017 CAO, the three experiments are performed on 27 November 2017, 4 December 2017, 11 December 2017, and 18 December 2017.

3. Evolutions of the two central U.S. CAO events

a. Evolution of the extratropical circulation pattern

We begin by analyzing the synoptic evolution of the selected CAOs. Figure 1 shows T2M and 500-hPa GPH anomalies across the NH for both cases in MERRA-2. The December 2017/18 CAO featured a strong anomalous ridge over Alaska ($\sim +30$ dam) leading up to and during the first 3 days of the CAO, coinciding with anomalous troughing over Canada/the northern United States and widespread negative T2M anomalies extending into the central and eastern United States (Figs. 1a,b). From 29 to 31 December, the anomalous ridge over the U.S. West Coast intensified, shifting colder air southeastward, while the downstream trough extended into the North Atlantic (Fig. 1c). Between 1 and 6 January 2018, ridging expanded along the U.S. West Coast with downstream troughing across eastern North America, producing above-normal T2M in the west and below-normal T2M in the east (Figs. 1d,e). Although the CAO mainly spanned 26 December–2 January in the central United States, strong negative T2M anomalies persisted beyond this window.

The synoptic evolution of the 27–31 January 2004 central U.S. CAO contrasts markedly from the December 2017/January 2018 event (Figs. 1f–j). Prior to onset, a strong anomalous ridge (>20 dam) developed over southern Greenland, flanked by anomalous troughing over the North Atlantic (Fig. 1f). Simultaneously, anomalous ridging over the East Siberian Sea and troughing from Canada into the central United States produced negative T2M anomalies across northwestern and northeastern North America. During the event's first 3 days, cold anomalies progressed from northwest Canada into the northern U.S. Great Plains (Fig. 1g). The Greenland/northern Canada ridge persisted through 4 February, with a deep trough to the south forming a pronounced negative NAO-like pattern from 27 January to 1 February 2004 (Figs. 1g,h). As the CAO subsided, cold anomalies lingered but weakened (Figs. 1h–j). Together, the two events highlight key differences in anomaly duration and

dominant high-latitude blocking: Alaskan ridging in 2017/18 versus Greenland/northern Canada ridging in 2004. While both blocking regimes can drive CAOs in the central and eastern United States, they reflect distinct links to stratospheric variability (Lee et al. 2019; Millin et al. 2022).

b. Evolution of the stratospheric and tropical circulation patterns

We now examine the stratospheric and tropical forcings associated with the two CAO events. Figure 2 shows the lagged evolution of polar cap (60° – 90° N) averaged u standardized anomalies, RI, and MJO phase and amplitude for the two CAOs. About 1 month before the December 2017/January 2018 CAO, upper-stratospheric u anomalies were positive, indicating a stronger-than-normal SPV (Fig. 2a). Three–four weeks prior, negative u anomalies developed in the troposphere, followed by mid- to upper-stratospheric weakening around 6 December. These anomalies persisted through the CAO onset before returning to a stronger vortex state. Simultaneously, the RI increased from $+3\sigma$ to near $+5\sigma$ (Fig. 2b, black line), coinciding with strong Siberian positive $v'T'$ pulses and subsequent negative Canadian $v'T'$ pulses at 100 hPa, indicative of stratospheric wave reflection. Two active MJO periods, in phases 6–8, occurred concurrently with or slightly lagged the Siberian heat flux pulses. These phases are known to enhance North Pacific and Atlantic blocking (Fig. 2c, Henderson et al. 2016; Henderson and Maloney 2018). Together, the RI and MJO timing support the idea that tropical convection helps drive the extratropical wave pattern associated with AkR-type CAOs in the central United States (Millin et al. 2022).

The stratospheric and tropical evolution for the January 2004 CAO differs notably from that of December 2017/18. The 2004 event followed a major SSW on 5 January, marked by strong negative polar cap u anomalies in the mid- to upper stratosphere (Fig. 2d). These anomalies gradually descended, producing two surface “drips,” one around 15 January and another near CAO onset on 27 January, coinciding with strong Greenland/northern Canada ridging (Fig. 1), suggesting the SSW may have influenced this CAO. The RI reached a moderate peak ($+3\sigma$, Fig. 2e) from a brief $v'T'$ pulse over Siberia on 29 December (Fig. 2e, red dashed line), likely weakening the SPV. However, the absence of a corresponding negative Canadian heat flux response suggests wave reflection did not occur. Unlike the 2017/18 case, stratospheric wave reflection appears unimportant during this event. Still, the January 2004 CAO was preceded by a highly active MJO (amplitude >2), in phases 6–8 before shifting into phases 1–3 about a week prior to the event (Fig. 2f).

4. Model experiment results

Our analyses of the evolution of the atmospheric circulation in each CAO event reveal that each event features different tropospheric weather regimes, stratospheric mechanisms, and tropical–extratropical interactions. We next examine the role of the stratosphere and tropics in subseasonal forecasts of these CAOs through our nudging experiments to investigate their impacts collectively and in isolation.

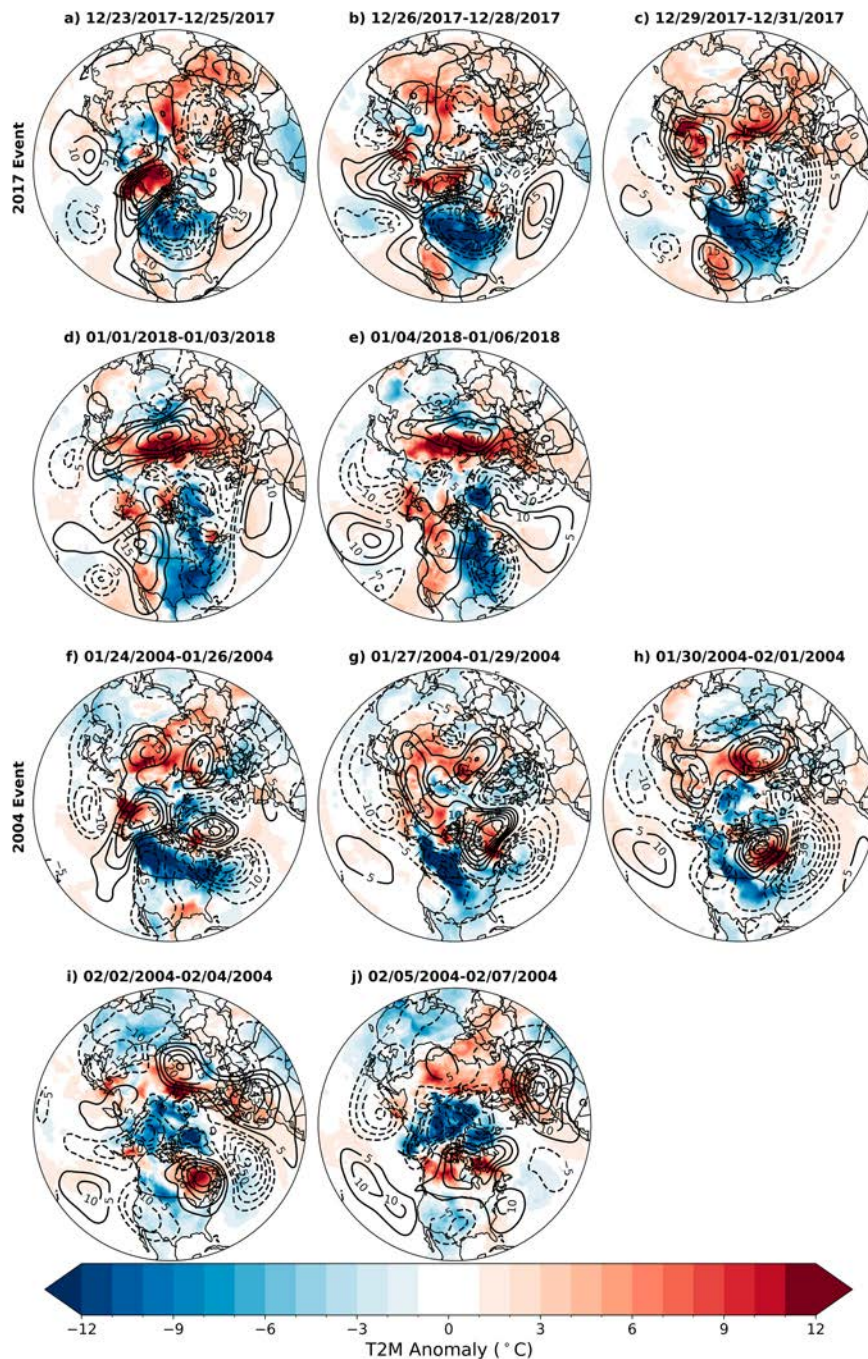


FIG. 1. MERRA-2 T2M anomalies ($^{\circ}\text{C}$; shading) and 500-hPa GPH anomalies (dam; contours) averaged for 3-day periods from (a) 23–25 Dec 2017 through (e) 4–6 Jan 2018 for the December 2017 central U.S. CAO event and (f) 24–26 Jan 2004 through (j) 5–7 Feb 2004 for the January 2004 central U.S. CAO event.

a. December 2017/January 2018 CAO analysis

1) FORECASTS OF THE REFLECTION INDEX

For the December 2017/January 2018 CAO, we first examine the RI in our forecast experiments. Figure 3 illustrates the forecast spread of the RI and its two components for the

target period 14–25 December 2017. This target period corresponds to the period of peak RI and the strongest downward wave response prior to the CAO (Fig. 2b). We show the results for three subseasonal CAO forecast lead times for each experiment (FREE, NUDGE-T, and NUDGE-S). For the 27 November 2017 initialization (~4 weeks before the CAO

