

Evaluating Medium-Range Forecast Performance of Regional-Scale Circulation Patterns

OMON A. OBAREIN¹,^a CAMERON C. LEE,^a ERIK T. SMITH,^b AND SCOTT C. SHERIDAN^a

^a *Department of Geography, Environmental Science and Design Research Institute, ClimRISE Laboratory, Kent State University, Kent, Ohio*

^b *Electric Power Research Institute, Washington, D.C.*

(Manuscript received 16 August 2022, in final form 5 May 2023, accepted 17 May 2023)

ABSTRACT: Accurate subseasonal-to-seasonal (S2S) weather forecasts are crucial to making important decisions in many sectors. However, significant gaps exist between the needs of society and what forecasters can produce, especially at weekly and longer lead times. We hypothesize that by clustering atmospheric states into a number of predefined categories, the noise can be reduced and, consequently, medium-range forecasts can be improved. Self-organizing map (SOM)-based clustering was used on daily mean sea level pressure (MSLP) data from the North American Regional Reanalysis to categorize the synoptic-scale circulation for eastern North America from 1979 to 2016 into 28 discrete patterns. Then, using two goodness-of-fit metrics, the relative skill of four different forecasting methods over a 90-day lead time was studied: 1) a circulation pattern (CP) forecast, 2) raw forecast output from the Climate Forecast System (CFS) operated by the National Centers for Environmental Prediction (NCEP), 3) a simple climatology forecast, and 4) a simple persistence forecast. As expected, forecast skill of both the CP forecast and the raw CFS forecast generally decreased rapidly from the first day, coming to parity with the skill of climatology after 10–12 days when using correlation, and at 7–10 days when using the root-mean-square error (RMSE). Most importantly, this study found that the CP forecast was the most skillful forecast method over the 8–11-day lead time when using RMSE. On a spatial basis, the skill of the CP forecast and the raw CFS decreases latitudinally from north to south. This study thus demonstrates the potential utility of categorical or circulation pattern-based forecasting at 1–2-week lead times.

KEYWORDS: Ensembles; Forecast verification/skill; Forecasting techniques; Numerical weather prediction/forecasting; Short-range prediction

1. Introduction

Many industries and sectors of society use guidance provided by subseasonal-to-seasonal (S2S) weather forecasts, which cover the period between 2 weeks and several months (He et al. 2021), to inform decisions. The ability to make informed decisions regarding changes in the risk of extreme weather events (e.g., floods, drought), resource management, agriculture, forestry, public health, and energy, among others, lies in effective S2S forecasts (Craig et al. 2021; Vitart et al. 2017; He et al. 2021; Hwang et al. 2019). However, significant gaps exist between the needs of society and what forecasters can produce, especially over longer lead times (White et al. 2017; Vitart et al. 2015). Thus, collaborative efforts to improve the quality of S2S forecasts have increased rapidly in the past decade (Brunet et al. 2010; Robertson et al. 2015).

Evaluation of forecast model performance is crucial in improving S2S weather forecasts. Atmospheric scientists have sought to improve the accuracy of forecast models by using verification methods that are fundamental to determining a model's skill (Robertson and Vitart 2018). Near-term climate

predictions are verified by performing many past predictions called hindcasts or reforecasts (Doblas-Reyes et al. 2013). The National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFS) carried out a comprehensive reforecast from 1982 to 2011 using initial conditions from the NCEP reanalysis dataset from 1979 to 2011 (Saha et al. 2014). The aim of the reforecasts was to obtain consistent calibrations, better bias diagnosis, and skill estimates of S2S forecasting with the CFS, thereby improving forecast skill (Hamill et al. 2006).

Following the CFS, many studies have used reforecasts to evaluate and recalibrate forecast models of various climate and environmental variables. For example, reforecasts have been used to verify the skill with which reanalysis products reproduce soil moisture conditions and accurately model land-atmosphere interactions (Dirmeier 2013; Prodhomme et al. 2016). The predictive skill of summer monsoon precipitation (Zuo et al. 2013), the sensitivities of surface air temperature predictions to the uncertainty in initial conditions of the land surface (Shin et al. 2020), and the prediction of wintertime Arctic Oscillation (Riddle et al. 2013), among others, have all utilized reforecast datasets. One of the most common methods of inferring forecast skill is by comparing its mean error or correlation to that of a persistence forecast or a climatology forecast (Li and Robertson 2015; Walsh 1984), the latter of which generally becomes most skillful at around 10–17 days of lead time (Vitart 2014).

There are several reasons why S2S forecasts have proven unskillful after 1–2 weeks despite advances in numerical

¹ Denotes content that is immediately available upon publication as open access.

Corresponding author: Omon A. Obarein, oobarein@kent.edu

weather prediction (Wedam et al. 2009; Huang et al. 2020). Primary among them are the compounding effects of small imprecisions in the initial conditions [i.e., the so-called butterfly effect (Lorenz 1963)], which can lead to significant errors in the forecast model, due to the chaotic nature of the atmosphere. This problem is best mitigated through ensemble forecasting (Robertson and Vitart 2019; Wilks 2020), whereby a finite ensemble of initial states is first chosen (rather than the single observed initial state) based on the sampling of uncertainties related to observational errors, and then each member of the ensemble is integrated forward in time according to the governing equations of the forecast model (Vannitsem et al. 2018). The spread of the ensemble member forecasts at the final forecast lead time provides the range of possible future states of the atmosphere and the uncertainty in the ensemble forecast. The probability of occurrence of individual ensemble forecasts can thus be evaluated, and the ensemble mean of these forecasts reflects the actual level of forecast skill (Wilks and Vannitsem 2018; Robertson and Vitart 2019).

In addition to sensitive dependence on initial conditions, the accuracy of S2S forecasts is influenced by the spatiotemporal collinearity in gridded fields of weather variables assimilated into forecast models, including the CFS (Sheridan and Lee 2011; Dormann et al. 2013). Data collinearity can lead to less-skilled forecasts (Sousa et al. 2007; Dormann et al. 2013) because the high level of unpredictable noise makes atmospheric models too nebulous to capture the underlying atmospheric condition, thus making it difficult to forecast. Finally, our inability to fully represent the equations that govern weather (Ferranti et al. 2015) and compounding errors introduced by the necessity of parameterizing nontrivial sub-gridscale processes (Shutts and Pallarès 2014) also contributes to larger errors in S2S forecast models. A working hypothesis of the present study is that by classifying atmospheric states into discrete categories so that all categories represent the whole spectrum of atmospheric variability, the noise can be reduced, and the predictable signals amplified. While classifying and identifying typically occurring atmospheric circulation patterns (CPs) have been done extensively in prior synoptic climatological research, few studies have focused on the potential skillfulness of forecasting predefined CPs and comparing such a forecast's performance to that of raw output from a numerical forecast model.

Synoptic CPs or weather patterns have proven useful in understanding the linkages of local surface weather characteristics and environmental outcomes to specific sets of atmospheric states—a fundamental aim of synoptic climatology (Yarnal 1993; Yarnal et al. 2001; Sheridan and Lee 2011; Gilabert and Llasat 2018). For example, synoptic CPs have been shown to significantly associate with weather and surface characteristics such as tornadoes, hurricanes, and forest fires (Tochimoto 2022; Lee 2012; Bagaglini et al. 2021); pollution and air quality (Sheridan et al. 2013; Zhou et al. 2018); and coastal flooding and sea level fluctuations (Sheridan et al. 2017, 2019; Pirhalla et al. 2022; Neal et al. 2018). Synoptic CPs have also been useful in understanding extreme temperature, precipitation, and dewpoint events (Adams et al. 2021; Richardson et al. 2020a; Neal et al. 2020; Mastrantonas et al. 2021; Lee et al. 2021), and especially their links to or influence on human health outcomes such as

mortality and morbidity (Huang et al. 2020; García-Herrera et al. 2022). Thus, in addition to potentially improving weather forecasts at weekly and longer lead times, since these numerous other human and environmental outcomes are influenced by CPs, forecasting predefined CPs may also lead to more skillful predictions of these outcomes.

