Geophysical Research Letters

RESEARCH LETTER
10.1029/2018GL081314

Special Section:
Bridging Weather and Climate: Subseasonal-to-Seasonal (S2S) Prediction

Subseasonal Week 3–5 Surface Air Temperature Prediction During Boreal Wintertime in a GFDL Model

Baoqiang Xiang[1,2], Shian-Jiann Lin[1], Ming Zhao[1], Nathaniel C. Johnson[1,3], Xiaosong Yang[1,2], and Xianan Jiang[4,5]

[1]NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, [2]Cooperative Programs for the Advancement of Earth System Science, University Corporation for Atmospheric Research, Boulder, CO, USA, [3]Cooperative Institute for Climate Science, Princeton University, Princeton, NJ, USA, [4]Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA, USA, [5]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Key Points:
- Significant prediction skill is found in a GFDL model for subseasonal wintertime week 3–5 surface air temperature prediction
- Three most predictable modes (ENSO, NAO, and the Eurasia Meridional Dipole mode) are identified that show a prediction skill of 5 weeks
- The NAO and the Eurasia Meridional Dipole modes are linked to the initial stratospheric polar vortex which is an important predictability source

Supporting Information:
- Supporting Information S1

Correspondence to:
B. Xiang,
baoqiang.xiang@noaa.gov

Citation:
Xiang, B., Lin, S.-J., Zhao, M., Johnson, N. C., Yang, X., & Jiang, X. (2019). Subseasonal week 3–5 surface air temperature prediction during boreal wintertime in a GFDL model. *Geophysical Research Letters*, *46*, 416–425. https://doi.org/10.1029/2018GL081314

Received 13 NOV 2018
Accepted 8 DEC 2018
Accepted article online 17 DEC 2018
Published online 8 JAN 2019

©2018. American Geophysical Union. All Rights Reserved.

Abstract With a Geophysical Fluid Dynamics Laboratory (GFDL) coupled model, the subseasonal prediction of wintertime (December–February) surface air temperature (SAT) is investigated through the analysis of 11-year hindcasts. Significant subseasonal week 3–5 correlation skill exists over a large portion of the global land domain, and the predictability originates primarily from the eight most predictable SAT modes. The first three modes, identified as the El Niño-Southern Oscillation mode, the North Atlantic Oscillation mode, and the Eurasia Meridional Dipole mode, can be skillfully predicted more than 5 weeks in advance. The North Atlantic Oscillation and Eurasia Meridional Dipole modes are strongly correlated with the initial stratospheric polar vortex strength, highlighting the role of stratosphere in subseasonal prediction. Interestingly, the Madden-Julian Oscillation is not essential for the subseasonal land SAT prediction in the Northern Hemisphere extratropics. The spatial correlation skill exhibits considerable intraseasonal and interannual fluctuations, indicative of the importance to identify the time window of opportunity for subseasonal prediction.

Plain Language Summary Subseasonal prediction has important societal and economic impacts, while it remains very challenging as a prediction frontier. In this study, we use a dynamic model with simple initialization method to investigate the subseasonal week 3–5 prediction of the surface air temperature (SAT) in boreal winter. It is encouraging to note that the model has a significant week 3–5 SAT prediction skill over a large portion of the land regions. Using a statistical method, we further examine the potential predictability sources and find eight most predictable modes with the first three identified as the El Niño/Southern Oscillation, the North Atlantic Oscillation, and the Eurasia Meridional Dipole Mode. Interestingly, although the Madden-Julian Oscillation is the dominant intraseasonal mode over the tropics, it has relatively weak impacts on the SAT prediction in the Northern Hemisphere extratropics. The strong fluctuation of the prediction skill varies from week to week and year to year highlighting the importance to identify the time window of opportunity for subseasonal prediction. The role of stratospheric polar vortex in subseasonal SAT prediction is also shown.

1. Introduction

Subseasonal prediction, usually referring to the timescale between 2 weeks and 2 months, is emerging as one of the top priorities in prediction research, given its connection with the growing societal demand associated with public health, disaster preparedness, water management, and so on. Subseasonal timescales have been long regarded as a "predictability desert," and the subseasonal prediction is still at its infancy and developing stage. How to improve the current subseasonal prediction skill is one of the most challenging but important tasks in the weather and climate prediction community.

Despite being in its early stages, considerable progress has been achieved in the past two decades for subseasonal prediction. These achievements originate from the improved model physics, resolution, initialization, as well as our improved understanding of the predictability sources. This progress is also substantially prompted by the multiagency and multi-institute joint efforts, such as the North-American Multi-Model Ensemble (NMME; Kirtman et al., 2013), IntraSeasonal Variability Hindcast Experiment (ISVHE; Neena et al., 2014), the Subseasonal Experiment (SubX; http://cola.gmu.edu/kpegion/subx/), and

the WWRP/WCRP subseasonal to seasonal prediction project (S2S; Vitart et al., 2016). Development of statistical models is also instrumental in identifying the subseasonal predictability sources (e.g., Black et al., 2017; Johnson et al., 2013), which, in turn, provides guidance on the further model development and more effective observational strategies.

In this study, our focus is on evaluating the predictability and exploring the most predictable modes (MPMs) for subseasonal wintertime surface air temperature (SAT) prediction using a Geophysical Fluid Dynamics Laboratory (GFDL) coupled model system. We consider lead times ranging from 1 to 7 weeks, with an emphasis on week 3–5 predictions, given that this timescale lies between the traditional targets of weather and climate prediction.

2. Model, Experiments, and Methodology

2.1. Model and Hindcast Experiments

The dynamical forecast model used here is based on the Forecast-oriented Low Ocean Resolution version of GFDL model (Vecchi et al., 2014), but with a new double-plume convection scheme (Zhao et al., 2018). The horizontal resolution for the atmospheric (ocean) model is about 50 km (100 km). The atmosphere (ocean) has 32 (50) vertical levels. Initial conditions were generated through a nudging technique for both the atmosphere and the ocean models. The atmospheric fields include winds, temperature, geopotential height, and surface pressure, which are nudged toward the National Centers for Environmental Prediction Global Forecast System analysis data. The sea surface temperature is nudged to the National Oceanic and Atmospheric Administration Optimum Interpolation 1/4 Degree Daily Sea Surface Temperature Analysis (OISST, v2; Reynolds et al., 2007). In this study, we analyze forecasts that were initialized in December–February. We have one case every 5 days and in total 198 (11 years × 3 months × 6 days/month) forecast cases during an 11-year period (2003–2013). For each case, there are six ensemble members with each integrated for 50 days. This forecast system has been used to study the Madden-Julian Oscillation (MJO) prediction (Xiang et al., 2015) and tropical cyclone genesis predictions (Jiang et al., 2018; Xiang et al., 2014).