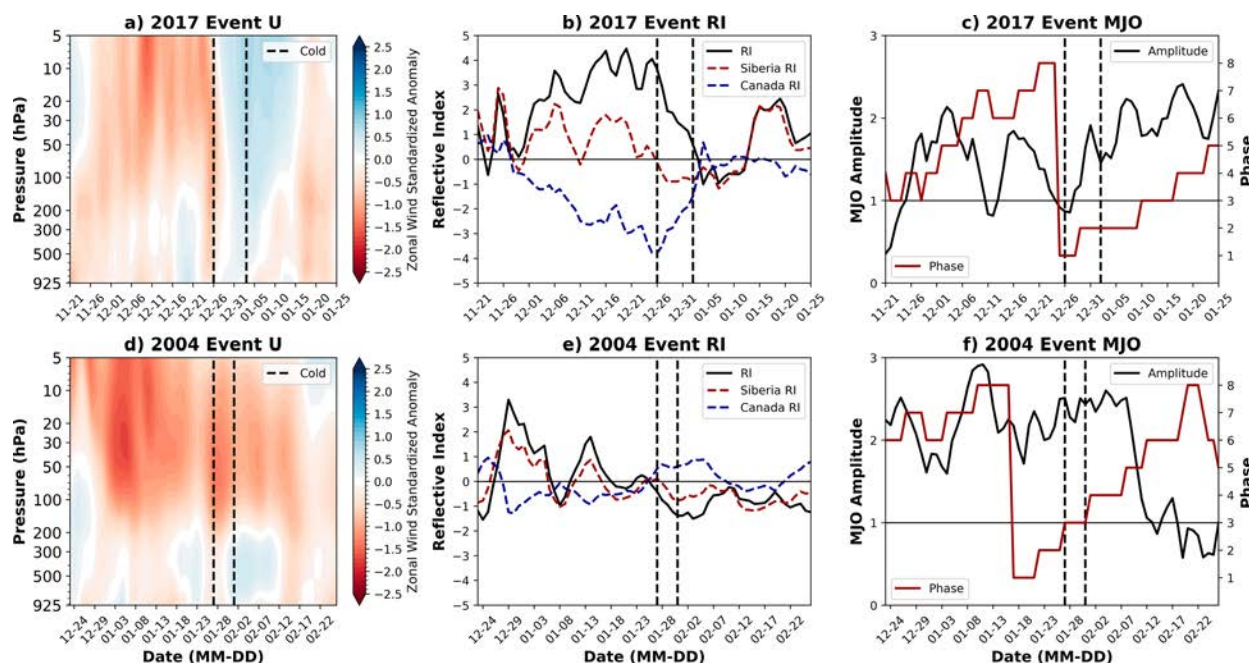


FIG. 2. Evolution of several key S2S indices for the December 2017 central U.S. CAO event: (a) polar cap (60° – 90° N) averaged u standardized anomalies throughout the atmosphere (shading) from MERRA-2, (b) the RI (black line) and its Siberian (red dashed line) and Canadian components (blue dashed line) from MERRA-2, and (c) the MJO amplitude (black solid line) and phase (red solid line) from the BoM. (d)–(f) As in (a)–(c), respectively, but for the January 2004 central U.S. CAO. Black dashed lines indicate the start and end of the CAO periods. In (a) and (d), positive zonal wind anomalies are in blue and negative anomalies are in red, representing a stronger-than-average polar vortex and weaker-than-average polar vortex, respectively.

onset), the forecast RI in the FREE run is centered around 0 (i.e., no reflection), with the maximum of the distribution reaching about $+3\sigma$, lower than the MERRA-2 observed value of $\sim +4\sigma$ (Fig. 3a, gray). In contrast, the NUDGE-T experiment shifts the forecast RI distribution toward higher values, reaching a maximum closer to $+4\sigma$ (Fig. 3a, red). This finding suggests that, even though the extratropical stratosphere is not directly nudged in the NUDGE-T experiment, the tropical atmosphere contributes to a better forecast of the RI for this event. This contribution may be better representation of MJO-related convection during the event in phases 6–8 (Fig. 2c), which is known to influence upward wave activity flux into the polar stratosphere (e.g., Garfinkel et al. 2014). However, the precise reason for this improvement is unknown. The NUDGE-S experiment performs best for this initialization, showing a much narrower ensemble spread and values close to, and even exceeding, the observed evolution (Fig. 3a, blue). Indeed, in the NUDGE-S experiment, there is a closely similar evolution of the polar cap u standardized anomalies in the stratosphere to that of MERRA-2 (Fig. S1 in the online supplemental material and Fig. 2a). A similar pattern of ensemble spread appears in the 4 December 2017 initialization forecast RI (~ 3 weeks before the CAO, Fig. 3b), though the FREE experiment exhibits an increase in the value of the forecast RI. By the 11 December 2017 initialization (~ 2 weeks before the CAO and 3 days before the RI forecast period), all experiments begin to converge, which is expected as nudging has less impact at shorter lead times to

the RI averaging period, due to the ability of the FREE experiment to better capture the shorter-term variability (Fig. 3c).

Examining the forecasts for the two components of the RI separately reveals similar conclusions to those of the total RI (Figs. 3d–f for the Siberian component; Figs. 3g–i for the Canadian component). Ultimately, across initializations, the NUDGE-S experiment consistently outperforms the other experiments in simulating both components of the RI, with narrower spreads as well. Interestingly, the NUDGE-T experiment exhibits the highest Siberian RI forecast among the experiments in the 27 November 2017 and 4 December 2017 initializations (Figs. 3d,e), indicating a potential strong association between tropical forcing and upward wave fluxes over Siberia. Closer to the event, all experiments converge to the observed values in both components. Thus, given the large experimental differences on the longer subseasonal time scales, the representation of the stratosphere and tropics may have a substantial impact on the subseasonal forecast of reflection-induced cold for this event.

2) TEMPERATURE AND TROPOSPHERIC CIRCULATION FORECAST

Next, we investigate the evolution of the T2M and 500-hPa GPH fields for this CAO event. For these comparisons among the model experiments, we choose to investigate the 27 November 2017 initialization run only. To compare the ensemble-mean tropospheric circulation and T2M forecasts, we use the FREE experiment as a baseline and subtract its

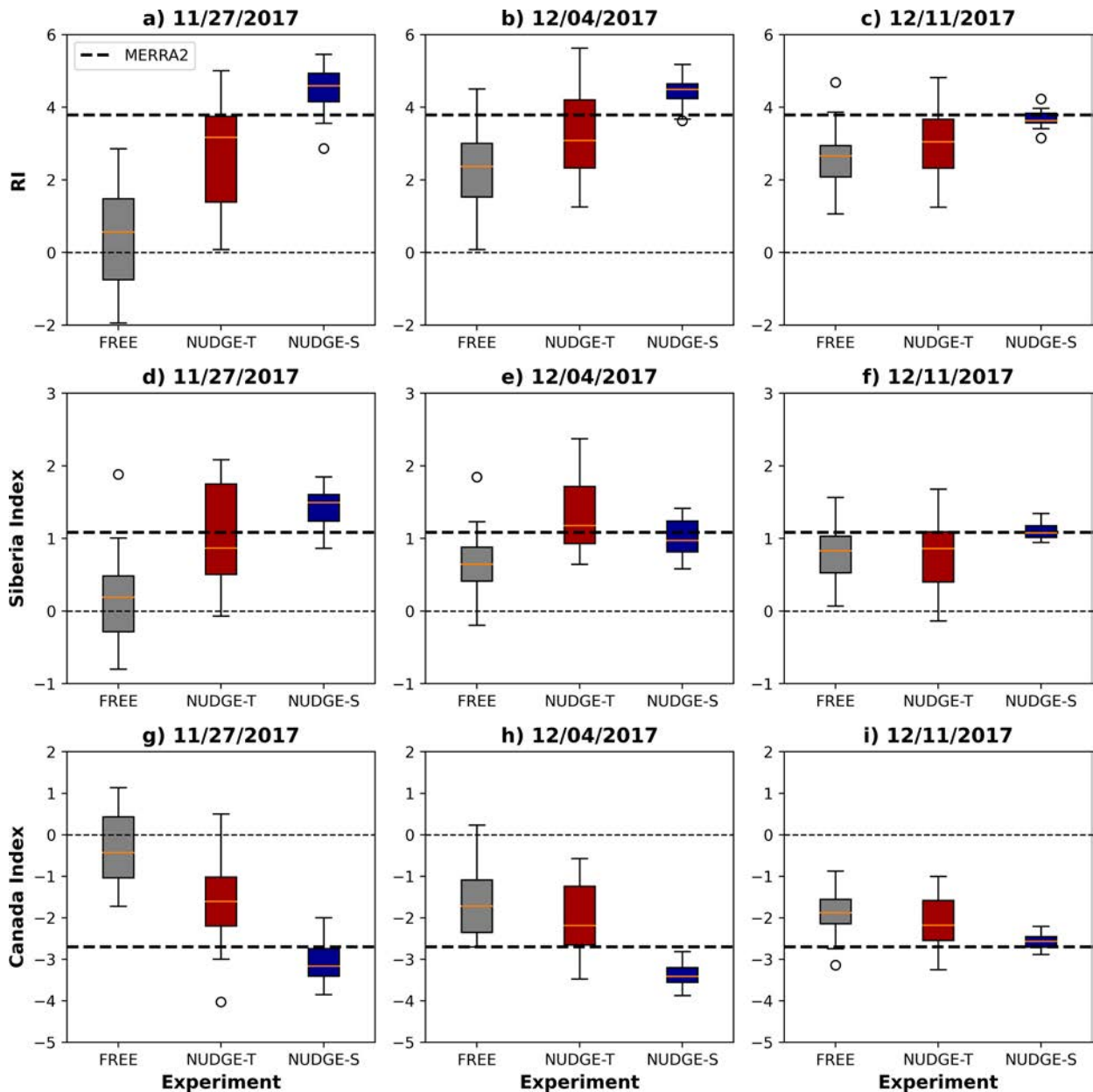


FIG. 3. Boxplots of (a)–(c) the RI, (d)–(f) the Siberian component of the RI, and (g)–(i) the Canadian component of the RI, each averaged between 14 and 25 Dec 2017, for forecast initializations of 27 Nov 2017, 4 Dec 2017, and 11 Dec 2017, respectively. Gray boxes denote the FREE experiments, red boxes denote the NUDGE-T experiments, and blue boxes denote the NUDGE-S experiments. Orange lines represent the ensemble median; the lower and upper bounds of the box indicate the 25th and 75th percentiles, respectively; whiskers extend from the box to the data point lying within 1.5 times the interquartile range from the box; and circles are outliers.

forecast values for the fields from MERRA-2 (to see the errors between the FREE experiment and the observed evolution) and the two nudged experiments forecast values. We focus on forecast fields during the CAO event (26 December 2017–6 January 2018). Figure 4 shows the differences in full-field T2M (shading, °C) and 500-hPa GPH (contours, dam) for MERRA-2, NUDGE-S, and NUDGE-T relative to the FREE experiment. The FREE run largely underestimates 500-hPa GPH across the North Pacific, Alaska, the southwestern United

States, and the Arctic Ocean during the first 3 days of the CAO, where strong ridging is observed (Fig. 1), while overestimating GPH over Canada and the northern United States, where a deep trough is present (Fig. 4a). As a result, the FREE run is not nearly as cold as MERRA-2 across much of Canada, the eastern United States, and the northern United States (Fig. 4a, negative values in shaded contours). From 29 to 31 December 2017, similar differences persist, but the FREE run better captures the GPH across the Arctic Ocean