Some recent studies have demonstrated the utility of forecasting predefined CPs or weather patterns. For example, the Met Office's 30 weather patterns, defined from the *k*-means clustering of daily mean sea level pressure (MSLP), have been applied to forecasting associated climate variables and environmental outcomes like flooding (Neal et al. 2018; Richardson et al. 2020a), precipitation and drought (Richardson et al. 2020b), volcanic ash flow (Harrison et al. 2022), and lightning activity (Wilkinson and Neal 2021), among others. In all these studies, ensemble members are assigned to the closest matching weather pattern for each day in the forecast and grid point (Neal et al. 2016; Richardson et al. 2020a). This ensemble prediction approach ensures that a probabilistic forecast output can be produced.

In this study, the ability of a single numerical model (the CFS) to forecast the correct predefined categorical circulation pattern, as compared to the observed continuous gridded field, was examined using two goodness-of-fit metrics. As mentioned above, fewer categorical abstractions of the weather have a higher signal to noise ratio than raw or continuous observations, so we hypothesize that forecast skill may be extended over the current limit by a few more days using the CP-based method. Since the aim here is a proof of concept to determine the relative skill of a self-organizing map (SOM)-based CP forecasting method, we chose to use a single observed initial state (as opposed to an ensemble) for a single variable (MSLP) over a single region to simplify the analysis. The performance of the CP forecast against the raw CFS forecasts, simple climatology forecasts, and simple persistence forecasts over a 90-day period was evaluated.

2. Data and methods

a. Study area and data

Daily-scale MSLP from the North American Regional Reanalysis (NARR; Mesinger et al. 2006) was used to categorize the synoptic-scale circulation for eastern North America from January 1979 through December 2016. Because categorizing synoptic-scale patterns across too large a region may negatively impact classifications (Sheridan et al. 2019), the study area is subcontinental (Fig. 1). In addition, daily MSLP from the Climate Forecast System (CFS) Reanalysis (CFSR), and from CFS, version 2 (CFSv2), 9-month operational forecasts from April 2011 through December 2016 (Saha et al. 2014) were also acquired. Since the spatial resolution of the CFS ($0.5^\circ \times 0.5^\circ$) is different than that of the NARR ($\sim 0.3^\circ \times 0.3^\circ$), all CFS data were spatially interpolated to the NARR grid by using a Delaunay triangulation, which tends to be more efficient and less prone to artifacts than other interpolation methods (Smith et al. 2021; Sheridan et al. 2019; Amidror 2002). As this study's interest is primarily subseasonal, only the first



FIG. 1. The eastern North American region used in the SOM-based atmospheric pattern classifications.

90 days of the 9-month CFS operational forecast data were used. All CFSv2 data were bias corrected by subtracting the mean 1981–2010 difference between CFSR and NARR data from each grid point for every day in the CFSv2 dataset.

b. SOM methodology

Compared to more traditional clustering or classification strategies, the use of SOMs in synoptic climatological research has become widespread over the past decade or longer (e.g., Jaye et al. 2019; Berkovic et al. 2021; Smith et al. 2021). A SOM is an unsupervised, neural network clustering methodology that projects high-dimensional input data onto a two-dimensional space in order to unearth patterns (Hewitson and Crane 2002; Kohonen 1990). One unique aspect of the SOM methodology is that, unlike other clustering algorithms (e.g.,

EOFs, *k*-means, etc.) that have discrete patterns, SOMs can organize atmospheric circulation into a continuum of synoptic classes by locating nodes that span the entire climatic data space. This process creates intuitive pattern visualizations of atmospheric states and of the relationship between the nodes (Hewitson and Crane 2002; Solidoro et al. 2007; Liu and Weisberg 2011; Sheridan and Lee 2011; Sheridan et al. 2017).

In this study, SOMs were used to classify daily MSLP CPs for the study area from 1979 to 2016, using NARR data. A SOM consists of an array of nodes that spans the two primary dimensions of the data space. The total number of nodes, or patterns, is determined by the number of patterns in the horizontal and vertical dimensions. Figure 2 shows the MSLP nodes arrayed in a continuum with high pressure dominant patterns or nodes, mostly over land, aligned on one side, while low pressure dominant patterns, especially over the North Atlantic, are aligned on the other. The transitional nodes, where moderately high and low pressure patterns alternate between land and ocean, lie within the SOMs space.

Each day in the dataset was assigned to one of 28 SOM nodes based on the similarity of the MSLP pattern on that day with all other days. The final number of nodes was determined by examining multiple different SOM sizes or dimensions to identify the SOM that best partitioned atmospheric variability in the region's domain. The optimal SOM dimensions were determined using cluster-validation metrics like the Davies–Bouldin index (Davies and Bouldin 1979) and the distributed variability skill score (Lee 2017) that quantified the within-node and across-node variability, with the goal of minimizing the within-node variability while maximizing the variability across nodes (Yarnal 1993; Hewitson and Crane

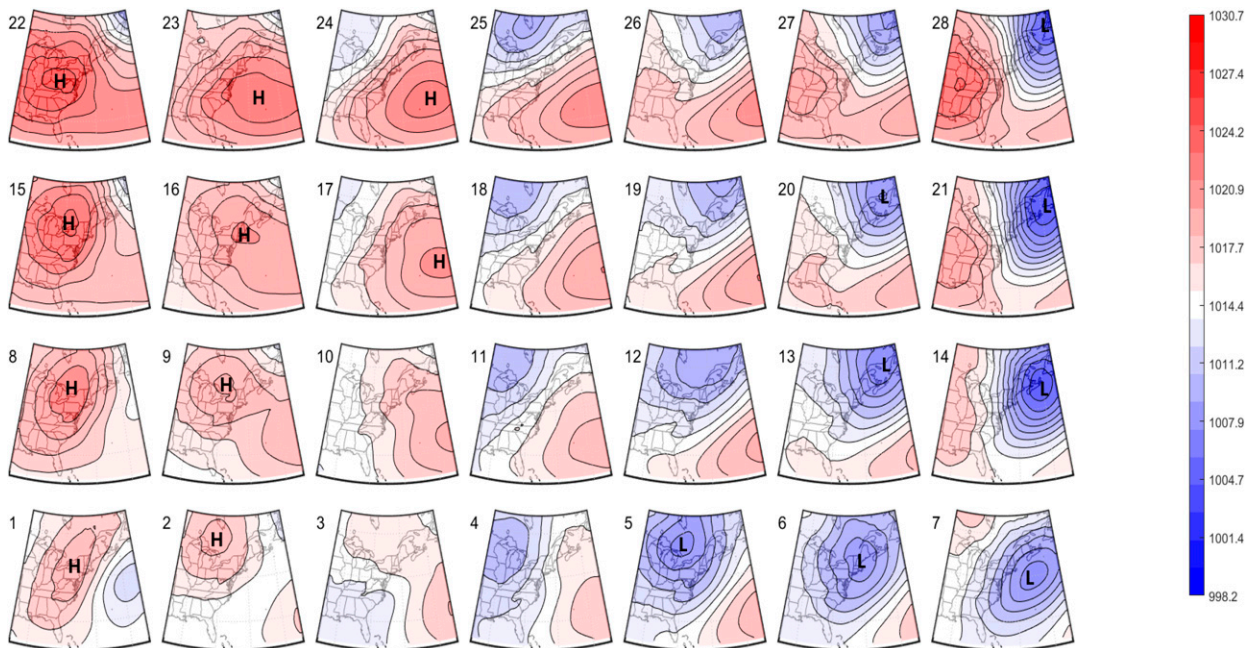


FIG. 2. SOM of MSLP for the eastern North American region (NARR: 1979–2016) plotted as filled contours (hPa). Modified from Sheridan et al. (2019).

2002). Before the SOM-based clustering, the raw MSLP data were standardized and subjected to an *S*-mode principal component analysis (PCA). PCA reduces the computational time (via dimension reduction) and helps mitigate spatial collinearity before clustering (Smith et al. 2021). All principal components (PCs) with eigenvalues greater than one were retained and used as inputs into the SOM neural network training.