The verification SAT and 500 hPa geopotential height data are from the National Centers for Environmental Prediction Global Forecast System analysis data. Although we focus on SAT predictions, we briefly examine observed precipitation data from the Global Precipitation Climatology Project v2.2 (Huffman et al., 2016) in association with the MJO.

The observational anomalous fields are calculated by removing the time mean and first three harmonics of the climatological annual cycle from the observational data (2002–2013) following Wheeler and Hendon (2004). The hindcast anomalies are obtained by removing the hindcast climatology as a function of hindcast starting date and lead time. Because our focus here is the subseasonal prediction instead of the prediction of subseasonal variability, we do not subtract the previous 120-day mean of anomalies so that variability at the interannual and even longer timescales is retained for both observations and hindcasts. This is different from our previous MJO prediction study (Xiang et al., 2015).

2.2. Average Predictability Time Analysis

To elaborate on the sources of predictability, we adopt the Average Predictability Time (APT) analysis proposed by DelSole and Tippett (2009a, 2009b). This analysis identifies patterns that maximize predictability. In this application, a standard measure of predictability is defined as

$$P(\tau) = 1 - \frac{\sigma_\tau^2}{\sigma_{clim}^2}, \tag{1}$$

where σ_τ^2 represents the variance of hindcast anomalies and the σ_{clim}^2 is the time-averaged climatological variance defined as the variance of all ensemble forecasts from different forecast lead times (τ). $P(\tau)$ typically has a value near 1 for the shortest lead times and decreases to near 0 as the forecast variance approaches the climatological variance. The APT is further defined as twice the sum of $P(\tau)$ over all lead times (weeks 1–7 in the present application):

$$APT = 2\sum_{\tau=1}^{\infty}\left(1 - \frac{\sigma_\tau^2}{\sigma_{clim}^2}\right). \tag{2}$$

Through a series of derivations, maximizing APT leads to a generalized eigenvalue problem (DelSole & Tippett, 2009a, 2009b):

$$2\sum_{\tau=1}^{\infty}(\sum_{clim}-\sum_{\tau})\boldsymbol{q}=\lambda\sum_{clim}\boldsymbol{q}. \quad (3)$$

The eigenvectors $\boldsymbol{q}$ decompose the time series into individual modes with the first maximizing APT and the second maximizing APT subject to being uncorrelated with the first one, and so on. The eigenvalue λ gives APT value of one independent component. With the eigenvector $\boldsymbol{q}$, the spatial pattern is obtained by regressing the time series of $\boldsymbol{q}^T\boldsymbol{x}(\tau,t,e)$ onto $\boldsymbol{x}(\tau,t,e)$. Here $\boldsymbol{x}(\tau,t,e)$ is the state variable at fixed lead time τ, start time t, and ensemble member e. The decomposition based on APT is analogous to empirical orthogonal function analysis except that it decomposes predictability instead of variance. For more details, we refer the reader to DelSole and Tippett (2009a, 2009b) and several following studies (Jia et al., 2015; Yang et al., 2013).

To assess the significance of the anomalous correlation, the effective sample size (N_{eff}) is calculated by considering the autocorrelation in the observed and forecast data (Bretherton et al., 1999; Yang et al., 2013):

$$N_{eff}=\frac{N}{1+2\sum_{j=1}^{N-1}\left(1-\frac{j}{N}\right)\rho_{\mathrm{obs}}(\mathrm{j})\rho_{mod}(j)}, \quad (4)$$

where N is the number of sample pairs and $\rho_{obs}(j)$ and $\rho_{mod}(j)$ are the sample autocorrelation of observation and model hindcast at lag j.

3. Results

The variance of wintertime SAT anomalies is first examined for both observations and model hindcasts. The hindcasts reasonably capture both the spatial pattern and the magnitude of the observational SAT anomalies, with its strong variability in northern Eurasia, Alaska, Canada, as well as the Arctic region (supporting information Figure S1).

3.1. Prediction of Week 3–5 SAT Anomalies

We next examine the prediction skill from the unfiltered model predictions during week 3–5 (Figures 1a–1c). This model shows encouraging results with a significant correlation skill over the majority of regions during week 3, except the southwestern United States, northern Eurasia, Arctic, and the northeast Atlantic Ocean (Figure 1a). The forecast skill generally decreases with the increase of forecast lead time, while the correlation skill remains significant at the 5% level over a large portion of the land domain for the week 4–5, such as eastern China, Canada, Alaska, southeastern United States, Amazon, and North Africa (Figures 1b and 1c).

3.2. Identification of the MPMs

To reveal the predictability sources, the APT analysis is performed with respect to the predicted weekly mean SAT anomalies for six ensemble members from week 1 to 7. We choose 15 modes to maximize APT that explain about 62% of the total SAT variance. Given the large sample of the hindcast data, the results are not sensitive to either the number of total modes (we test 20 modes) or the choice of the forecast period (we test week 1–4).

To demonstrate the importance of the leading modes in subseasonal prediction, the hindcast anomalies are reconstructed using the spatial patterns and time series of the first eight MPMs. The reason to choose these eight modes is that their corresponding time series can be skillfully predicted more than 3 weeks in advance. Note that these eight MPMs can explain 37% of the total variance. Compared to the unfiltered (raw) model, the reconstructed anomalies show a fairly comparable land SAT prediction skill at week 3–5 (Figure 1).