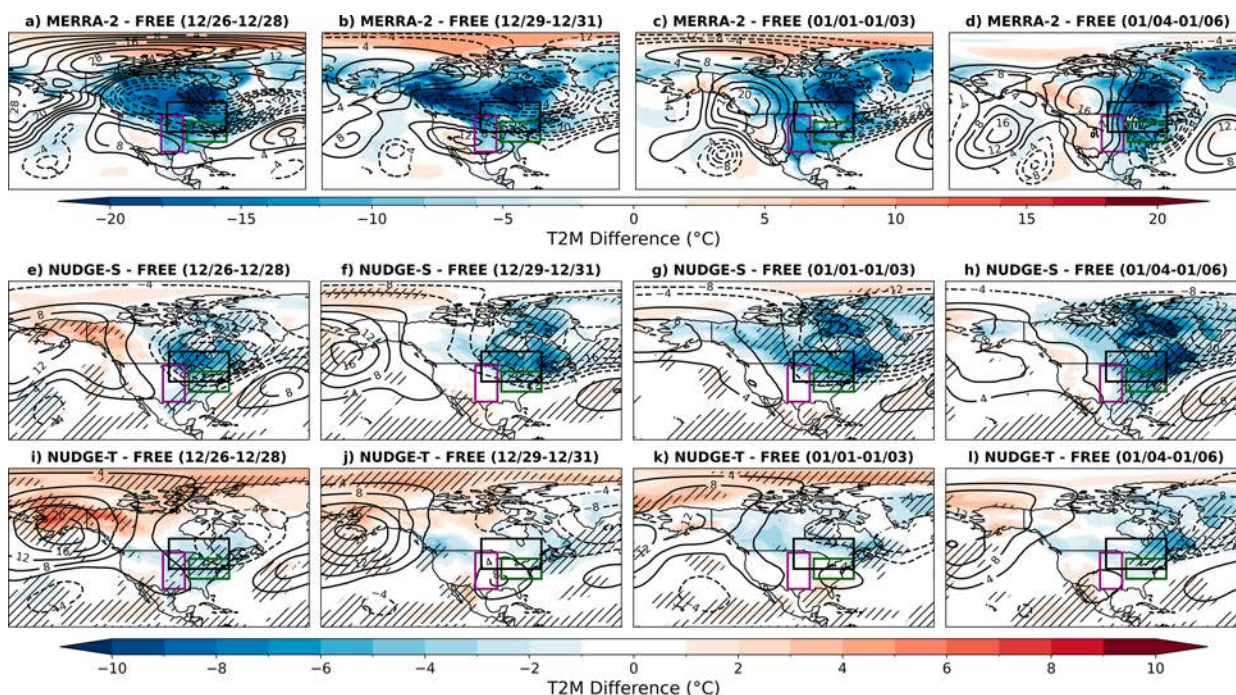


FIG. 4. Difference in T2M ($^{\circ}\text{C}$; shading) and 500-hPa GPH (dam; contours) from the ensemble mean FREE run for (a)–(d) MERRA-2, (e)–(h) the NUDGE-S ensemble mean, and (i)–(l) the NUDGE-T ensemble mean, for the 27 Nov 2017 initialization at 3-day averaged periods from 26–28 Dec 2017 through 4–6 Jan 2018, respectively. Hatching shows statistical significance of the T2M difference ($p < 0.05$) via 5000-iteration bootstrapping with replacement. Black boxes represent the area that usually has the coldest T2M response to reflection events in Messori et al. (2022) (40° – 55°N , 260° – 290°E), purple boxes represent the Great Plains (30° – 48°N , 257° – 268°E), and green boxes represent the Northeast United States (35° – 45°N , 270° – 290°E).

(Fig. 4b). Throughout 1–6 January 2018, the FREE run fails to accurately simulate the west/east dipole in ridging/troughing, respectively. The difference between MERRA-2 and the FREE run resembles the pattern in Fig. 1, with anomalous ridging across the U.S. West Coast and anomalous troughing over eastern North America. This results in lower T2M values in MERRA-2 extending farther south across the central United States and, particularly from 4 to 6 January, the northeastern United States (Figs. 4c,d). The underestimation of extremely high or low 500-hPa GPH or T2M values is also more likely on subseasonal time scales versus reanalysis, as S2S models often struggle to maintain the magnitude of extremes at longer lead times and thus explains why probabilistic forecasts are often used (Vitart and Robertson 2018).

Figures 4e–h show the differences in T2M and 500-hPa GPH between the NUDGE-S and FREE experiments. From 26 to 31 December 2017, the NUDGE-S experiment simulates higher GPH over Alaska and lower GPH over northeast Canada and the northern United States compared to the FREE experiment (Figs. 4e,f). Additionally, the NUDGE-S experiment produces colder T2M over northeastern Canada, the eastern United States, and the northern United States during these periods. Notably, the NUDGE-S run is significantly colder across regions sensitive to extreme cold following a stratospheric reflection event (Messori et al. 2022), including the northeastern United States (black and green boxes in Fig. 4, respectively). The difference pattern closely

resembles that of MERRA-2 (Figs. 4a,b), though of smaller magnitude, suggesting that nudging the stratosphere improves the simulation of tropospheric circulation and near-surface temperature on subseasonal time scales for this CAO event. From 1 to 6 January 2018, the NUDGE-S experiment continues to enhance ridging over Alaska, extending into the U.S. West Coast, similar to the MERRA-2 difference maps (Figs. 4g,h). The NUDGE-S run also produces lower GPH across the eastern half of the United States than the FREE experiment, accompanied by significantly negative T2M differences across Canada and the Northeast United States, which extend southward from 4 to 6 January 2018 (Fig. 4h). Despite the overall improvement in subseasonal forecasts with stratospheric nudging, the U.S. Great Plains shows little temperature response, suggesting that this region is less sensitive to improved stratospheric representation with wave reflection, at least for this particular event (Fig. 4, purple boxes).

The NUDGE-T forecast consistently exhibits a stronger signal for higher GPH over Alaska than both the FREE and NUDGE-S forecasts, suggesting that improved representation of the tropics in the CESM2 subseasonal model may enhance anomalous ridge development over Alaska and the North Pacific (Figs. 4i–l). The stronger ridging in NUDGE-T versus the FREE experiment also supports the RI differences in Siberia (Fig. 3d), where enhanced ridging likely amplifies the upward heat flux response from that region. However, while the Siberian component of the RI aligns with this explanation,

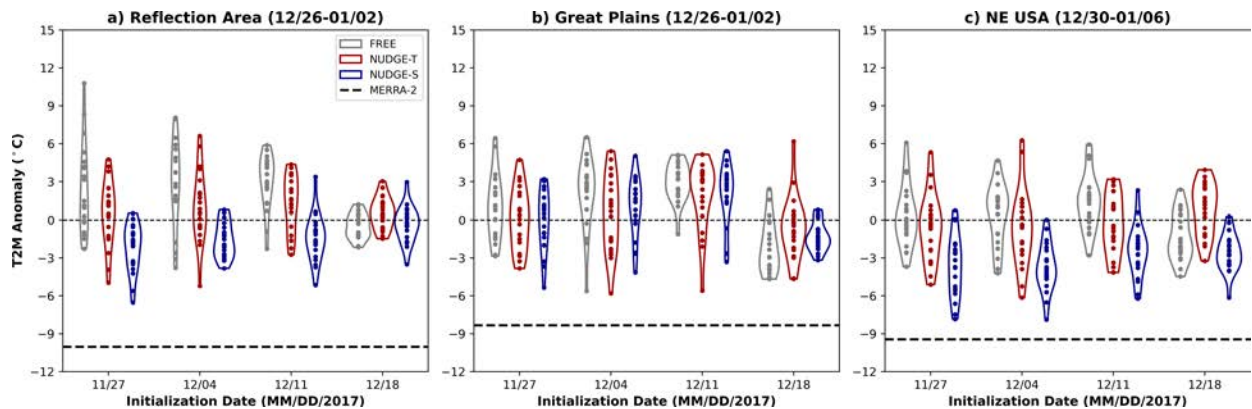


FIG. 5. Violin plots of the ensemble forecast of area-averaged T2M anomalies ($^{\circ}\text{C}$) for weekly CESM2 CAM6 initializations from 27 Nov 2017 to 18 Dec 2017 for (a) the area most prone to cold events due to stratospheric wave reflection ($40^{\circ}\text{--}55^{\circ}\text{N}$, $260^{\circ}\text{--}290^{\circ}\text{E}$) from Messori et al. (2022) averaged between 26 Dec 2017 and 2 Jan 2018, (b) the Great Plains ($30^{\circ}\text{--}48^{\circ}\text{N}$, $257^{\circ}\text{--}268^{\circ}\text{E}$) averaged between 26 Dec 2017 and 2 Jan 2018, and (c) the Northeast United States ($35^{\circ}\text{--}45^{\circ}\text{N}$, $270^{\circ}\text{--}290^{\circ}\text{E}$) averaged between 30 Dec 2017 and 6 Jan 2018. Blue, red, and gray distributions represent the NUDGE-S, NUDGE-T, and FREE experiments, respectively. Dashed black lines represent the MERRA-2 values.

the downward wave response into Canada is weaker than in the NUDGE-S experiment (Fig. 3g). This weaker signal is evident in Figs. 4i–l; although the ridging response in the NUDGE-T experiment is stronger across Alaska than in the other experiments, it fails to capture the magnitude of the anomalous trough over Canada and the northern United States. By extension, improvements to T2M forecasts across the central and eastern United States remain modest, particularly between 26 and 28 December 2017 (Fig. 4i). In fact, between 1 and 3 January 2018, the NUDGE-T experiment even produces higher T2M than the FREE experiment in the eastern United States (Fig. 4k).

Figure 5 presents violin plots of ensemble T2M anomalies for four initialization dates between 27 November and 18 December 2017 for three regions: the wave reflection-sensitive area (26 December 2017–2 January 2018), the Great Plains (26 December 2017–2 January 2018), and the Northeast United States (30 December 2017–6 January 2018). In both the reflection-sensitive area (Fig. 5a) and the Northeast United States (Fig. 5c), the NUDGE-S experiment consistently produces the coldest temperature response, with most of the distribution falling below average, especially in the Northeast United States. While none of the experiments fully capture the magnitude of the MERRA-2 T2M anomalies (Fig. 5, black dashed line), the NUDGE-T run is generally the next best performer, with the FREE run performing the worst, particularly at longer lead times (e.g., 27 November and 4 December 2017 initialization). This ranking is especially pronounced in the reflection-sensitive area, underscoring the influence of subseasonal stratospheric representation on extreme cold temperature forecasts in North America (Fig. 5a). For the 18 December initialization, the stratospheric wave reflection and associated negative heat flux anomalies largely precede the forecast start date (Fig. 2b). Consequently, much of the downward influence of the stratosphere on the tropospheric circulation has already occurred, limiting the ability of stratospheric nudging to modify the subsequent surface temperature evolution at lead times of 8–15 days. As a

result, the nudged experiment closely follows the FREE run, which does not develop a strong cold anomaly, leading to a weaker cold signal compared to earlier initializations (i.e., 27 November) that precede the reflective period. Across all experiments, the largest differences occur for initializations on the subseasonal lead times. Closer to the CAO event, even the FREE run gains skill, reducing the impact of the nudging. As noted earlier, T2M anomalies in the Great Plains remain relatively unaffected by experiment type, suggesting that reflection-induced cold events in this region depend on dynamically more than just improving the representation of the stratosphere (Fig. 5b).