Once historical or observed (NARR) CPs are defined, then, to determine the ability of the CFS to forecast SOM-based patterns, the squared Euclidean distance between the raw CFS output (for each lead time and initialization date) and the 28 CPs' averaged MSLP field was calculated. The CP or node with the smallest distance was selected as the best matching pattern (BMP), and then the actual CP forecast was simply the mean NARR MSLP field for that pattern (averaged over the 1981–2010 period). For example, if there were 100 days classified into pattern 3, then the MSLP at every grid point on those 100 days is averaged to get the mean NARR MSLP field for that pattern. Then, for any day from the CFSv2 model output with a BMP of pattern 3, the mean MSLP field is used as the actual forecast (instead of the CFSv2 model output of the MSLP field). If two (or more) patterns had distances smaller than a standard radial distance (SRD) of 40.29 (or one standard deviation unit of distance between all cluster nodes) from the BMP, then the forecast was a combined mean MSLP field of all the CPs falling within this distance. The idea is, if a forecasted MSLP field is well defined by a single pattern, then only that pattern will be in the forecast, but if the forecasted MSLP field resembles two or more different patterns (i.e., it is “borderline”), then the forecast will reflect that. This process is repeated for each lead time (1–90 days) and each initialization day ($n = 2101$) from April 2011 to December 2016.

c. Forecast model comparison and evaluation

Using root-mean-square error (RMSE) and Pearson's correlation (r) metrics, the skill of the 90-day CFS CP forecast model to reproduce NARR data was calculated for all 90 forecast lead days throughout the study period (2101 days, from April 2011 to December 2016). This 90-day CFS CP forecast skill was also compared to that of three other models: the skill of the bias-corrected CFS forecast MSLP model output (hereafter referred to as raw CFS), the seasonal MSLP climatology, and simple persistence. The seasonal climatology was calculated from the MSLP NARR data from 1981 to 2010. Persistence, though not used as often as climatology, is a good reference to measure shorter-term forecast skill (Lipperheide et al. 2015; Jacox et al. 2019, and others). Persistence was calculated from the NARR data and is the MSLP NARR value on the day of forecast initialization held constant throughout the 90 forecast days.

To calculate time-based (temporal) forecast skill, the evaluation metrics were calculated over time ($n = 2101$ days for each correlation or RMSE calculation) for each grid point. Then, to get a lead time-based result for the whole domain, these metrics were averaged across all 10590 locations (grid

points), producing a 90-day CFS forecast skill assessment. Oppositely, for space-based (pattern) forecast skill, the evaluation metrics (r and RMSE) are calculated over space ($n = 10590$ locations) for each initialization date. Then, to get a lead time-based result (for the entire period), these metrics are averaged across the 2101 initialization days. Finally, forecast skill is evaluated using anomaly correlation (i.e., the anomaly of either the CP forecast or the raw CFS model values from the NARR data) on a temporal basis.

Across the study area, the spatial distribution of the skill of the CFS CP forecast using the RMSE and r was calculated by averaging CFS CP forecast in each of the 10590 locations for certain lead windows. Furthermore, the difference in the skill of the CFS CP forecast and the raw forecast, on one hand, and the CFS CP forecast and climatology, on the other hand, were calculated. These results were mapped to identify spatial variation in skill differential.

3. Results and discussion

a. Forecast skill as determined from Pearson's correlation coefficient

The forecast skill of the varying methods is shown in Fig. 3 (Pearson's correlation) and Fig. 4 (RMSE). Although 90 days of forecast skill were evaluated, the figures focus on the 0–20-day lead window since, after this period, the skill of all the forecast models is effectively constant. For correlation, the ability of all forecasts, including the CP forecasts, to replicate NARR observations expectedly decreases sharply from the first forecast day, where r is strongly positive ($r = 0.7$ – 0.9), until coming to parity with the skill of climatology (at about 9–13 days) before a gradual decline below the skill of the climatology forecast by the end of the forecast period. The high forecast skill for the CP forecast in the first 10 days and the drastic decrease in skill after is consistent with many current studies of S2S weather forecasting using more traditional techniques (e.g., Weyn et al. 2021; Li and Robertson 2015; Toth and Buizza 2019). It is also in keeping with NOAA's current outlook for most weather variables (Pegion et al. 2019). Forecast skill has continued to improve, with skill improving by one day per year over the last two decades (Johnson et al. 2014; Pegion et al. 2019).

While the raw CFS model output is more skillful than the CP forecast in the first 8–10 days, the CP forecast is better than raw model output after ~10 days of lead time, especially when examining pattern correlations (Fig. 3b). Nonetheless, the raw and CP forecasts are less skillful than climatology at longer lead times. The persistence forecast is generally the worst-performing forecast after a lead time of 1 day (lead-1), though when examining anomaly correlations (Fig. 3c), declining performance of raw CFS and CP-based forecasts come to parity with persistence after about 2 weeks of lead time.

b. Forecast skill as determined from RMSE

When evaluating RMSE (Fig. 4), many similarities with the correlation results are evident. The most significant differences, however, are noted around 9–11 lead days, when the CP-based forecast is nearly identical to or better than all of

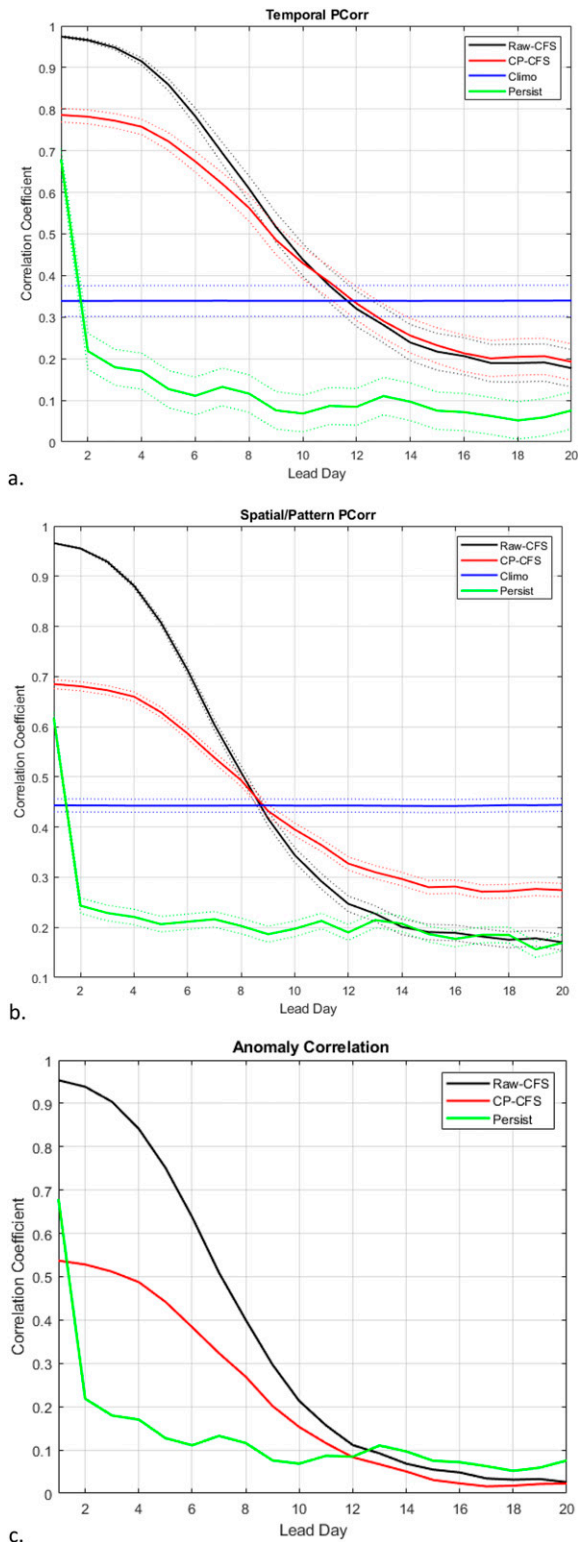


FIG. 3. Averaged Pearson's correlation (y axis) by lead day (x axis) for (a) temporal, (b) spatial/pattern, and (c) anomaly assessments of forecast skill for the different forecast models. The dotted lines show the 5th–95th confidence intervals.