The first three MPMs are highlighted here as they dominate the longer timescale prediction (week 4–5). The first MPM (MPM1) is characterized by pronounced positive SAT anomalies over the equatorial central and eastern Pacific, accompanied by a seesaw pattern both in North America and the Eurasian continent (Figure 2a). Superimposed on Figure 2a are the week 3 correlation coefficients between the time series of

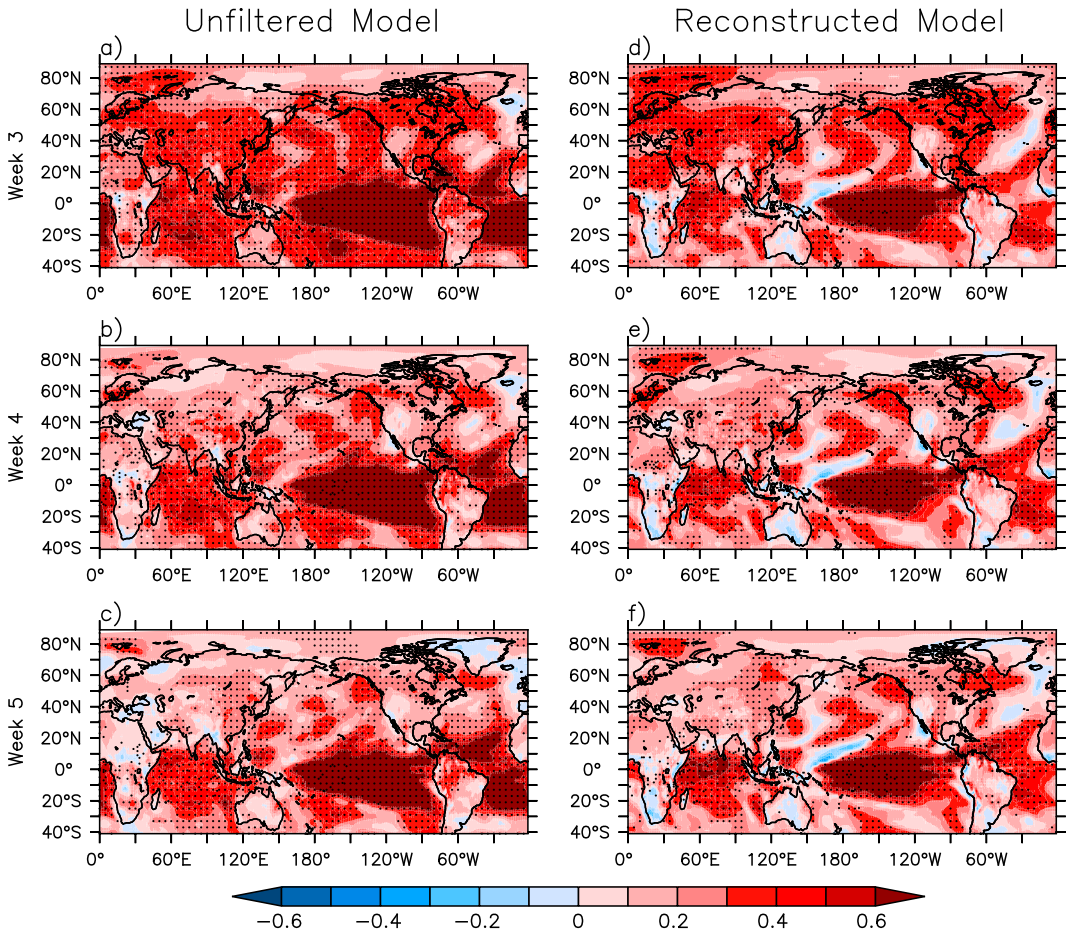

Figure 1. The correlation skill of the unfiltered surface air temperature anomalies with a lead time of (a) 3, (b) 4, and (c) 5 weeks. (d)–(f) are similar to (a)–(c) but for the reconstructed model forecast skill using the first eight most predictable modes. The dots denote the region with the correlation skill significant at the 5% significance level considering the effective sample size.

this mode and the 500 hPa geopotential height (H500) anomalies from model hindcasts. Clearly evident are strong positive correlations over the whole tropics together with a Pacific-North America teleconnection pattern forced by the equatorial Pacific warming. The SAT and midtropospheric circulation patterns of MPM1 resemble the characteristic patterns of the warm phase of the El Niño-Southern Oscillation (ENSO) variability. Hereinafter, we refer to this mode as the ENSO mode. The large loading of this mode in the United States is consistent with the finding (DelSole et al., 2017) that ENSO plays the dominant role in the wintertime SAT prediction (week 3–4) over the Contiguous United States.

We note, as discussed below, that this mode also exhibits pronounced subseasonal variability, which indicates that it also captures subseasonal Pacific-North America-like variability unrelated to ENSO. In the North Atlantic basin, a meridional dipole pattern is reminiscent of a negative phase of North Atlantic Oscillation (NAO; Figure 2a). This is consistent with the fact that the interannual variability in the tropical Pacific is dominated by the central Pacific El Niño after the late 1990s (Xiang et al., 2013), which may have more impact on the NAO than the eastern Pacific El Niño (Graf & Zanchettin, 2012). One way that ENSO impacts the North Atlantic and European sector is through an indirect pathway from the stratosphere (e.g., Butler et al., 2014; Polvani et al., 2017).

For the second MPM (MPM2), pronounced negative SAT anomalies are present in the northeast Canada, Greenland, and northern Africa, together with the positive anomalies in the Eurasia continent and Arctic (Figure 2b). At week 3, the associated anomalous H500 displays an enhanced subtropical high in the North Atlantic and a strong Icelandic low, giving rise to stronger-than-average westerlies across the