3) NORTH AMERICAN WINTER WEATHER REGIME FORECAST

To further examine differences in tropospheric evolution across our experiments, we analyze the ensemble-mean evolution of five North American winter weather regimes: 1) the ArH regime; 2) the AkR regime; 3) the West Coast Ridge (WCR) regime, featuring ridging along the U.S. West Coast; 4) the Pacific trough (PT) regime, associated with an enhanced Aleutian low; and 5) the Arctic low (ArL) regime, featuring pan-Arctic troughing (see Fig. S2 for the composite maps and daily frequencies of these regimes in MERRA-2). Figure 6 illustrates the ensemble-mean evolution of the WRI for these five regimes in MERRA-2 and the three experiments, initialized on 27 November 2017. In MERRA-2, the period begins in a PT regime (Fig. 6a, green line) before rapidly transitioning to a dominant WCR regime from approximately 7–17 December 2017 (Fig. 6a, blue line), with WRI values increasing to near $+3\sigma$. Concurrently, the ArL regime (Fig. 6a, purple line) weakens, with its WRI falling to -2σ , reflecting its strong spatial anticorrelation with the WCR pattern (Fig. S2). In the 2 weeks preceding the onset of the central U.S. CAO, the WRIs for the AkR (Fig. 6a, orange line) and WCR (Fig. 6a, blue line) regime signals strengthen,

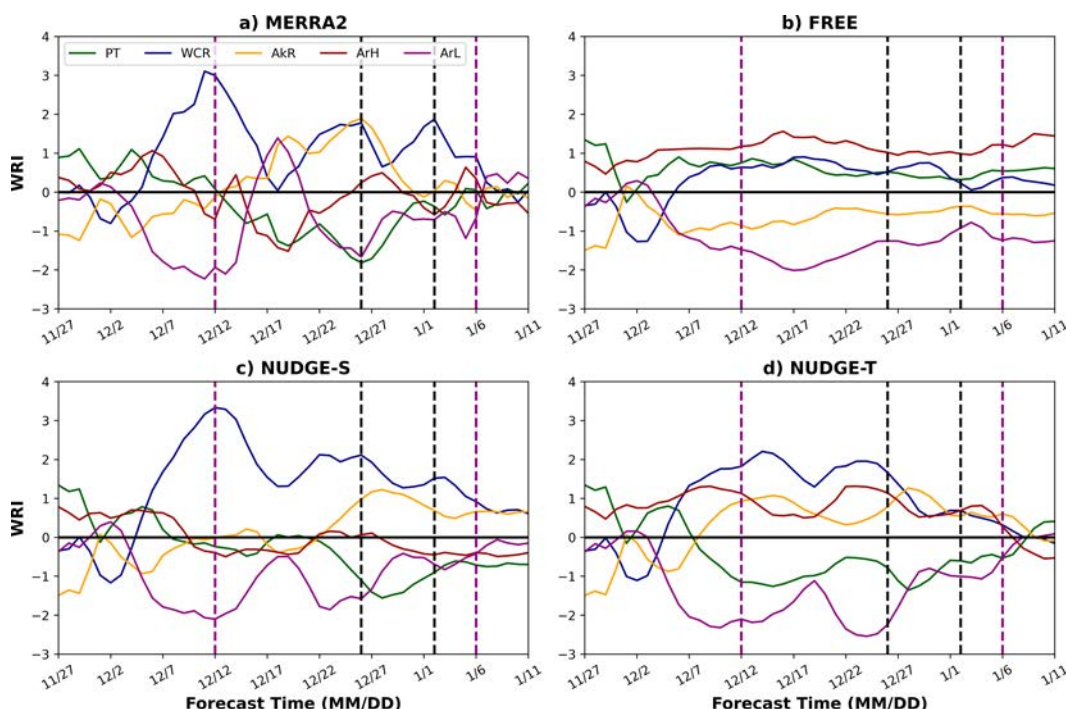


FIG. 6. Evolution of the ensemble mean WRI for the five North American winter weather regimes for the 27 Nov 2017 CESM2 CAM6 initialization in (a) the MERRA-2 evolution, (b) the FREE experiment, (c) the NUDGE-S experiment, and (d) the NUDGE-T experiment. The green line represents the PT, blue line represents the WCR, orange line represents the AkR, red line represents the ArH, and solid purple line represents the ArL. Black dashed lines show the start and end dates of the 2017/18 CAO. Purple dashed lines represent the start and end of the period used for correlation analysis in Fig. S5.

with the AkR regime emerging as the dominant regime at CAO onset. As the CAO progresses, the AkR regime weakens, and the WCR regime once again becomes the dominant pattern, consistent with Fig. 1.

The FREE experiment does not accurately simulate the evolution of winter weather regimes compared to MERRA-2 (Fig. 6b). Throughout most of the forecast, the dominant regime remains the ArH regime (WRI values of $+1\sigma$ to $+2\sigma$; Fig. 6b, red line), followed closely by the WCR (Fig. 6b, blue line) and PT (Fig. 6b, green line) regimes. However, the FREE run fails to establish the WCR regime as one of the dominant regimes during the CAO period (Fig. 6b). By contrast, the NUDGE-S experiment shows a substantial improvement in the WRI forecast, closely resembling the MERRA-2 evolution (Fig. 6c). Although it does not fully capture the reduction and subsequent rebound of the WCR regime before the CAO, it maintains the WCR regime as the dominant regime throughout the event. These results further reinforce our earlier findings; i.e., nudging the stratosphere toward the observed state on subseasonal time scales improves the representation of North American tropospheric circulation, leading to a more accurate CAO forecast. Finally, the NUDGE-T ensemble-mean WRI evolution also shows an increase in the WCR WRI and a decrease in the ArL WRI between 7 and 27 December 2017 (Fig. 6d). However, the experiment shows higher AkR and ArH WRI values compared to the other

experiments until the CAO begins, with the AkR regime becoming the dominant signal around 29 December 2017 (Fig. 6d). This forecast evolution aligns with prior research illustrating that active MJO convection in phases 6 and 7 enhances ridging (and blocking) across the North Pacific (e.g., Garfinkel et al. 2014; Henderson et al. 2016; Henderson and Maloney 2018). Indeed, ensemble-mean maps of the 200-hPa streamfunction anomalies leading up to, and including, the CAO onset show that the NUDGE-T simulation features a stronger wave pattern originating from the central tropical Pacific, aligning with strong ridging in the 500-hPa GPH anomaly field across Alaska and northwest Canada (Figs. S3 and S4).

We also correlate MERRA-2 and ensemble member WRIs from 12 December 2017 to 6 January 2018 (Fig. S5). Stratospheric nudging markedly improves WRI prediction, especially for the dominant WCR regime (mean correlation for all ensemble members: $r = 0.68$ for NUDGE-S; Fig. S5b vs $r = -0.06$ for FREE; Fig. S5a), but tropical nudging provides improvement to a lesser extent (mean correlation for all ensemble members: $r = 0.20$ for NUDGE-T; Fig. S5c). While tropical variability contributes to the large-scale circulation during this event, its influence is likely primarily mediated through the stratosphere, and therefore NUDGE-T does not produce the full evolution. As such, this CAO is best characterized as stratosphere-driven, with accurate representation