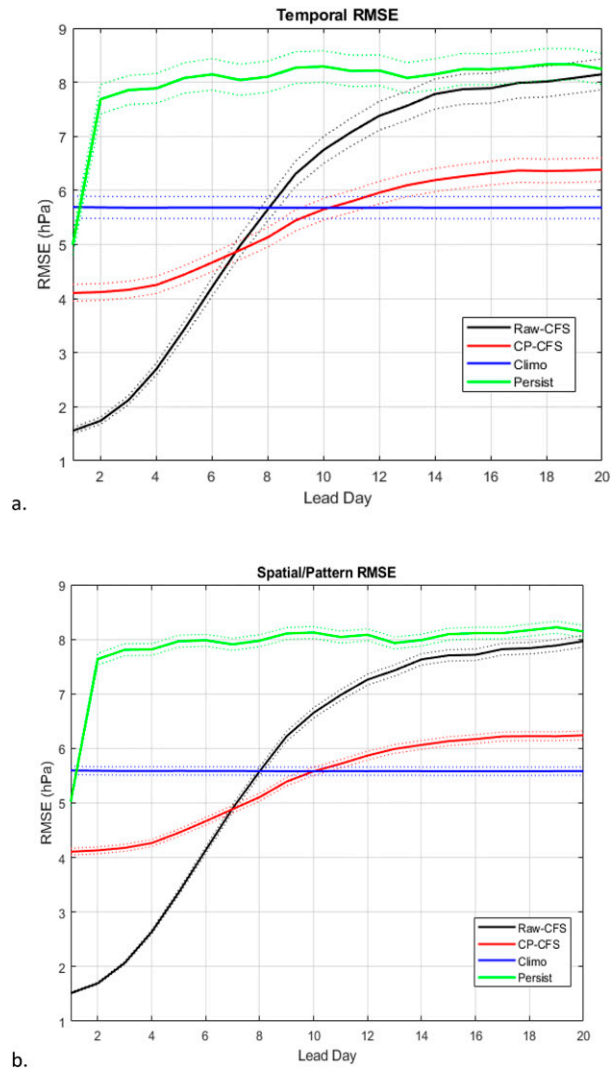
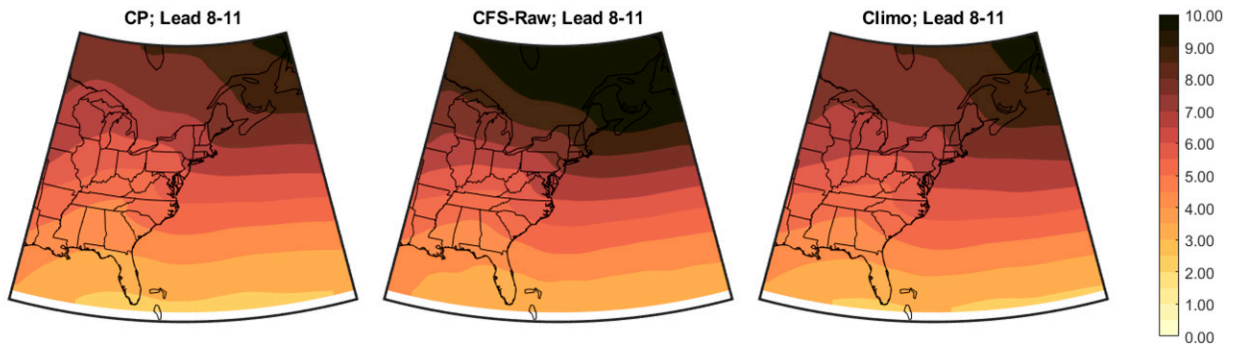


FIG. 4. Averaged RMSE (y axis; hPa) by lead day (x axis) for (a) temporal and (b) spatial/pattern assessments of forecast skill for the different forecast models. The dotted lines show the 5th–95th confidence intervals.

the other forecast types. Most notable is the rapid divergence of the CP-based forecast from that of the raw CFS forecast after about 7 days and continuing through the entire 20-day lead time (as well as the entire 90-day lead time); the CP-based forecast beats the raw CFS output consistently by about 1.5–2 hPa after week 2 and remains close to parity with the climatology forecast afterward. These results are consistent whether evaluating pattern or spatial RMSE or temporal RMSE.

On a spatial basis, at these lead times (8–11 days), forecasts generally improve inversely with latitude; that is, lower latitudes have better skill when evaluating by either correlation or RMSE (Figs. 5 and 6). That said, these spatial results help elucidate where the different forecasts' strengths lie. The CP-based forecast appears to perform best (relative to

a.



b.

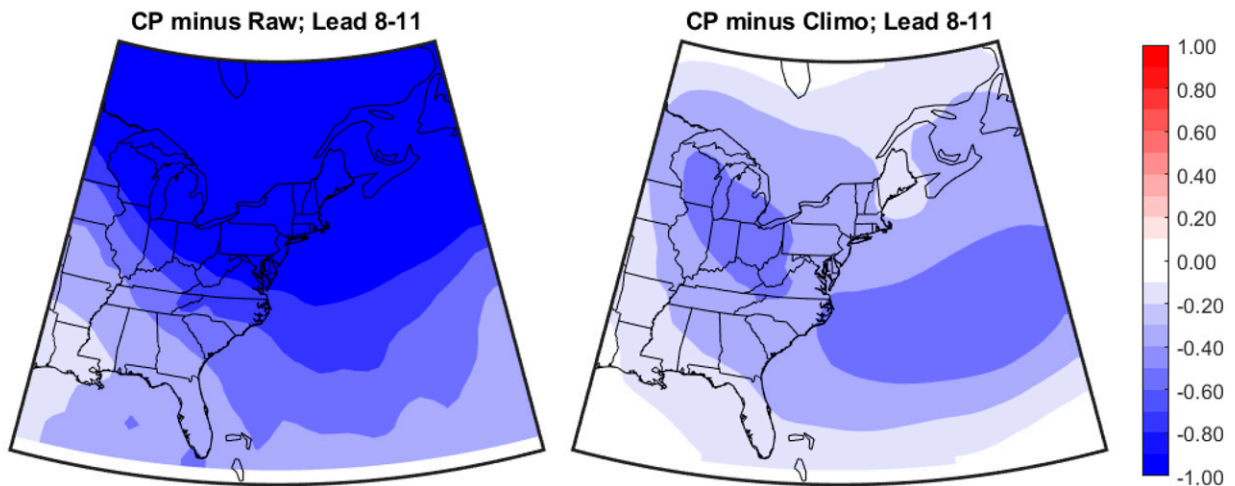


FIG. 5. (a) Averaged temporal RMSE (hPa) by location for (left) CP, (center) raw CFS, and (right) climatology forecasting methods at the 8–11-day lead time. (b) The difference in averaged temporal RMSE between (left) CP and raw CFS and between (right) CP and climatology. Lower scores (blue colors) indicate that the CP RMSE is better (i.e., lower).

climatology and raw CFS) out in the open Atlantic Ocean, where RMSEs are close to 1 hPa better than raw CFS and 0.3 hPa better than climatology (Fig. 5). With regard to correlations, again, the CP forecast appears to be more skillful largely over the open Atlantic, especially versus climatology, where it is up to 0.4 points better, with negligible differences from raw CFS at these lead times (Fig. 6). Another area of strength for the CP forecast is the Great Lakes region, where climatology forecasts struggle, especially when examining correlations.

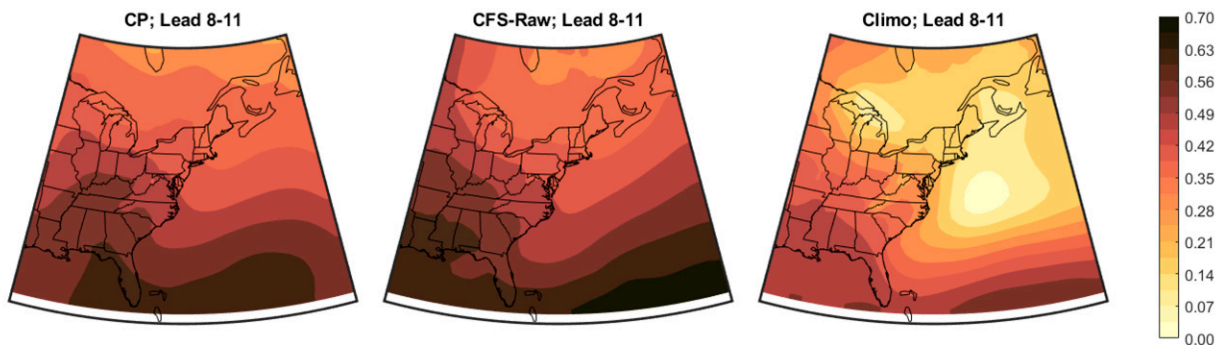
Furthermore, confidence intervals were produced for all models, based on 1000 bootstrap resamples of the RMSE and correlation coefficient values. In general, the forecast estimates for all four models in the two evaluation methods appear to be statistically significant.