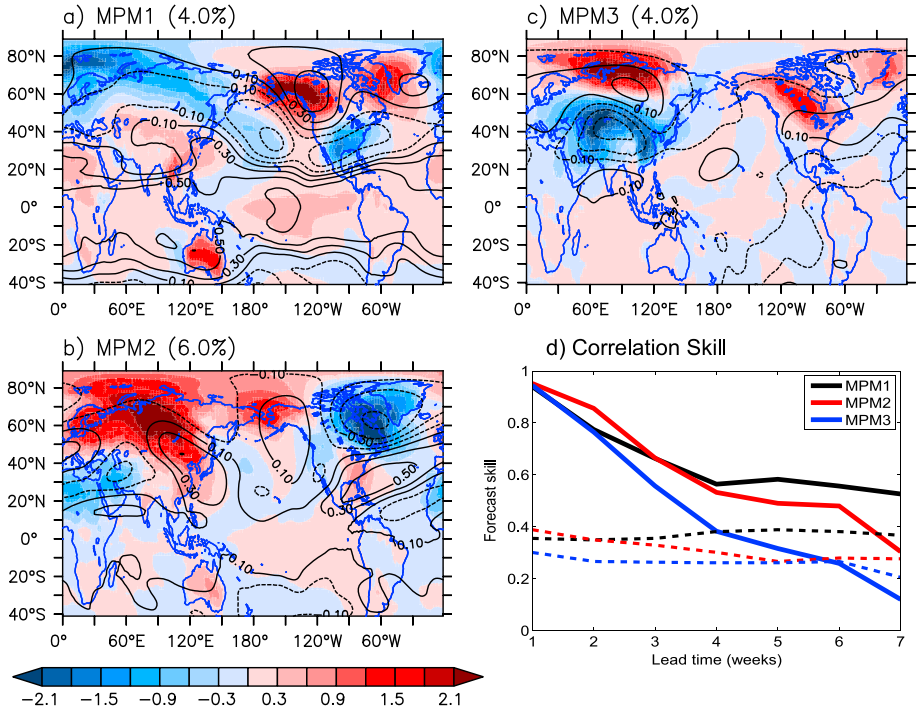

Figure 2. (a) The spatial pattern of the first MPM (MPM1) of the weakly mean surface air temperature anomalies (°C) during the wintertime (December–February). The contours show the correlation coefficient between the time series of MPM1 and the 500 hPa geopotential height (H500) anomalies at week 3. Both surface air temperature and H500 are from model hindcast results. (b) and (c) are similar to (a) but for the second and third MPMs (MPM2, MPM3). The variance explained by these modes is shown in the parentheses. (d) The correlation skill between the observed and predicted time series of these three modes as a function of lead times (solid lines). The dashed lines indicate the 5% significance level after considering the effective sample size. MPM = most predictable mode.

midlatitudes. The gross features of this mode are reminiscent of the NAO-like (positive phase) circulation pattern that brings warm air to the northern Europe and the eastern United States. Hereinafter, this mode is referred as to the NAO mode.

The third MPM (MPM3) features a meridional dipole structure in both H500 and SAT over Eurasia, and predominantly positive SAT anomalies centered in North America (Figure 2c). This mode bears a resemblance to the second empirical orthogonal function mode shown in Lin (2018). The formation of this mode is partially linked to the stratospheric polar vortex, as discussed later. Since the major loading of this mode is in the Eurasian continent, we refer this mode as to the Eurasia Meridional Dipole (EMD) mode.

The time series of these three MPMs exhibit a wide range of timescales, consisting of subseasonal and interannual components (Figure S2). Figure 2d shows the correlation coefficient of the time series between observations and hindcasts as a function of forecast lead times. Here the observational time series are obtained by projecting observed SAT anomalies onto the spatial patterns of these three modes. For the first two modes, the correlation coefficients decrease gradually but remain significant at nearly 7 weeks lead time at the 5% significance level (Figure 2d). For the third mode (the EMD mode), the significant correlation skill extends to 5 (6) weeks at the 5% (10%) level.

Compared to the first three modes, the other modes (modes 4 to 15) have a shorter timescale of variability. The modes 4–8 primarily represent the middle to high latitude wave train pattern or the teleconnection pattern associated with the tropical convective forcing (Figure S3), that can be skillfully predicted at 3 weeks lead time at the 5% significance level. The other higher frequency modes (9–15) can only be skillfully predicted at 2 weeks lead time. Another noteworthy feature is that the relative contributions of these eight modes on the regional SAT prediction differ dramatically. For example, for the Eurasia continent, most of the prediction skill originates from the NAO and the EMD modes, while for North America, the modes 4–

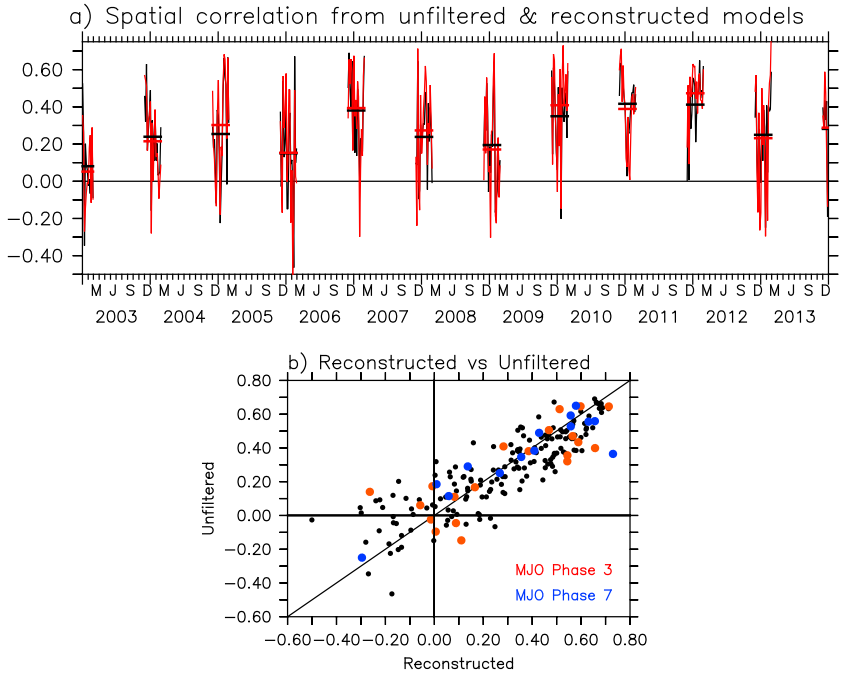

Figure 3. (a) The time evolution of the spatial correlation skill over the Northern Hemisphere extratropical land region (30–90°N) from unfiltered model hindcasts (black) and the reconstructed model using the eight most predictable modes (red). The corresponding wintertime mean skill is shown as the black and red bars. (b) The scatter diagram of spatial correlation skill from the reconstructed model (*x* axis) versus the unfiltered model (*y* axis). Red (blue) dots are the correlation skill with the initially strong MJO events in phase 3 (7). All results are at week 3. MJO = Madden-Julian Oscillation.