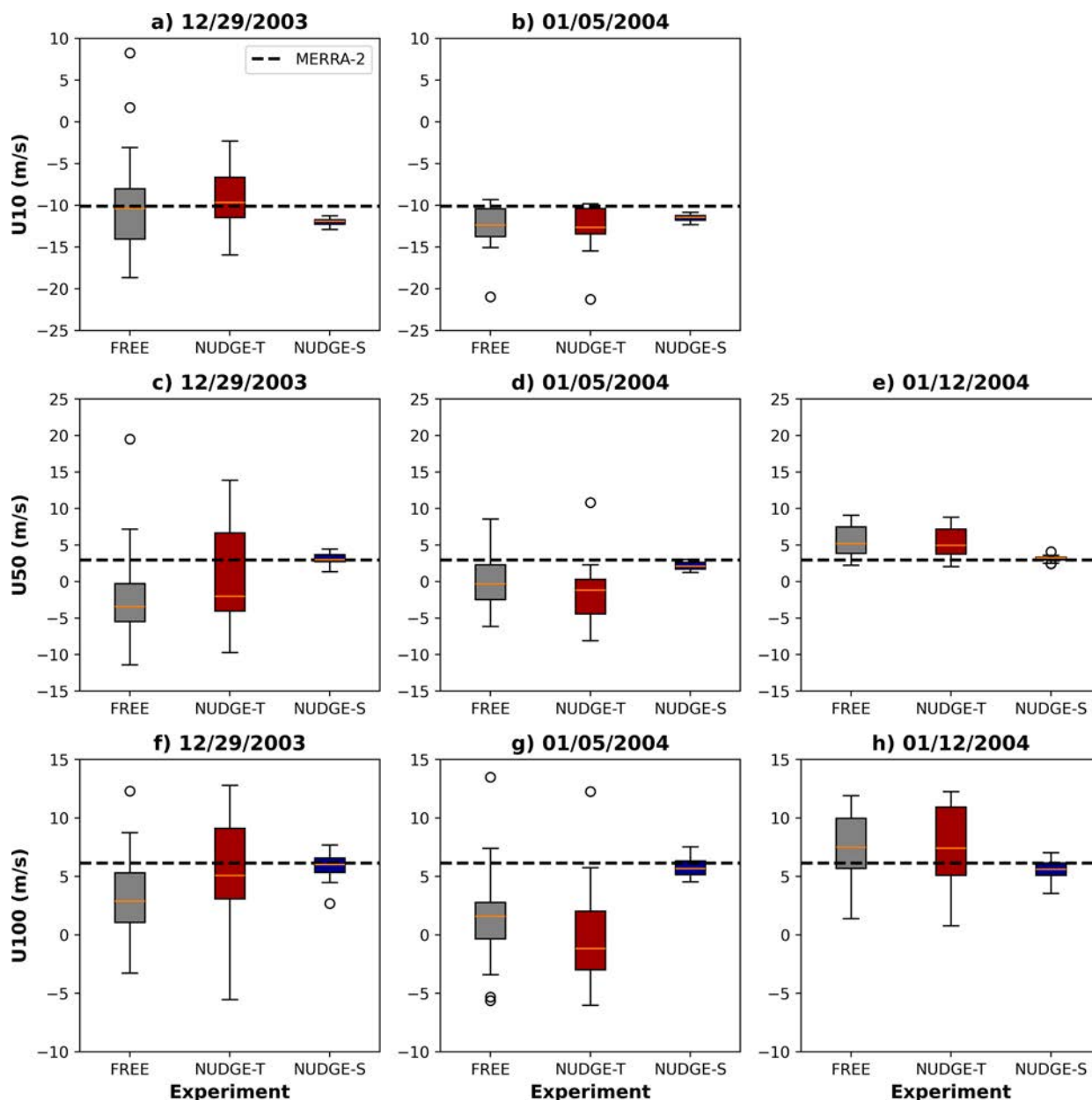


FIG. 7. Boxplots of the 60° – 75° N u at (a),(b) 10 hPa between 5 and 11 Jan 2004, (c)–(e) 50 hPa between 12 and 18 Jan 2004, and (f)–(h) 100 hPa between 19 and 25 Jan 2004, for model initializations of 29 Dec 2003, 5 Jan 2004, and 12 Jan 2004, respectively. Orange lines represent the ensemble median; the lower and upper bounds of the box indicate the 25th and 75th percentiles, respectively; whiskers extend from the box to the data point lying within 1.5 times the interquartile range from the box; and circles are outliers.

of stratospheric dynamics being the dominant requirement for reproducing the observed surface response.

b. January 2004 CAO analysis

1) POLAR CAP U FORECASTS

We next assess the relative roles of tropical and stratospheric circulations on forecasts of the January 2004 CAO. Since this event occurred following a major SSW on 5 January 2004, Fig. 7 presents the ensemble-mean forecast evolution of

the 60° – 75° N-averaged u at 10 hPa ($U10$) between 5 and 11 January 2004 (Figs. 7a,b), 50 hPa ($U50$) between 12 and 18 January 2004 (Figs. 7c–e), and 100 hPa ($U100$) between 19 and 25 January 2004 (Figs. 7f–h). We average u across 60° – 75° N to better capture jet stream variability, and the staggered averaging windows represent the downward propagation of the SSW in Fig. 2d. For $U10$, the NUDGE-S experiment maintains a narrow distribution centered near the observed MERRA-2 values ($\sim -10 \text{ m s}^{-1}$, dashed line) across the 29 December 2003 and 5 January 2004 initialization

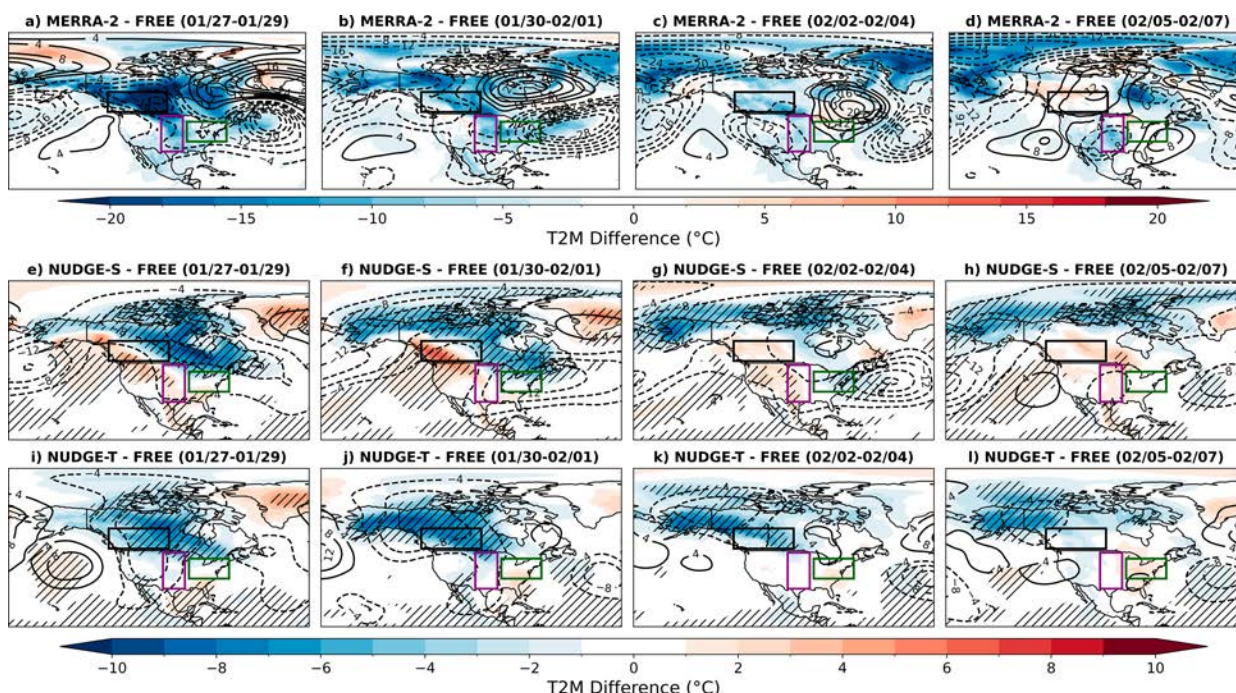


FIG. 8. Difference in T2M ($^{\circ}\text{C}$; shading) and 500-hPa GPH (dam; contours) from the ensemble mean FREE run for (a)–(d) MERRA-2, (e)–(h) the NUDGE-S ensemble mean, and (i)–(l) the NUDGE-T ensemble mean, for the 29 Dec 2003 initialization at 3-day averaged periods from 27–29 Jan 2004 through 5–7 Feb 2004, respectively. Hatching shows statistical significance at the 95th percentile for the difference in the T2M via 5000-iteration bootstrapping with replacement. Black boxes represent western Canada (50° – 60°N , 230° – 260°E); purple and green boxes are the same as in Fig. 4.

dates, consistent with the experimental design (Figs. 7a,b). However, for the 29 December 2003 initialization (prior to the SSW), the NUDGE-T and FREE experiments exhibit greater uncertainty in the forecast of the winds with the NUDGE-T experiment showing the highest variability, particularly for U50 and U100 (Figs. 7a,c,f, red). On the 5 January 2004 initialization (the central date of the major SSW), the FREE and NUDGE-T ensembles predominantly predict weaker-than-observed zonal winds which is especially evident in the lower stratosphere for U50 and U100, whereas the NUDGE-S experiment remains constrained to the observed evolution (Figs. 7b,d,g). These differences suggest that, without a nudged extratropical stratosphere, the FREE and NUDGE-T experiments tend to push the, especially lower, stratospheric u toward stronger easterlies/weaker westerlies, highlighting the uncertainty in subseasonal stratospheric evolution. By the 12 January 2004 initialization (Figs. 7e,h), both the NUDGE-T and FREE experiments converge more toward the observed value, displaying similar ensemble distributions as forecast lead time decreases and skill improves.

2) TEMPERATURE AND TROPOSPHERIC CIRCULATION FORECAST

Given the larger ensemble spread in the forecasts for U10, U50, and U100 in the NUDGE-T and FREE experiments for the 29 December 2003 initialization and that this initialization is prior to the major SSW, we use this initialization to explore

differences in the tropospheric evolution across the experiments when forecasting the January 2004 CAO event. Figure 8 is similar to Fig. 4 but uses the 29 December 2003 initialization to evaluate the forecasts for 27 January–7 February 2004. During the first 3 days of the CAO event (27–29 January 2004), the FREE run substantially underestimates the magnitude of the GPH across Greenland, northern Canada, and the Arctic Ocean compared to MERRA-2, which are all areas of anomalous ridging in the true evolution (cf. Figs. 1g and 8a). Additionally, the FREE run does not capture the magnitude of the troughing south of Greenland/northern Canada and near the Aleutian Islands, leading to a warmer bias across western Canada (Fig. 8, black box), the northern United States, and the Great Plains (Fig. 8, purple box). Between 30 January and 1 February 2004, many of the same errors persist (Fig. 8b), resulting in an overall warm bias across Canada and the central/eastern United States compared to MERRA-2. Although the CAO ended on 2 February 2004, the FREE run continues to underestimate the negative T2M anomalies in the central United States (Figs. 8c,d). Additionally, the FREE experiment struggles to simulate both the anomalous ridging over northeast Canada and the Rossby wave train that develops across the United States, which sustains more modest cold across the central United States after the end of the CAO by making the GPH pattern less progressive.

The NUDGE-S experiment features higher GPH than the FREE run across Greenland between 27 January 2004 and 1 February 2004, highlighting the ability of the model to simulate

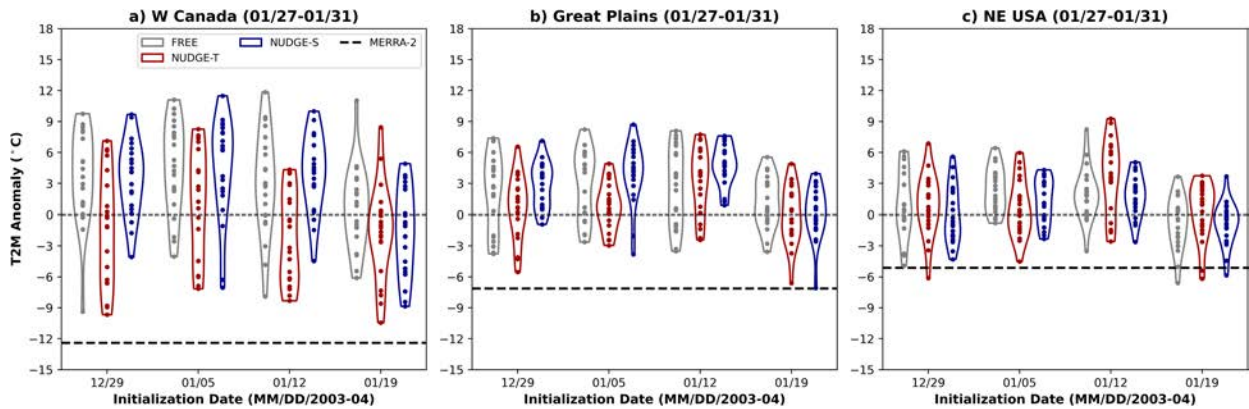


FIG. 9. Violin plots of the ensemble forecast of area-averaged T2M anomalies ($^{\circ}\text{C}$) between 27 and 31 Jan 2004 for weekly CESM2 CAM6 initializations from 29 Dec 2003 to 19 Jan 2004 for (a) western Canada (50° – 60°N , 230° – 260°E), (b) the Great Plains (30° – 48°N , 257° – 268°E), and (c) the Northeast United States (35° – 45°N , 270° – 290°E). Blue, red, and gray distributions represent the NUDGE-S, NUDGE-T, and FREE experiments, respectively; dashed black lines represent the MERRA-2 values.

riding in the region that often experiences anomalous ridging after stratospheric weakenings (Figs. 8e,f, Lee et al. 2022; Millin et al. 2022). Over the same period, the NUDGE-S experiment better represents the lower GPH across the East Coast of the United States than the FREE run. Interestingly, the NUDGE-S experiment shows significantly lower T2M across much of the eastern half of Canada, associated with lower GPH there than in the FREE run. However, the NUDGE-S experiment produces significantly higher T2M than the FREE run across western Canada and the Great Plains (Figs. 1g,h and 8e,f). Between 2 and 7 February 2004, higher T2M persists across western Canada and the Great Plains in the NUDGE-S experiment, with negative T2M departures weakening across much of the rest of Canada, and modest negative T2M departures from the FREE run in the Northeast United States (Figs. 8g,h). Thus, by nudging the stratosphere, the forecast improves somewhat across the eastern half of Canada but not much improvement across western Canada and the central United States later.