Together, the results from this study suggest that CP-based predictions not only show forecast skill but, for certain lead

times, especially at the edge of current numerical model capability, this CP forecast may, indeed, be slightly better than raw model output, persistence, and climatology forecasts. In a practical sense, at least for MSLP along the East Coast of the United States, a CP-based forecast fits a useful window between where numerical model skill drops off and before the climatological signal is most useful. This window of added skill from using a CP forecast over other models is shown in Fig. 7, for temporal correlations (Fig. 7a) and temporal RMSE (Fig. 7b).

In Fig. 8, we further examined the four forecast models based on how their underlying MSLP circulation patterns evolve over different lead days (0, 3, 6, 9, 12, and 15), relative to the observed NARR MSLP. We randomly selected 15 May 2011 as a case study out of the 2101 days. The similarity between the NARR MSLP pattern and the CP and raw CFS forecast pattern depict the expectedly high forecast skill in the

a.



b.

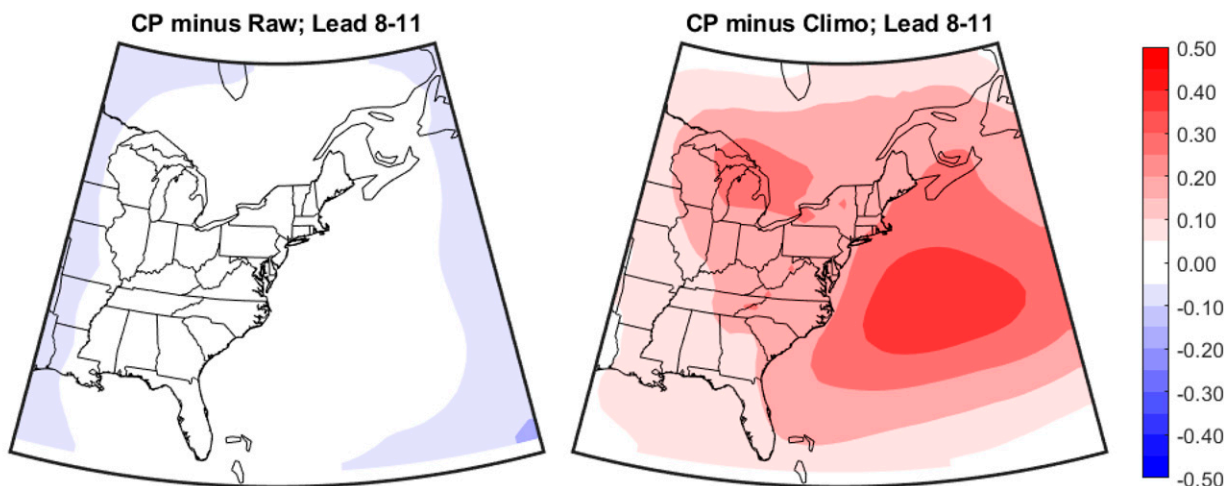


FIG. 6. As in Fig. 5, but for Pearson's correlations. In (b), higher scores (red colors) indicate that the CP correlation is better (i.e., higher).

first forecast lead day. As forecast skill decreases over the next 2 weeks, the forementioned similarity expectedly disappears. In this case study, while raw CFS becomes unskillful, the CP forecast acquires the skill of climatology, which can be seen in the similarity between the climatology forecast pattern and the CP forecast pattern.

In addition to juxtaposing the observed and forecast models, forecast anomalies were estimated by subtracting the observed MSLP fields from each of the four forecast model MSLP values at the reference lead days (Fig. 9). Anomalies generally increase with forecast lead days. More importantly, across all forecast models and lead days, positive anomalies dominate, especially over the U.S. coastal eastern seaboard and the Atlantic Ocean, suggesting that the lack of forecast skill in the MSLP S2S forecast may be due to overestimation. Positive North Atlantic MSLP forecast errors have been reported by Kolstad et al. (2020).

A medium-range forecast [defined by White et al. (2017) as a 3–10-day lead window] is critical to applied research since

day-to-day weather fluctuations impact almost every aspect of human life, as well as plant and animal ecosystems. Thus, improved forecast skill in the 8–11-day lead window has numerous sectoral applications. In the humanitarian sector, accurately forecasting extreme weather events like coastal flooding, heat waves, and hurricanes 1–2 weeks in advance can activate disaster preparedness actions that are time dependent. These actions may include prepurchase and preallocation of aid, evacuation, and other mitigation plans (White et al. 2017). Heat waves and cold spells, for example, are two of the most impactful extreme events, so more accurate early-warning systems a few days further in advance can potentially save lives. The energy sector is another that would benefit from improved medium-range weather forecasts; improved medium-range forecasts would help power suppliers prepare for anomalous cold and hot days (heating and air-conditioning demand) that could help alleviate energy price spikes and prevent damage to infrastructure (Soret et al. 2019). Moreover, as a larger percentage of the energy resource mix is composed of wind and solar, improved operational

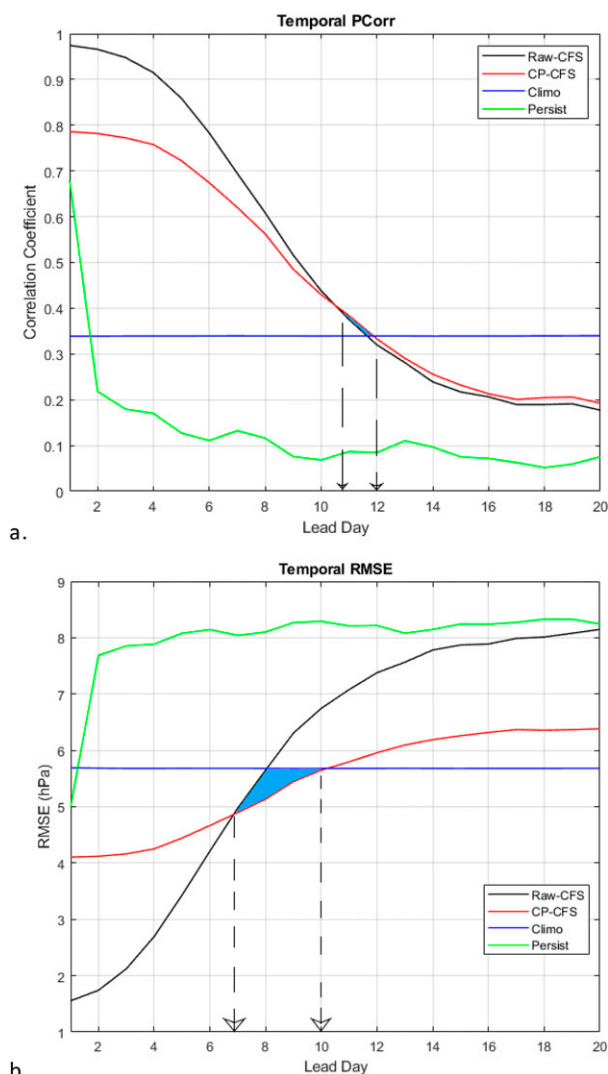


FIG. 7. Example illustration of the skill of the 20-day CFS SOM forecast over the raw CFS forecasts for (a) temporal correlations and (b) temporal RMSE. The turquoise-shaded region is the value-added skill of the SOM-based forecast over just using the output of the raw CFS forecast.

forecasting techniques will become critical (Bloomfield et al. 2021; Pinson 2013; Lynch et al. 2014). Improved forecasts beyond 1 week in advance can help operators plan for periods when wind and solar resources may fail to meet demand (Füss et al. 2015). Other areas where improved medium-range forecasts are potentially useful include instances of inland and coastal flooding, which pertains to the insurance and agricultural industries.