8 are also important. This is inferred by comparing the results with the reconstructed model from the first three MPMs (Figure 1 vs Figure S4).

The subseasonal forecast skill also depends on forecast initial conditions. Figure 3a shows the time evolution of the spatial correlation skill over the Northern Hemisphere extratropical land region (30–90°N) for the 3-week lead forecast. The prediction skill experiences prominent fluctuations at the interannual and intraseasonal timescales. During some years (e.g., 2006–2007, 2009–2010, 2010–2011, 2011–2012, 2012–2013), the prediction skill is generally higher than the other years. The reconstructed model largely captures the interannual fluctuations of the prediction skill. Figure 3b shows the scatter plot of the spatial correlation skill (over the Northern Hemisphere extratropical land region) between the reconstructed model and the unfiltered model. The unfiltered model forecast generally outperforms the reconstructed forecast particularly when the prediction skill is relatively low. For the cases with higher prediction skill ($r > 0.3$ for both the unfiltered and filtered forecasts), the reconstructed model displays a slightly higher skill compared to the unfiltered model (0.52 vs 0.47). Understanding the cause of this fluctuation is of importance in order to identify the time window of opportunity for a skillful subseasonal forecast. Compared to the cases with higher week 3 spatial correlation skill, the cases with lower spatial correlation skill does show a symmetrically lower prediction skill (at week 1–5) for the first three MPMs (not shown).

Characteristics of precursor signals for these modes are vital to understand their predictability sources. Figure 4 shows that the observational stratospheric polar vortex strength (1 week before forecast) is significantly correlated (at 5% significance level) with the time series of MPM2/MPM3 at week 3 from model hindcasts. Here the stratospheric polar vortex index is used to measure the strength of polar vortex, defined as the zonal mean zonal wind at 50 hPa averaged over 60–90°N (Jia et al., 2017). Actually, significant correlations are also found from week 1 to week 7 for both modes. Note that other modes are not significantly correlated with the observational stratospheric polar vortex index. It suggests that the predictability sources for both modes (NAO and EDM) originate, at least partially, from the stratosphere, highlighting the importance of the stratosphere in S2S prediction.

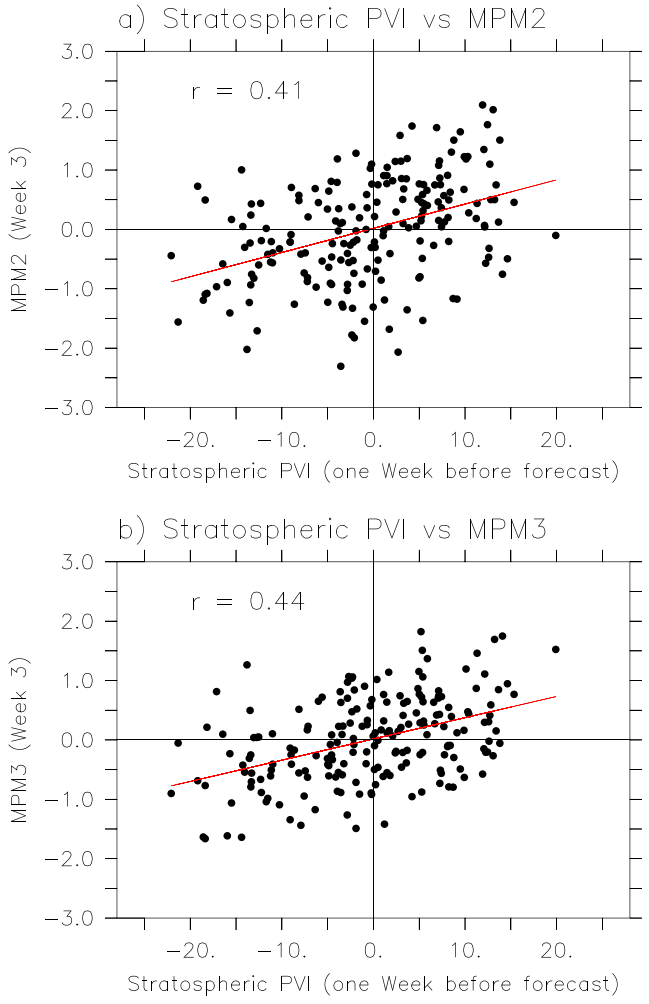

Figure 4. The scatter diagram between the stratospheric PVI at 1 week before the forecast, and the time series of (a) MPM2 at week 3, (b) MPM3 at week 3. The corresponding correlation coefficients, significant at 5% significance level for both modes, are shown in the upper-left corner. Note that the time series of MPM2 and MPM3 are nearly orthogonal. The stratospheric PVI is defined as the zonal mean zonal wind (m/s) at 50 hPa averaged over 60–90°N. MPM = most predictable mode; PVI = polar vortex index.

3.3. MJO and SAT Prediction

As the dominant tropical intraseasonal mode, the MJO exerts its impacts not only in the tropics but also in the extratropics through the excitation of poleward propagating Rossby waves (e.g., Donald et al., 2006; Moon et al., 2011). More importantly, the MJO serves as one of the crucial predictability sources to bridge the forecast gap between the weather and seasonal predictions.

The relative high MJO prediction skill in the wintertime (about 17–31 days) in current dynamic models (e.g., Kim et al., 2014; Neena et al., 2014; Vitart, 2017), measured by the bivariate correlation of the Real-time Multivariate MJO (RMM) index (Wheeler & Hendon, 2004), makes it possible to examine its direct impacts on the predictions of tropical cyclones and other fields. Indeed, dynamical model results demonstrate that the presence of MJO substantially influences the prediction of tropical cyclone genesis (e.g., Chen & Lin, 2011; Jiang et al., 2018; Lee et al., 2018; Xiang et al., 2014) as well as the extratropical NAO prediction (e.g., Lin, Brunet, & Fontecilla Juan, 2010; Vitart, 2017).