By contrast, the NUDGE-T experiment consistently produces lower T2M than the FREE run across much of western and central Canada and into the northern United States throughout the entire period (Figs. 8i–l). The difference is most notable across western Canada throughout 27 January 2004–1 February 2004, where the NUDGE-T experiment produces significantly lower T2M across this key region where the extremely negative T2M anomalies were in the observations (Fig. 1). Indeed, the NUDGE-T run consistently exhibits higher 500-hPa GPH than the FREE run to the west of Alaska and just off the U.S. West Coast (Figs. 8i–l). This area of higher 500-hPa GPH aligns with that shown in the difference between MERRA-2 and the FREE run, where anomalous ridging develops in the observed evolution (Fig. 1). The simulation of more ridging in the North Pacific likely makes the overall tropospheric pattern less progressive and leads to the better simulation of lower T2M across Canada and the northern United States, despite the relatively weaker simulation of strong ridging across Greenland and northern Canada versus the NUDGE-S experiment. Regardless, the simulation

of lower T2M in the Great Plains and Northeast United States is weak throughout, although not with positive T2M departures as in the NUDGE-S experiment (Figs. 8i–l). Thus, the NUDGE-T experiment tends to perform better for this case study for Canada and the northern United States, despite the NUDGE-S simulation showing the most accurate polar cap u evolution in the stratosphere following the major SSW (Fig. S6).

Figure 9 shows violin plots of ensemble T2M anomalies averaged between 27 and 31 January 2004. For the 29 December 2003 initialization in western Canada, the NUDGE-T ensemble distribution of T2M anomalies shows a clear shift toward more negative anomalies when compared to the NUDGE-S and FREE run, with much of the NUDGE-S experiment ensemble above 3°C (Fig. 9a), aligning with Fig. 8. Furthermore, the NUDGE-T experiment has some of the ensemble members with the most negative T2M anomalies in the Great Plains, with the NUDGE-S again showing the warmest ensemble forecast (Fig. 9b). In the Northeast United States, there is very little difference between experiments during the CAO (Fig. 9c). Similar results hold for the 5 and 12 January 2004 initializations across all regions, before the experiments converge toward more negative T2M anomalies on the 19 January 2004 initialization (Fig. 9). Overall, the difference between experiments for this event is modest in these three regions versus the 2017/18 CAO event, highlighting the complex relationship between SSWs, tropical variability, and CAOs in North America.

3) NORTH AMERICAN WINTER WEATHER REGIME FORECAST

Figure 10 shows the North American winter WRIs for the three experiments and in MERRA-2 for the 29 December 2003 initialization. The MERRA-2 evolution of the WRI shows a dominant AkR regime during the first 10 days of the forecast before transitioning to a WCR between 6 and 13 January 2004. The North American weather pattern then transitions to a dominant ArH regime around 18 January 2004 (Fig. 10a). During the first 2 weeks of the forecast, all

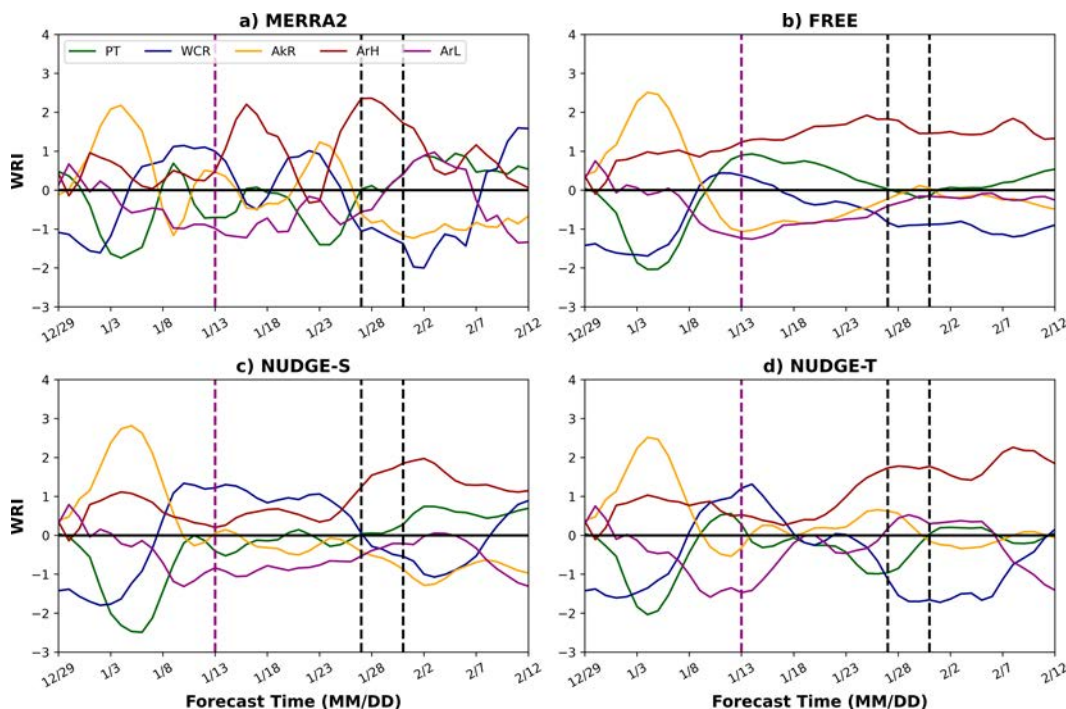


FIG. 10. As in Fig. 6, but for the 29 Dec 2003 CESM2 CAM6 initialization. Black dashed lines indicate the start and end dates of the 2004 CAO event. The purple dashed line represents the start of the period used for correlation analysis in Fig. S7.

experiments produce a WRI evolution very similar to reanalysis, with a dominant AkR regime at the beginning of January (Figs. 10b–d). After the first 2 weeks, the FREE run maintains a dominant ArH regime (Fig. 10b, red line), while the NUDGE-S experiment better captures the emergence of the WCR regime before the ArH regime emerges in the days leading to the CAO and during the CAO (Fig. 10c). The NUDGE-T experiment also captures the initial emergence of the WCR regime around 13 January 2004 (Fig. 10d), and then the ArH regime emerges as the dominant regime 7–10 days prior to the CAO. Therefore, even though each experiment was unable to capture the extreme magnitude of the observed 500-hPa GPH and T2M on longer subseasonal time scales, the model forecasts somewhat accurately capture the regime evolution for this CAO event. Interestingly, the NUDGE-T experiment is the only experiment that captures the increased AkR WRI before the CAO start, likely corresponding to the higher GPH across the west of Alaska and the North Pacific relative to the FREE run (Figs. 8j–m). Correlations between ensemble member forecast WRIs and MERRA-2 during the 2 weeks prior to and through the 2004 CAO show mixed skill in the FREE run: strong for ArL and WCR but weaker for ArH and AkR (Fig. S7). Stratospheric nudging markedly improves WRI correlations across the board; for example, ensemble-mean correlation coefficients for the ArH WRI go from $r = -0.02$ in the FREE run (Fig. S7a) to $r = 0.46$ for the NUDGE-S run (Fig. S7b). Tropical nudging has more mixed results, with the best improvements seen for the WCR WRI ($r = 0.33$ for the FREE run; Fig. S7a vs $r = 0.60$ for the NUDGE-T run; Fig. S7c).

5. Discussion and conclusions

Wintertime CAOs in the central United States are extreme events that have significant socioeconomic and environmental impacts that often also affect neighboring highly populated regions, such as the eastern United States. Past works have established significant statistical relationships between modes of climate variability and North American CAOs (e.g., Lee et al. 2019; Matthias and Kretschmer 2020; Millin et al. 2022). Our study borrows from parallel studies on European CAO events (e.g., Kautz et al. 2020; Huang et al. 2022) and addresses the gap in the literature on the mechanisms of how these features may impact the subseasonal prediction of North American CAOs. We employ nudging experiments in the CESM2 CAM6 subseasonal model simulations of two dynamically different central U.S. CAO events to compare and contrast the contributions of tropical versus stratospheric circulation variability on the subseasonal forecasts of these CAOs.

In the case of the December 2017/January 2018 CAO, the NUDGE-S experiment results in robust forecasts of cold temperatures throughout Canada and the eastern United States on subseasonal time scales (Figs. 4 and 5). Additionally, the NUDGE-S experiment also improves the forecasts of the North American winter weather regimes (proxies for sensible weather patterns) in the 2 weeks leading up to and including the CAO (Fig. 6 and Fig. S5). Thus, better representations of the stratosphere, including the wave reflection mechanism, result in an improved subseasonal forecast for the 2017/18 CAO versus the FREE run. Despite this improvement, the T2M

response in the central United States was relatively modest, highlighting that other processes may have played a role in the cold development there. The NUDGE-T experiment also improved the simulation of the wave reflection during December 2017, particularly the increase in upward Siberian wave activity flux (Fig. 3). This experiment also corresponded to a better representation of the Alaskan ridge and a modest improvement in the T2M forecast across the eastern United States (Figs. 4 and 5). Indeed, this finding agrees with prior work showing that active MJO convection in the western and central tropical Pacific is associated with increased ridging near Alaska (Henderson et al. 2016; Henderson and Maloney 2018; Millin et al. 2022). As such, there may be importance to both tropical and stratospheric processes in the evolution of this CAO event.