Perhaps as importantly, the application of categorical atmospheric patterns has a long history. That is, CPs have been utilized to relate atmospheric variability with myriad outcomes, ranging from sea level variability to human health outcomes, air quality, and tornadoes, among others. Thus, even in the absence of superior skill (compared to the skill of climatology

or raw model output) in modeling the atmosphere itself, a skillful CP-based forecast of categorical patterns may still very well be useful to decision-makers in these climate-related sectors, especially at these lead times.

Forecasting categorical variables, where only a finite set of outcomes can occur, is not completely novel, since some weather variables like precipitation type are routinely predicted in this manner. Forecasting predefined SOM-based categorical weather variables and atmospheric circulation patterns that are clustered and created from point-based, continuous weather observations, like this study has done is, however, new. The sparse focus on this approach is surprising given that most continuous data-based weather forecasts are routinely presented using categorical terms in order to provide as intuitive an understanding as possible of the weather for the need of society (National Weather Service 2022). These may, for example, be the occurrence of precipitation types, cloud cover categories (e.g., mostly sunny or mostly cloudy), or binary/trinary below-normal (near normal) and above-normal temperature, humidity, or precipitation over a given window. Besides individual weather variables, other climate system dynamics like El Niño–Southern Oscillation (ENSO), with continuous indices of sea surface temperatures (SSTs) (e.g., the Niño series) and sea level pressure [e.g., the Southern Oscillation index (SOI)] are used to discretize El Niño and La Niña years into, for example, weak, strong, very strong, and neutral categories based on standard deviation thresholds (Hanley et al. 2003; Barnston 2015). Phases of most other teleconnections are similarly broken down into categories.

Forecasting categorical variables is important because it allows for the creation of thresholds that focus on the main aspect of weather that is important for the end user (Neal et al. 2016; Roman 2017). This study sought to infuse objective thresholds into the forecasting process itself rather than as a means of summarizing the forecast results. In our case, the thresholds are homogenous, predefined, categorical atmospheric CPs that represent the whole continuum of atmospheric variability. More importantly, the smaller set of discrete categories forecasted are meaningful because they have been objectively predefined and can, therefore, be a priori related to specific local-scale applications or outcomes (e.g., extreme heat or flooding); relative to the noisiness of continuous weather forecasts at these lead times, this aspect of CPs makes them more readily usable. The value-added skill of the CFS CP forecast over the raw CFS forecast from roughly the eighth to the eleventh forecast day agrees with our working hypothesis that the predictability of S2S forecasts can increase by using CPs. It is also noteworthy that, although less than the skill of climatology, the CP forecasting outperforms the CFS model output in most of the 90-day forecast lead times.

As discussed earlier, ensemble forecasting produces many likely future scenarios, for example, when ensemble members are matched with their closest weather-pattern pair (Neal et al. 2016). While the research herein uses only a single numerical model to forecast predefined CPs (rather than a probabilistic, ensemble-based approach), we argue that our

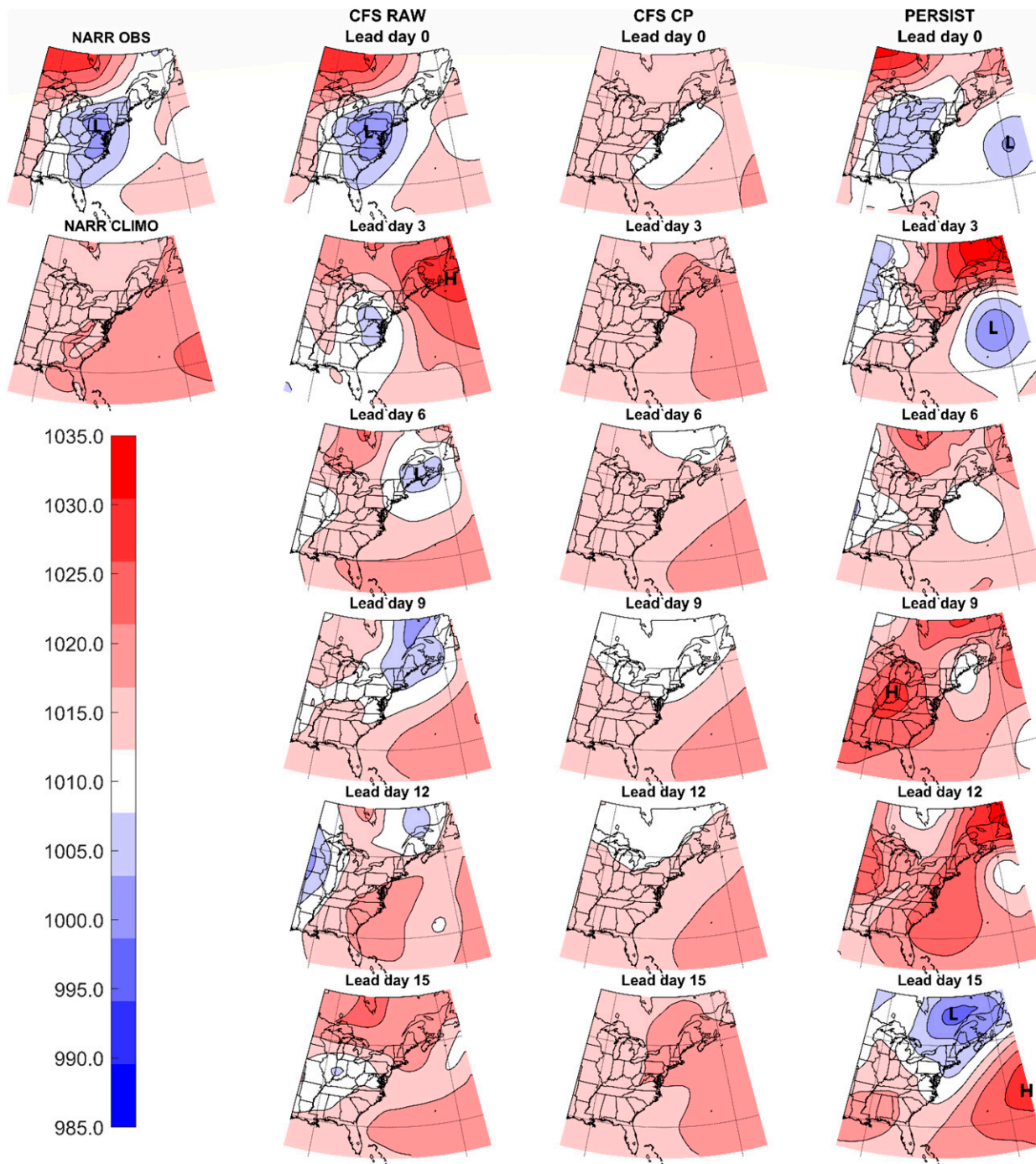


FIG. 8. Underlying MSLP pattern of the four forecast models on various lead days (raw CFS in the second column, CFS CP in the third column, and persistence in the fourth column) relative to the observed (NARR) MSLP pattern on 15 May 2011 (shown at top left), plotted as filled contours (hPa).