For this GFDL model forecast system, wintertime MJO prediction skill is about 27 days and even beyond 40 days for some specific phases (Xiang et al., 2015). Figure 5 features composite results of anomalous SAT, H500, and tropical precipitation 3 weeks after initially strong MJO (RMM $>$ 1.0) in phases 3 and 7. We focus on phases 3 and 7 because the associated anomalous convection featuring an east-west dipole pattern between the Indian Ocean and the western Pacific is particularly effective at inducing extratropical circulation anomalies (Lin, Brunet, & Mo, 2010). Consistent with a skillful MJO prediction (Xiang et al., 2015), the tropical precipitation anomalies are well predicted for both phases (Figure 5). The week 3 composite H500 anomalies following MJO phase 3 (7) projects onto the positive (negative) phase of NAO starting from MJO phase 3 (7), which is consistent with previous work (Cassou, 2008; Lin et al., 2009; Riddle et al., 2013). Note that the temperature anomalies have the largest magnitude over North America and northern Eurasia, showing a dynamically coherent pattern with circulation change. For initial MJO phase 7, the predicted SAT and H500 anomalies are somewhat weaker than the observational composites. Similar results are seen for the composite results after subtracting the interannual signals (Figure S5), confirming the robustness of the results.

However, the contribution of MJO on the Northern Hemispheric SAT prediction is likely secondary. Although the composite results possess strikingly similar patterns between hindcasts and observations (Figure 5), the spatial correlation skill of the land SAT anomalies in the Northern Hemisphere extratropics is rather scattered for those initialized with strong MJO in phases 3 and 7 (Figure 3b). For both phases, the mean correlation skill is not statistically distinguishable from the average of all events. All other phases starting with strong MJO events are also examined while they do not appear to differ significantly from the inactive MJO cases in terms of the SAT prediction skill. This can be understood from several different perspectives. First, the MJO-induced SAT anomalies are relatively weak with its magnitude only about half of the total variance (Figure 5 vs Figure S1). This is in accord with the relatively weak NAO response forced by MJO (about half of the total variance of NAO; Vitart, 2017). Second, the impacts from MJO are confined to a limited area including Canada and the Arctic region for both phases (Figure 5). Third, the mean state bias, such as the jet location in the Northern Hemisphere, may result in biased extratropical teleconnections even with the "correct" tropical forcing due to MJO (e.g., Henderson et al., 2017). Therefore, a skillful prediction of predictors does not necessarily entail a skillful prediction of its climate impacts.

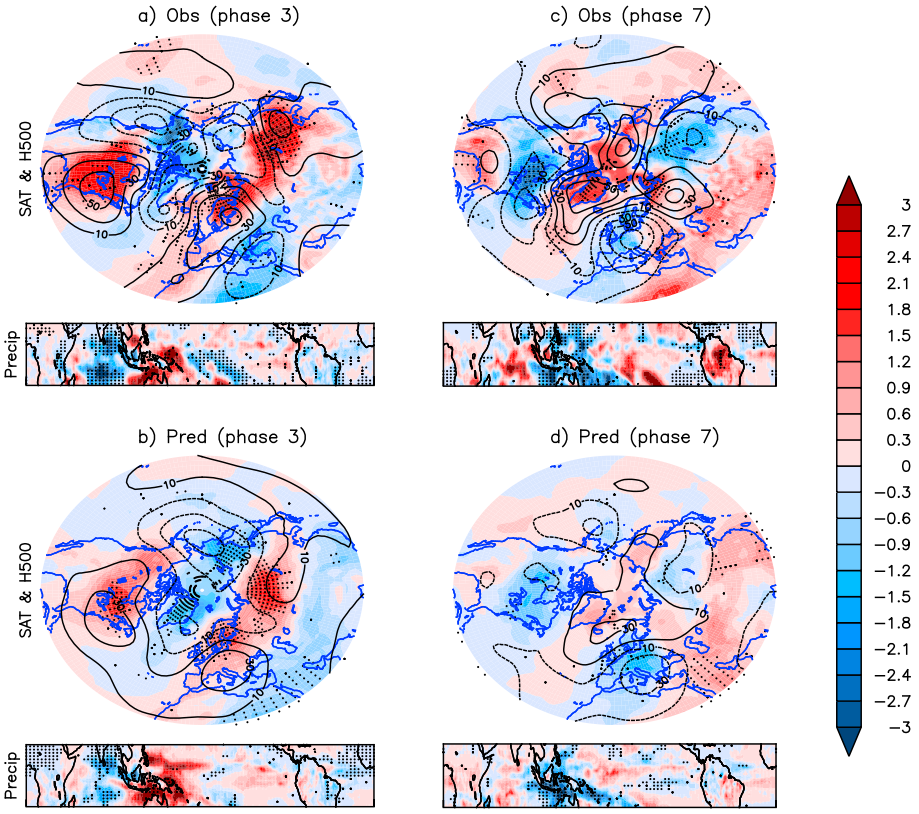

Figure 5. (a) Upper panel: Composite results of observational SAT (shading; °C) and H500 anomalies (contours; m) 3 weeks after an initially strong MJO in phase 3. Stippling represents the regions with more than 70% of events with the same sign as the composite mean for SAT anomalies. Lower panel: Similar to upper panel but for the composite precipitation anomalies (mm/day). (b) is similar to (a) but for model hindcast results. (c) and (d) are similar to (a) and (b) but for composite results 3 weeks after the initially strong MJO in phase 7. Note that 20 (14) MJO events are found for the composite analysis for MJO phase 3 (7). SAT = surface air temperature; MJO = Madden-Julian Oscillation.

4. Summary and Discussion

Using a GFDL coupled model forecast system, the subseasonal SAT prediction performance is examined based on wintertime hindcasts covering the 2003–2013 period. Results show a significant correlation skill of weekly mean SAT anomalies over a large portion of land domains for the week 3–5 forecast. The forecast skill primarily originates from eight MPMs with the first three as the ENSO mode, the NAO mode, and the EMD mode.

The potential contribution of MJO to the subseasonal SAT prediction in the Northern Hemisphere extratropical land region (30–90°N) is moderate and regional. The prediction skill varies from week to week and year to year. Thus, it is thus critical to identify the time window of opportunity for subseasonal prediction during which the forecast skill is relatively high.

At the seasonal timescale, the prediction skill of the NAO is strongly influenced by stratospheric conditions related to the polar vortex strength (e.g., Jia et al., 2017; Scaife et al., 2015; Wang et al., 2017). Here it is found that the NAO and the EMD modes are both strongly correlated with the stratospheric polar vortex strength from the initial condition. It suggests that the stratospheric variability is an important component of subseasonal SAT prediction, consistent with Tripathi Om et al. (2014).