For the January 2004 CAO event, we found that both the NUDGE-S and NUDGE-T experiments produced little improvement in T2M forecasts across the United States when compared to the FREE run (Figs. 8 and 9). In fact, the T2M response in the Great Plains in the NUDGE-S experiment was warmer than in the FREE run during the CAO period. This finding is somewhat contradictory to canonical relationships between major SSW events and midlatitude CAOs, though the connection between SSWs to North American cold is likely less robust than that for Europe (Domeisen et al. 2020; Davis et al. 2022; Millin et al. 2024). Despite less improvement in the T2M response than in the December 2017/18 nudged runs, the NUDGE-S and NUDGE-T experiments improved forecasts of the North American winter weather regimes compared to the FREE run (Fig. 10 and Fig. S7). One possible explanation for this apparent contradiction in skill between temperature and weather regime forecasts is the timing of CAO events. CAOs have large-scale forcing elements, but their timing and magnitude rely on smaller-scale features including the synoptic timing of troughs, dynamics and thermodynamics (e.g., diabatic cooling, cold air advection), tropopause polar vortices, and synoptic-scale wave breaks (Cavallo and Hakim 2009, 2010, 2013; Millin and Furtado 2022). Furthermore, not every major SSW propagates downward to yield surface impacts (e.g., Karpechko et al. 2017). The January 2019 (e.g., Rao et al. 2019) and January 2021 (e.g., Davis et al. 2022) major SSWs are two recent examples of events not directly connected to North American winter weather. Ultimately, the subseasonal prediction problem linking major SSW activity to CAOs in North America may be two-pronged. Major SSWs may increase the probability of CAOs and thus, from a probabilistic perspective, there is increased forecast confidence for such cold events weeks post-SSW (Baldwin and Dunkerton 2001; Domeisen et al. 2020; Furtado et al. 2021). However, timing the CAO event and its magnitude likely lies at the synoptic-scale forecasting time scale.

Several caveats to our work exist. This study only considered two events, limiting the generalization of our findings on the contributions of the stratosphere and tropics on the evolution of North American CAOs. As such, we recommend reproducing our experiment design for more CAO events. An additional caveat is that we selected events only from a central U.S. CAO database, whereas events from other domains

within the United States may also be fruitful. Additionally, we provided nudged experiments for the stratosphere and tropics separately to isolate their impacts, but the nudging of both the stratosphere and tropics together would be interesting to consider. Furthermore, future work should investigate the impact of the lower boundary of the nudging level in the stratosphere on the resulting forecast evolution. Sensitivity experiments with nudging confined to higher stratospheric levels produce a weaker tropospheric response for our cases but retain a similar large-scale circulation and WRI evolution (not shown). As discussed in Davis et al. (2022) and Cohen et al. (2023), wave amplification associated with North American CAOs depends not simply on the presence of upward or downward wave activity flux but on the convergence/divergence of wave activity in the lower stratosphere and in the vicinity of the tropopause, which governs whether reflected or transmitted waves reinforce tropospheric circulation anomalies. By changing the nudging constraint in the lower stratosphere, these sensitivity experiments likely underestimate near-tropopause wave activity flux convergence/divergence and therefore yield a muted response.

Future work should extend this analysis to look at the impact of the stratosphere and tropics on blocking frequency in subseasonal forecasts, which are hallmark features for many extreme weather events during the cold season such as extreme precipitation and even winter warm spells (Grace et al. 2024). Further understanding the stratospheric and tropical circulations will be crucial to S2S predictability of surface extremes and their potential impacts. One forecast metric may be the winter weather regimes themselves, as their large-scale nature may make them “easier” targets for subseasonal forecasts, both in dynamical forecast systems and data-driven models. Indeed, future work should also consider employing nudging on the full reforecast suite to assess how the representation of these features affects the overall skill of the model during the winter season.

Acknowledgments. We thank the National Science Foundation (NSF) Oklahoma Established Program to Stimulate Competitive Research (EPSCoR) project Grant OIA-1946093 for the financial support that enabled the completion of this work. We also thank Sasha Glanville, Jim Edwards, and Kai Huang for their help in setting up the CESM2 CAM6 model for the nudging experiments. Finally, we acknowledge that this work has been partly completed using the Derecho and Casper supercomputers at the National Center for Atmospheric Research [Computational and Information Systems Laboratory (CISL) 2023].

Data availability statement. MERRA-2 data can be downloaded using the MDISC at <https://disc.gsfc.nasa.gov/datasets/project=MERRA-2>. MERRA-2 meteorology nudging files for CESM2 CAM6 can be found at <https://wiki.ucar.edu/spaces/camchem/pages/452821200/Meteorology+for+physics-based+nudging>. All code used for the analyses in this paper can be found at <https://doi.org/10.5281/zenodo.19486711>. CESM2 S2S hindcasts can be found on NCAR’s Derecho supercomputer at

the file path /glade/campaign/cesm/development/cross-wg/S2S/CESM2/S2SHINDCASTS/. Model experiment data are available on request.

REFERENCES

- Afarghan-Gerstman, H., and D. I. V. Domeisen, 2020: Pacific modulation of the North Atlantic storm track response to sudden stratospheric warming events. *Geophys. Res. Lett.*, **47**, e2019GL085007, <https://doi.org/10.1029/2019GL085007>.
- 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>.
- Bloomfield, H. C., D. J. Brayshaw, P. L. M. Gonzalez, and A. Charlton-Perez, 2021: Pattern-based conditioning enhances sub-seasonal prediction skill of European national energy variables. *Meteor. Appl.*, **28**, e2018, <https://doi.org/10.1002/met.2018>.
- Bolinger, R. A., and Coauthors, 2022: An assessment of the extremes and impacts of the February 2021 South-Central U.S. Arctic outbreak, and how climate services can help. *Wea. Climate Extremes*, **36**, 100461, <https://doi.org/10.1016/j.wace.2022.100461>.
- BoM, 2025: Tropical monitoring and outlooks. Accessed 26 March 2025, <https://www.bom.gov.au/climate/mjo/#tabs=Monitoring>.
- Butler, A. H., J. P. Sjöberg, 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>.
- Cavallo, S. M., and G. J. Hakim, 2009: Potential vorticity diagnosis of a tropopause polar cyclone. *Mon. Wea. Rev.*, **137**, 1358–1371, <https://doi.org/10.1175/2008MWR2670.1>.
- , and —, 2010: Composite structure of tropopause polar cyclones. *Mon. Wea. Rev.*, **138**, 3840–3857, <https://doi.org/10.1175/2010MWR3371.1>.
- , and —, 2013: Physical mechanisms of tropopause polar vortex intensity change. *J. Atmos. Sci.*, **70**, 3359–3373, <https://doi.org/10.1175/JAS-D-13-088.1>.
- 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, A. J., and L. M. Polvani, 2007: A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *J. Climate*, **20**, 449–469, <https://doi.org/10.1175/JCLI3996.1>.
- CISL, 2023: Derecho: HPE Cray EX System (university community computing). NSF National Center for Atmospheric Research, <https://doi.org/doi:10.5065/qx9a-pg09>.
- Cohen, J., L. Agel, M. Barlow, C. I. Garfinkel, and I. White, 2023: Response to limited surface impacts of the January 2021 sudden stratospheric warming. *Nat. Commun.*, **14**, 3289, <https://doi.org/10.1038/s41467-023-38772-3>.
- Davis, N. A., J. H. Richter, A. A. Glanvilles, J. Edwards, and E. LaJoie, 2022: Limited surface impacts of the January 2021 sudden stratospheric warming. *Nat. Commun.*, **13**, 1136, <https://doi.org/10.1038/s41467-022-28836-1>.
- Dias, J., S. N. Tulich, M. Gehne, and G. N. Kiladis, 2021: Tropical origins of weeks 2–4 forecasts errors during Northern Hemisphere cool season. *Mon. Wea. Rev.*, **149**, 2975–2991, <https://doi.org/10.1175/MWR-D-21-0020.1>.
- Ding, X., G. Chen, P. Zhang, D. I. V. Domeisen, and C. Orbe, 2023: Extreme stratospheric wave activity as harbingers of cold events over North America. *Nat. Commun. Earth Environ.*, **4**, 187, <https://doi.org/10.1038/s43247-023-00845-y>.
- Domeisen, D. I. V., and Coauthors, 2020: The role of the stratosphere in subseasonal to seasonal prediction: 1. Predictability of the stratosphere. *J. Geophys. Res. Atmos.*, **125**, e2019JD030920, <https://doi.org/10.1029/2019JD030920>.
- Doss-Gollin, J., D. J. Farnham, U. Lall, and V. Modi, 2021: How unprecedented was the February 2021 Texas cold snap? *Environ. Res. Lett.*, **16**, 064056, <https://doi.org/10.1088/1748-9326/ac0278>.
- Erner, I., and A. Karpechko, 2024: Factors influencing subseasonal predictability of northern Eurasian cold spells. *Quart. J. Roy. Meteor. Soc.*, **150**, 2955–2975, <https://doi.org/10.1002/qj.4744>.
- Furtado, J. C., J. Cohen, E. J. Becker, and D. C. Collins, 2021: Evaluating the relationship between sudden stratospheric warmings and tropospheric weather regimes in the NMME phase-2 models. *Climate Dyn.*, **56**, 2321–2338, <https://doi.org/10.1007/s00382-020-05591-x>.
- Garfinkel, C. I., S. B. Feldstein, D. W. Waugh, C. Yoo, and S. Lee, 2012: Observed connection between stratospheric sudden warmings and the Madden-Julian Oscillation. *Geophys. Res. Lett.*, **39**, L18807, <https://doi.org/10.1029/2012GL053144>.
- , J. J. Benedict, and E. D. Maloney, 2014: Impact of the MJO on the boreal winter extratropical circulation. *Geophys. Res. Lett.*, **41**, 6055–6062, <https://doi.org/10.1002/2014GL061094>.
- , and Coauthors, 2025: A process-based evaluation of biases in extratropical stratosphere–troposphere coupling in subseasonal forecast systems. *Wea. Climate Dyn.*, **6**, 171–195, <https://doi.org/10.5194/wcd-6-171-2025>.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Grace, T. M., K. Pegion, and J. B. Basara, 2024: Atmospheric and surface dynamics during winter warm spells in the Southern Great Plains: Insights from the 2021 case study. *J. Geophys. Res. Atmos.*, **129**, e2024JD040856, <https://doi.org/10.1029/2024JD040856>.
- Green, M. R., and J. C. Furtado, 2019: Evaluating the joint influence of the Madden-Julian oscillation and the stratospheric polar vortex on weather patterns in the Northern Hemisphere. *J. Geophys. Res. Atmos.*, **124**, 11 693–11 709, <https://doi.org/10.1029/2019JD030771>.
- Harnik, N., 2009: Observed stratospheric downward reflection and its relation to upward pulses of wave activity. *J. Geophys. Res.*, **114**, D08120, <https://doi.org/10.1029/2008JD010493>.
- Henderson, S. A., and E. D. Maloney, 2018: The impact of the Madden-Julian oscillation on high-latitude winter blocking during El Niño–Southern Oscillation events. *J. Climate*, **31**, 5293–5318, <https://doi.org/10.1175/JCLI-D-17-0721.1>.
- , —, and E. A. Barnes, 2016: The influence of the Madden-Julian oscillation on Northern Hemisphere winter blocking. *J. Climate*, **29**, 4597–4616, <https://doi.org/10.1175/JCLI-D-15-0502.1>.
- Hitchcock, P., and I. R. Simpson, 2014: The downward influence of stratospheric sudden warmings. *J. Atmos. Sci.*, **71**, 3856–3876, <https://doi.org/10.1175/JAS-D-14-0012.1>.
- , and Coauthors, 2022: Stratospheric Nudging and Predictable Surface Impacts (SNAPSI): A protocol for investigating the role of stratospheric polar vortex disturbances in subseasonal to seasonal forecasts. *Geosci. Model Dev.*, **15**, 5073–5092, <https://doi.org/10.5194/gmd-15-5073-2022>.