SOM-based method is inherently probabilistic in that, even from a single model realization, it can produce a range of possible outcomes (equal to the range of the hundreds of different patterns that are contained within the pattern that is forecasted) at a much lower computational cost than

traditional ensembles. Moreover, it is important to note that if we used an ensemble, then the SOM-based forecast stemming from that ensemble will likely be that much better as well. That is, the SOM-based method should be at least as good as the underlying data from which it is

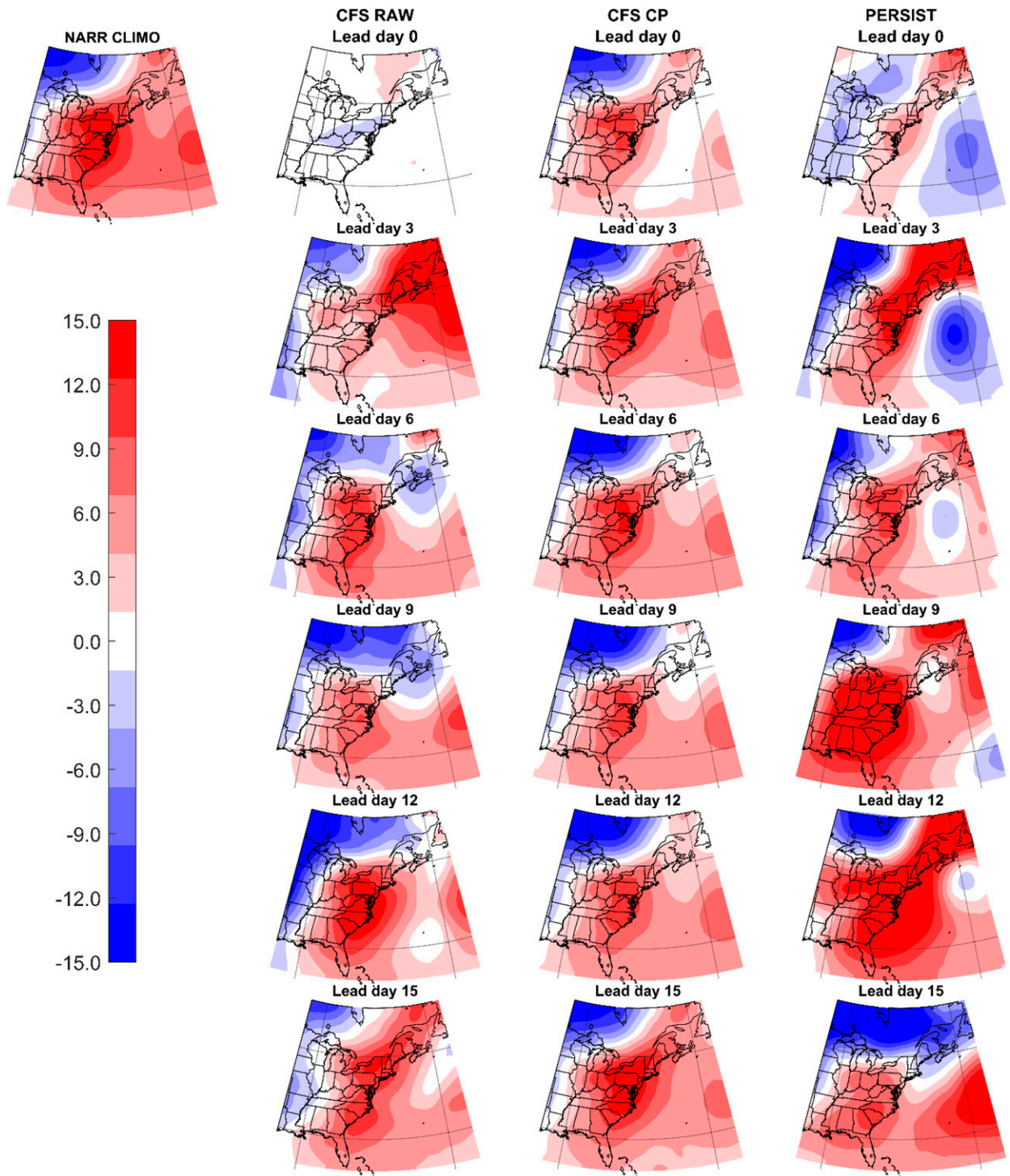


FIG. 9. Anomalies of MSLP forecasts models on 15 May 2011 for various lead days relative to observed (NARR) MSLP, plotted as filled contours (hPa) for raw CFS in the second column, CFS CP in the third column, and persistence in the fourth column.

produced. In effect, it does not matter what forecast model(s) is used as long as we compare the SOM-based forecasting method from model X to the raw forecast output from model X.

4. Summary and conclusions

Weather forecasts have been shown to be largely unskillful from week 2 onward, even with improving model ability to resolve the atmosphere and produce best estimates. This

problem is, in part, due to data initialization and parameterization schemes, high dimensionality, and low signal-to-noise ratio in climatic datasets. Numerical weather models may be our best guide to atmospheric evolution, but they are also increasingly noisy with time and hence limit forecast usefulness. We hypothesized that predicting a smaller set of clustered atmospheric states or CPs that are categorical (rather than continuous fields) may improve forecast skill. Using MSLP along the East Coast of the United States as a proof of concept, we demonstrate that using SOM-based CPs generated from CFS forecast output improves forecast skill at certain lead times and locations, especially from about 8 to 11 days. Specifically, the results showed the following:

- 1) Forecast skill generally decreased rapidly from lead-1 to the skill of climatology after 9–11 days when using correlation, and 8–11 days when using RMSE for CP forecasts and raw CFS forecasts, respectively. Thereafter, a gradual decrease in skill follows until about 3 weeks of lead time, with skill remaining roughly constant throughout the rest of the 90-day lead time.
- 2) For RMSE and spatial or pattern correlation, the skill of the CP forecasts eclipsed the skill of raw CFS output at a lead time of 8 days and stayed superior thereafter.
- 3) Of the four forecast methods compared herein, the CP-based forecasting was generally the most skillful from 8 to 11 days lead time when using RMSE. Prior to 8 days, the raw CFS is better, and after 11 days, the climatology is better. This lead-time window of value-added skill of the CP forecast supports our hypothesis regarding the utility of forecasting categorical predefined CPs rather than continuous fields.
- 4) When the skill of these forecasting methods is mapped, there is a general inverse latitudinal relationship, with the CP forecasts largely outperforming other forecasting methods over the open Atlantic Ocean and the Great Lakes.
- 5) The difference in the performance of the CFS CP forecast and the raw CFS forecast (in forecasting NARR MSLP data) is more vivid in the RMSE metric than in Pearson's r . Anomaly correlation of CFS output is slightly, but consistently, better than that of CP forecasting until week 3.

This research is part of a larger project exploring forecastability of daily-scale anomalous sea level variability around the entire United States, and MSLP in this $30^\circ \times 30^\circ$ latitudinal and longitudinal domain near the eastern coast of the United States plays an important role in this application (see Sheridan et al. 2019). As NWP models keep improving in each successive generation, it is expected that such improved skill in raw model output would, in turn, improve the CP-based forecasts as well, leading to the CP-based forecasting method becoming a multiday extension of forecast skill at whatever the contemporary skillful limit of NWP models becomes. However, forecast skill is expected to vary depending on the variable of interest and the region (e.g., see Figs. 5 and 6). Additionally, seasonality is a factor in S2S weather predictability since atmospheric variables show notable seasonal variations. Herein, producing long-term (and year-round) averages of

forecast skill may hide important seasonal differences, likely leading to some seasons being poorer than the annual averages reported here, but others being better.

Accordingly, research exploring atmospheric patterns of other meteorological variables in other domains (e.g., in the warmer Gulf of Mexico region and western North America) is currently underway. Our preliminary analysis already shows some differences in the predictability of these other weather variables, along with regional differences in forecast skill when using the SOM-based CPs. For example, wind speeds had the lowest forecast skill in all regions compared to MSLP and 700-hPa geopotential heights. Also, winter and summer forecasts were generally the most skillful and least skillful, respectively. In addition to our group's current research on CP forecasting, studies in other parts of the world (e.g., Richardson et al. 2020a; Harrison et al. 2022) may be able to provide deeper insight into the seasonal and regional differences in the predictability of predefined categorical CPs. It may also be beneficial to compare the performance of forecasted CPs created from different clustering techniques (e.g., EOFs, k -means, and others) and/or using different bias correction methodologies.

Acknowledgments. This research was supported by federal award NA17OAR4310113, entitled "Using a synoptic climatological framework to assess predictability of anomalous coastal sea levels in NOAA high priority areas" from the National Oceanic and Atmospheric Administration's (NOAA) Climate Program Office (CPO).

Data availability statement. The NARR data analyzed during this study are available on the National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR) website homepage, <https://www.emc.ncep.noaa.gov/mmb/rrean/index.html> (<https://doi.org/10.1175/BAMS-87-3-343>). The Climate Forecast System (CFS) reanalysis (CFRS), and the CFS version 2 (CFSv2) 9-month operational forecasts are available on the National Center for Atmospheric Research (NCAR) National Centers for Environmental Prediction (NCEP) website homepage, <https://rda.ucar.edu/datasets/ds094.0/> (<https://doi.org/10.1175/JCLI-D-12-00823.1>).