Note that the overall subseasonal prediction skill is relatively low despite of the significant correlation skill over a large portion of the land domains (Figure 1). Here the land model is not initialized although the land tends to be adjusted accordingly with atmospheric nudging. Realistic land model initialization (such as snow cover and snow depth) has been documented to play a principal role in the subseasonal prediction during the wintertime (Orsolini et al., 2013). This strengthens the need for understanding and quantifying the role of

land initialization in subseasonal SAT prediction. Further improvement of dynamic models and the initialization provides promise for the subseasonal prediction.

There are some caveats in this study. First, there is only one model used here. Whether the identified modes are present in other model forecast systems remains to be seen. Second, the current study covers a limited time sample (11 years). The amplitude and phase of NAO are sensitive to the period chosen. The MJO connections to high-latitude SAT rely on the background mean state that may differ among different periods. Thus, whether the current results are sensitive to the period chosen is also an open question. These questions can be addressed by examining the WWRP/WCRP S2S data sets with more models and longer records.

Acknowledgments
We are grateful for the insightful comments from Fanrong Zeng and Yongqiang Sun and two anonymous reviewers. X.J. acknowledges support by the NOAA Climate Program Office under awards NA15OAR4310177 and NA17OAR4310261. The GFS analysis data are available at https://www.ncdc.noaa.gov/data-access/model-data/model-data sets/global-forcast-system-gfs. The NOAA OISST data are available at https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html. The GPCP data can be downloaded from https://rda.ucar.edu/datasets/ds728.3/. The model hindcast data are obtained from Xiang et al. (2015).

References

Black, J., Johnson, N. C., Baxter, S., Feldstein, S. B., Harnos, D. S., & L'Heureux, M. L. (2017). The predictors and forecast skill of northern hemisphere teleconnection patterns for Lead times of 3–4 weeks. *Monthly Weather Review*, *145*(7), 2855–2877. https://doi.org/10.1175/MWR-D-16-0394.1

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., & Bladé, I. (1999). The effective number of spatial degrees of freedom of a time-varying field. *Journal of Climate*, *12*(7), 1990–2009. https://doi.org/10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2

Butler, A. H., Polvani, L. M., & Deser, C. (2014). Separating the stratospheric and tropospheric pathways of El Niño-Southern Oscillation teleconnections. *Environmental Research Letters*, *9*. https://doi.org/10.1088/1748-9326/9/2/024014

Cassou, C. (2008). Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Nature*, *455*, 523–527. https://doi.org/10.1038/nature07286

Chen, J.-H., & Lin, S.-J. (2011). The remarkable predictability of inter-annual variability of Atlantic hurricanes during the past decade. *Geophysical Research Letters*, *38*, L11804. https://doi.org/10.1029/2011GL047629

DelSole, T., & Tippett, M. K. (2009a). Average predictability time. Part I: Theory. *Journal of the Atmospheric Sciences*, *66*(5), 1172–1187. https://doi.org/10.1175/2008JAS2868.1

DelSole, T., & Tippett, M. K. (2009b). Average predictability time. Part II: Seamless diagnoses of predictability on multiple time scales. *Journal of the Atmospheric Sciences*, *66*(5), 1188–1204. https://doi.org/10.1175/2008JAS2869.1

DelSole, T., Trenary, L., Tippett, M. K., & Pegion, K. (2017). Predictability of week-3–4 average temperature and precipitation over the contiguous United States. *Journal of Climate*, *30*, 3499–3512. https://doi.org/10.1175/JCLI-D-16-0567.1

Donald, A., Meinke, H., Power, B., Maia Aline de, H. N., Wheeler Matthew, C., White, N., et al. (2006). Near-global impact of the Madden-Julian Oscillation on rainfall. *Geophysical Research Letters*, *33*, L09704. https://doi.org/10.1029/2005GL025155

Graf, H.-F., & Zanchettin, D. (2012). Central Pacific El Niño, the "subtropical bridge," and Eurasian climate. *Journal of Geophysical Research*, *117*, D01102. https://doi.org/10.1029/2011JD016493

Henderson, S. A., Maloney, E. D., & Son, S.-W. (2017). Madden–Julian Oscillation Pacific teleconnections: The impact of the basic state and MJO representation in general circulation models. *Journal of Climate*, *30*(12), 4567–4587. https://doi.org/10.1175/JCLI-D-16-0789.1

Huffman, G. J., Bolvin, D. T., & Adler, R. F. 2016. GPCP version 1.2 one-degree daily precipitation data set. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6D50K46

Jia, L., Yang, X., Vecchi, G., Gudgel, R., Delworth, T., Fueglistaler, S., et al. (2017). Seasonal prediction skill of northern extratropical surface temperature driven by the stratosphere. *Journal of Climate*, *30*, 4463–4475. https://doi.org/10.1175/JCLI-D-16-0475.1

Jia, L., Yang, X., Vecchi, G. A., Gudgel, R. G., Delworth, T. L., Rosati, A., et al. (2015). Improved seasonal prediction of temperature and precipitation over land in a high-resolution GFDL climate model. *Journal of Climate*, *28*(5), 2044–2062. https://doi.org/10.1175/JCLI-D-14-00112.1

Jiang, X., Xiang, B., Zhao, M., Li, T., Lin, S.-J., Wang, Z., & Chen, J.-H. (2018). Intraseasonal tropical cyclogenesis prediction in a global coupled model system. *Journal of Climate*, *31*, 6209–6227. https://doi.org/10.1175/JCLI-D-17-0454.1

Johnson, N. C., Collins, D. C., Feldstein, S. B., L'Heureux, M. L., & Riddle, E. E. (2013). Skillful wintertime North American temperature forecasts out to 4 weeks based on the state of ENSO and the MJO. *Weather and Forecasting*, *29*(1), 23–38.

Kim, H.-M., Webster, P., Toma, V. E., & Kim, D. (2014). Predictability and prediction skill of the MJO in two operational forecasting systems. *Journal of Climate*, *27*, 5364–5378. https://doi.org/10.1175/JCLI-D-13-00480.1

Kirtman, B. P., Min, D., & Infanti, J. M. (2013). The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bulletin of the American Meteorological Society*, *95*(4), 585–601.