- Huang, J., P. Hitchcock, W. Tian, and J. Sillin, 2022: Stratospheric influence on the development of the 2018 late winter European cold air outbreak. *J. Geophys. Res. Atmos.*, **127**, e2021JD035877, <https://doi.org/10.1029/2021JD035877>.
- Huang, K., J. H. Richter, and K. V. Pegion, 2023: Captured QBO-MJO connection in a subseasonal prediction system. *Geophys. Res. Lett.*, **50**, e2022GL102648, <https://doi.org/10.1029/2022GL102648>.
- 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>.
- Jung, T., M. J. Miller, and T. N. Palmer, 2010: Diagnosing the origin of extended-range forecast errors. *Mon. Wea. Rev.*, **138**, 2434–2446, <https://doi.org/10.1175/2010MWR3255.1>.
- Karpechko, A. Y., P. Hitchcock, D. H. W. Peters, and A. Schneidereit, 2017: Predictability of downward propagation of major sudden stratospheric warmings. *Quart. J. Roy. Meteor. Soc.*, **143**, 1459–1470, <https://doi.org/10.1002/qj.3017>.
- , F. Vitart, I. Statnaia, M. A. Balmaseda, A. J. Charlton-Perez, and I. Polichtchouk, 2024: The tropical influence on sub-seasonal predictability of wintertime stratosphere and stratosphere–troposphere coupling. *Quart. J. Roy. Meteor. Soc.*, **150**, 1357–1374, <https://doi.org/10.1002/qj.4649>.
- Kautz, L.-A., I. Polichtchouk, T. Birner, H. Garny, and J. G. Pinto, 2020: Enhanced extended-range predictability of the 2018 late-winter Eurasian cold spell due to the stratosphere. *Quart. J. Roy. Meteor. Soc.*, **146**, 1040–1055, <https://doi.org/10.1002/qj.3724>.
- Kodera, K., H. Mukougawa, and S. Itoh, 2008: Tropospheric impact of reflected planetary waves from the stratosphere. *Geophys. Res. Lett.*, **35**, L16806, <https://doi.org/10.1029/2008GL034575>.
- 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>.
- Lee, S. H., J. C. Furtado, and A. J. Charlton-Perez, 2019: Wintertime North American weather regimes and the Arctic stratospheric polar vortex. *Geophys. Res. Lett.*, **46**, 14892–14900, <https://doi.org/10.1029/2019GL085592>.
- , A. J. Charlton-Perez, S. J. Woolnough, and J. C. Furtado, 2022: How do stratospheric perturbations influence North American weather regime predictions? *J. Climate*, **35**, 5915–5932, <https://doi.org/10.1175/JCLI-D-21-0413.1>.
- Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708, [https://doi.org/10.1175/1520-0469\(1971\)028<0702:DOADOI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2).
- , and —, 1972: Description of global-scale circulation cells in the tropics with a 40–50 day period. *J. Atmos. Sci.*, **29**, 1109–1123, [https://doi.org/10.1175/1520-0469\(1972\)029<1109:DOGGSC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1972)029<1109:DOGGSC>2.0.CO;2).
- Mariotti, A., and Coauthors, 2020: Windows of opportunity for skillful forecasts subseasonal to seasonal and beyond. *Bull. Amer. Meteor. Soc.*, **101**, E608–E625, <https://doi.org/10.1175/BAMS-D-18-0326.1>.
- Matthias, V., and M. Kretschmer, 2020: The influence of stratospheric wave reflection on North American cold spells. *Mon. Wea. Rev.*, **148**, 1675–1690, <https://doi.org/10.1175/MWR-D-19-0339.1>.
- Messori, G., M. Kretschmer, S. H. Lee, and V. Wendt, 2022: Stratospheric downward wave reflection events modulate North American weather regimes and cold spells. *Wea. Climate Dyn.*, **3**, 1215–1236, <https://doi.org/10.5194/wcd-3-1215-2022>.
- Michel, C., and G. Rivière, 2011: The link between Rossby wave breakings and weather regime transitions. *J. Atmos. Sci.*, **68**, 1730–1748, <https://doi.org/10.1175/2011JAS3635.1>.
- Millin, O. T., and J. C. Furtado, 2022: The role of wave breaking in the development and subseasonal forecasts of the February 2021 Great Plains cold air outbreak. *Geophys. Res. Lett.*, **49**, e2022GL100835, <https://doi.org/10.1029/2022GL100835>.
- , —, and J. B. Basara, 2022: Characteristics, evolution, and formation of cold air outbreaks in the Great Plains of the United States. *J. Climate*, **35**, 4585–4602, <https://doi.org/10.1175/JCLI-D-21-0772.1>.
- , —, and C. Malloy, 2024: The impact of North American winter weather regimes on electricity load in the central United States. *npj Climate Atmos. Sci.*, **7**, 254, <https://doi.org/10.1038/s41612-024-00803-1>.
- Nakagawa, K. I., and K. Yamazaki, 2006: What kind of stratospheric sudden warming propagates to the troposphere? *Geophys. Res. Lett.*, **33**, L04801, <https://doi.org/10.1029/2005GL024784>.
- NCEI, 2025: Billion-dollar weather and climate disasters: Events. Accessed 27 February 2025, <https://www.ncei.noaa.gov/access/billions/events>.
- Pedragosa, F., and Coauthors, 2011: Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, **12**, 2825–2830.
- Pegion, K., and Coauthors, 2019: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.*, **100**, 2043–2060, <https://doi.org/10.1175/BAMS-D-18-0270.1>.
- Perlitz, J., and N. Harnik, 2003: Observational evidence of a stratospheric influence on the troposphere by planetary wave reflection. *J. Climate*, **16**, 3011–3026, [https://doi.org/10.1175/1520-0442\(2003\)016<3011:OEAOASI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<3011:OEAOASI>2.0.CO;2).
- , and —, 2004: Downward coupling between the stratosphere and troposphere: The relative roles of wave and zonal mean processes. *J. Climate*, **17**, 4902–4909, <https://doi.org/10.1175/JCLI-3247.1>.
- Quinting, J., C. M. Grams, E. K.-M. Chang, S. Pfahl, and H. Wernli, 2024: Warm conveyor belt activity over the Pacific: Modulation by the Madden–Julian Oscillation and impact on tropical–extratropical teleconnections. *Wea. Climate Dyn.*, **5**, 65–85, <https://doi.org/10.5194/wcd-5-65-2024>.
- Rao, J., C. I. Garfinkel, H. Chen, and I. P. White, 2019: The 2019 new year stratospheric sudden warming and its real-time predictions in multiple S2S models. *J. Geophys. Res. Atmos.*, **124**, 11 155–11 174, <https://doi.org/10.1029/2019JD030826>.
- Richter, J. H., and Coauthors, 2022: Subseasonal Earth system prediction with CESM2. *Wea. Forecasting*, **37**, 797–815, <https://doi.org/10.1175/WAF-D-21-0163.1>.
- Schutte, M. K., A. Portal, S. H. Lee, and G. Messori, 2025: Dynamics of stratospheric wave reflection over the North Pacific. *Wea. Climate Dyn.*, **6**, 521–548, <https://doi.org/10.5194/wcd-6-521-2025>.
- Smith, E. T., and S. C. Sheridan, 2018: The characteristics of extreme cold events and cold air outbreaks in the eastern United States. *Int. J. Climatol.*, **38**, e807–e820, <https://doi.org/10.1002/joc.5408>.
- , and —, 2020: Where do cold air outbreaks occur, and how have they changed over time? *Geophys. Res. Lett.*, **47**, e2020GL086983, <https://doi.org/10.1029/2020GL086983>.

- Thompson, D. W. J., and J. M. Wallace, 2000: Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate*, **13**, 1000–1016, [https://doi.org/10.1175/1520-0442\(2000\)013<1000:AMITEC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<1000:AMITEC>2.0.CO;2).
- UCAR, 2024: Meteorology for physics-based nudging. Accessed 26 March 2025, <https://wiki.ucar.edu/display/camchem/Meteorology+for+physics-based+nudging>.
- Vitart, F., and A. W. Robertson, 2018: The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *npj Climate Atmos. Sci.*, **1**, 3, <https://doi.org/10.1038/s41612-018-0013-0>.
- , and M. A. Balmaseda, 2024: Sources of MJO teleconnection errors in the ECMWF extended-range forecasts. *Quart. J. Roy. Meteor. Soc.*, **150**, 2028–2044, <https://doi.org/10.1002/qj.4688>.
- Wheeler, D. D., V. L. Harvey, D. E. Atkinson, R. L. Collins, and M. J. Mills, 2011: A climatology of cold air outbreaks over North America: WACCM and ERA-40 comparison and analysis. *Geophys. Res. Lett.*, **116**, D12107, <https://doi.org/10.1029/2011JD015711>.
- 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>.
- White, I., C. I. Garfinkel, E. P. Gerber, M. Jucker, V. Aquila, and L. D. Oman, 2019: The downward influence of sudden stratospheric warmings: Association with tropospheric precursors. *J. Climate*, **32**, 85–108, <https://doi.org/10.1175/JCLI-D-18-0053.1>.