Lee, C.-Y., Camargo, S. J., Vitart, F., Sobel, A. H., & Tippett, M. K. (2018). Subseasonal tropical cyclone genesis prediction and MJO in the S2S dataset. *Weather and Forecasting*, *33*(4), 967–988. https://doi.org/10.1175/WAF-D-17-0165.1

Lin, H. (2018). Predicting the dominant patterns of subseasonal variability of wintertime surface air temperature in extratropical northern hemisphere. *Geophysical Research Letters*, *45*, 4381–4389. https://doi.org/10.1029/2018GL077509

Lin, H., Brunet, G., & Derome, J. (2009). An observed connection between the North Atlantic Oscillation and the Madden–Julian Oscillation. *Journal of Climate*, *22*(2), 364–380. https://doi.org/10.1175/2008JCLI2515.1

Lin, H., Brunet, G., & Fontecilla Juan, S. (2010). Impact of the Madden-Julian Oscillation on the intraseasonal forecast skill of the North Atlantic Oscillation. *Geophysical Research Letters*, *37*, L19803. https://doi.org/10.1029/2010GL044315

Lin, H., Brunet, G., & Mo, R. (2010). Impact of the Madden-Julian Oscillation on wintertime precipitation in Canada. *Monthly Weather Review*, *138*, 3822–3839. https://doi.org/10.1175/2010MWR3363.1

Moon, J.-Y., Wang, B., & Ha, K.-J. (2011). ENSO regulation of MJO teleconnection. *Climate Dynamics*, *37*(5-6), 1133–1149. https://doi.org/10.1007/s00382-010-0902-3

Neena, J. M., Lee, J. Y., Waliser, D., Wang, B., & Jiang, X. (2014). Predictability of the Madden–Julian Oscillation in the intraseasonal variability hindcast experiment (ISVHE). *Journal of Climate*, *27*(12), 4531–4543. https://doi.org/10.1175/JCLI-D-13-00624.1

Orsolini, Y. J., Senan, R., Balsamo, G., Doblas-Reyes, F. J., Vitart, F., Weisheimer, A., et al. (2013). Impact of snow initialization on subseasonal forecasts. *Climate Dynamics*, *41*(7-8), 1969–1982. https://doi.org/10.1007/s00382-013-1782-0

Polvani, L. M., Sun, L., Butler, A. H., Richter, J. H., & Deser, C. (2017). Distinguishing stratospheric sudden warmings from ENSO as key drivers of wintertime climate variability over the North Atlantic and Eurasia. *Journal of Climate*, *30*, 1959–1969. https://doi.org/10.1175/JCLI-D-16-0277.1

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, *20*(22), 5473–5496. https://doi.org/10.1175/2007JCLI1824.1

Riddle, E. E., Stoner, M. B., Johnson, N. C., L'Heureux, M. L., Collins, D. C., & Feldstein, S. B. (2013). The impact of the MJO on clusters of wintertime circulation anomalies over the North American region. *Climate Dynamics*, *40*, 1749–1766. https://doi.org/10.1007/s00382-012-1493-y

Scaife, A. A., Karpechko, A. Y., Baldwin, M. P., Brookshaw, A., Butler, A. H., Eade, R., et al. (2015). Seasonal winter forecasts and the stratosphere. *Atmospheric Science Letters*, *17*(1), 51–56.

Tripathi Om, P., Baldwin, M., Charlton-Perez, A., Charron, M., Eckermann, S. D., Gerber, E., et al. (2014). The predictability of the extratropical stratosphere on monthly time-scales and its impact on the skill of tropospheric forecasts. *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 987–1003.

Vecchi, G. A., Delworth, T., Gudgel, R., Kapnick, S., Rosati, A., Wittenberg, A. T., et al. (2014). On the seasonal forecasting of regional tropical cyclone activity. *Journal of Climate*, *27*(21), 7994–8016. https://doi.org/10.1175/JCLI-D-14-00158.1

Vitart, F. (2017). Madden—Julian Oscillation prediction and teleconnections in the S2S database. *Quarterly Journal of the Royal Meteorological Society*, *143*(706), 2210–2220. https://doi.org/10.1002/qj.3079

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., et al. (2016). The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, *98*(1), 163–173.

Wang, L., Ting, M., & Kushner, P. J. (2017). A robust empirical seasonal prediction of winter NAO and surface climate. *Scientific Reports*, *7*(1), 279. https://doi.org/10.1038/s41598-017-00353-y

Wheeler, M. C., & Hendon, H. H. (2004). An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Monthly Weather Review*, *132*(8), 1917–1932. https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2

Xiang, B., Lin, S.-J., Zhao, M., Zhang, S., Vecchi, G., Li, T., et al. (2014). Beyond weather time-scale prediction for Hurricane Sandy and Super Typhoon Haiyan in a global climate model. *Monthly Weather Review*, *143*(2), 524–535.

Xiang, B., Wang, B., & Li, T. (2013). A new paradigm for the predominance of standing Central Pacific warming after the late 1990s. *Climate Dynamics*, *41*(2), 327–340. https://doi.org/10.1007/s00382-012-1427-8

Xiang, B., Zhao, M., Jiang, X., Lin, S.-J., Li, T., Fu, X., & Vecchi, G. (2015). The 3–4-week MJO prediction skill in a GFDL coupled model. *Journal of Climate*, *28*(13), 5351–5364. https://doi.org/10.1175/JCLI-D-15-0102.1

Yang, X., Rosati, A., Zhang, S., Delworth, T. L., Gudgel, R. G., Zhang, R., et al. (2013). A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. *Journal of Climate*, *26*(2), 650–661. https://doi.org/10.1175/JCLI-D-12-00231.1

Zhao, M., Golaz, J. C., Held, I. M., Guo, H., Balaji, V., Benson, R., et al. (2018). The GFDL global atmosphere and land model AM4.0/LM4.0: 1. Simulation characteristics with prescribed SSTs. *Journal of Advances in Modeling Earth Systems*, *10*, 691–734. https://doi.org/10.1002/2017MS001208