REFERENCES

- Adams, R. E., C. C. Lee, E. T. Smith, and S. C. Sheridan, 2021: The relationship between atmospheric circulation patterns and extreme temperature events in North America. *Int. J. Climatol.*, **41**, 92–103, <https://doi.org/10.1002/joc.6610>.
- Amidror, I., 2002: Scattered data interpolation methods for electronic imaging systems: A survey. *J. Electron. Imaging*, **11**, 157–176, <https://doi.org/10.1117/1.1455013>.
- Bagagnoli, L., R. Inghoso, and M. M. Miglietta, 2021: Synoptic patterns and mesoscale precursors of Italian tornadoes. *Atmos. Res.*, **253**, 105503, <https://doi.org/10.1016/j.atmosres.2021.105503>.
- Barnston, A., 2015: Why are there so many ENSO indexes, instead of just one? ENSO blog, accessed 21 July 2021, <https://www.climate.gov/news-features/blogs/enso/why-are-there-so-many-enso-indexes-instead-just-one>.

- Berkovic, S., O. Y. Mendelsohn, E. Ilotoviz, and S. Raveh-Rubin, 2021: Self-organizing map classification of the boundary layer profile: A refinement of eastern Mediterranean winter synoptic regimes. *Int. J. Climatol.*, **41**, 3317–3338, <https://doi.org/10.1002/joc.7021>.
- Bloomfield, H. C., D. J. Brayshaw, P. L. M. Gonzalez, and A. Charlton-Perez, 2021: Sub-seasonal forecasts of demand and wind power and solar power generation for 28 European countries. *Earth Syst. Sci. Data*, **13**, 2259–2274, <https://doi.org/10.5194/essd-13-2259-2021>.
- Brunet, G., and Coauthors, 2010: Collaboration of the weather and climate communities to advance subseasonal-to-seasonal prediction. *Bull. Amer. Meteor. Soc.*, **91**, 1397–1406, <https://doi.org/10.1175/2010BAMS3013.1>.
- Craig, G. C., and Coauthors, 2021: Waves to weather: Exploring the limits of predictability of weather. *Bull. Amer. Meteor. Soc.*, **102**, E2151–E2164, <https://doi.org/10.1175/BAMS-D-20-0035.1>.
- Davies, D. L., and D. W. Bouldin, 1979: A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.*, **1**, 224–227, <https://doi.org/10.1109/TPAMI.1979.4766909>.
- Dirmeyer, P. A., 2013: Characteristics of the water cycle and land–atmosphere interactions from a comprehensive reforecast and reanalysis data set: CFSv2. *Climate Dyn.*, **41**, 1083–1097, <https://doi.org/10.1007/s00382-013-1866-x>.
- Doblas-Reyes, F. J., and Coauthors, 2013: Initialized near-term regional climate change prediction. *Nat. Commun.*, **4**, 1715, <https://doi.org/10.1038/ncomms2704>.
- Dormann, C. F., and Coauthors, 2013: Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 27–46, <https://doi.org/10.1111/j.1600-0587.2012.07348.x>.
- Ferranti, L., S. Corti, and M. Janousek, 2015: Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quart. J. Roy. Meteor. Soc.*, **141**, 916–924, <https://doi.org/10.1002/qj.2411>.
- Füss, R., S. Mahringer, and M. Prokopczuk, 2015: Electricity derivatives pricing with forward-looking information. *J. Econ. Dyn. Control*, **58**, 34–57, <https://doi.org/10.1016/j.jedc.2015.05.016>.
- García-Herrera, R., J. M. Garrido-Perez, and C. Ordóñez, 2022: Modulation of European air quality by Euro-Atlantic weather regimes. *Atmos. Res.*, **277**, 106292, <https://doi.org/10.1016/j.atmosres.2022.106292>.
- Gilabert, J., and M. C. Llasat, 2018: Circulation weather types associated with extreme flood events in northwestern Mediterranean. *Int. J. Climatol.*, **38**, 1864–1876, <https://doi.org/10.1002/joc.5301>.
- Hamill, T. M., J. S. Whitaker, and S. L. Mullen, 2006: Reforecasts: An important dataset for improving weather predictions. *Bull. Amer. Meteor. Soc.*, **87**, 33–46, <https://doi.org/10.1175/BAMS-87-1-33>.
- Hanley, D. E., M. A. Bourassa, J. J. O'Brien, S. R. Smith, and E. R. Spade, 2003: A quantitative evaluation of ENSO indices. *J. Climate*, **16**, 1249–1258, [https://doi.org/10.1175/1520-0442\(2003\)16<1249:AQEOEI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)16<1249:AQEOEI>2.0.CO;2).
- Harrison, S. R., J. O. Pope, R. A. Neal, F. K. Garry, R. Kurashina, and D. Suri, 2022: Identifying weather patterns associated with increased volcanic ash risk within British Isles airspace. *Wea. Forecasting*, **37**, 1157–1168, <https://doi.org/10.1175/WAF-D-22-0023.1>.
- He, S., X. Li, T. DelSole, P. Ravikumar, and A. Banerjee, 2021: Sub-seasonal climate forecasting via machine learning: Challenges, analysis, and advances. *Proc. AAAI Conf. on Artificial Intelligence*, Online, AAAI, 169–177, <https://doi.org/10.1609/aaai.v35i1.16090>.
- Hewitson, B. C., and R. G. Crane, 2002: Self-organizing maps: Applications to synoptic climatology. *Climate Res.*, **22**, 13–26, <https://doi.org/10.3354/cr022013>.
- Huang, W. T. K., A. Charlton-Perez, R. W. Lee, R. Neal, C. Sarran, and T. Sun, 2020: Weather regimes and patterns associated with temperature-related excess mortality in the UK: A pathway to sub-seasonal risk forecasting. *Environ. Res. Lett.*, **15**, 124052, <https://doi.org/10.1088/1748-9326/abcba>.
- Hwang, J., P. Orenstein, J. Cohen, K. Pfeiffer, and L. Mackey, 2019: Improving subseasonal forecasting in the western US with machine learning. *KDD'19: Proc. 25th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, Anchorage, AK, Association for Computing Machinery, 2325–2335, <https://dl.acm.org/doi/proceedings/10.1145/3292500>.
- Jacox, M. G., M. A. Alexander, C. A. Stock, and G. Hervieux, 2019: On the skill of seasonal sea surface temperature forecasts in the California Current System and its connection to ENSO variability. *Climate Dyn.*, **53**, 7519–7533, <https://doi.org/10.1007/s00382-017-3608-y>.
- Jaye, A. B., C. L. Bruyère, and J. M. Done, 2019: Understanding future changes in tropical cyclogenesis using self-organizing maps. *Wea. Climate Extremes*, **26**, 100235, <https://doi.org/10.1016/j.wace.2019.100235>.
- 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>.
- Kohonen, T., 1990: The self-organizing map. *Proc. IEEE*, **78**, 1464–1480, <https://doi.org/10.1109/5.58325>.
- Kolstad, E. W., C. O. Wulff, D. I. Domeisen, and T. Woollings, 2020: Tracing North Atlantic Oscillation forecast errors to stratospheric origins. *J. Climate*, **33**, 9145–9157, <https://doi.org/10.1175/JCLI-D-20-0270.1>.
- Lee, C. C., 2012: Utilizing synoptic climatological methods to assess the impacts of climate change on future tornado-favorable environments. *Nat. Hazards*, **62**, 325–343, <https://doi.org/10.1007/s11069-011-9998-y>.
- , 2017: Reanalysing the impacts of atmospheric teleconnections on cold-season weather using multivariate surface weather types and self-organizing maps. *Int. J. Climatol.*, **37**, 3714–3730, <https://doi.org/10.1002/joc.4950>.
- , O. Obarein, S. C. Sheridan, E. T. Smith, and R. Adams, 2021: Examining trends in multiple parameters of seasonally-relative extreme temperature and dew point events across North America. *Int. J. Climatol.*, **41**, E2360–E2378, <https://doi.org/10.1002/joc.6852>.
- Li, S., and A. W. Robertson, 2015: Evaluation of submonthly precipitation forecast skill from global ensemble prediction systems. *Mon. Wea. Rev.*, **143**, 2871–2889, <https://doi.org/10.1175/MWR-D-14-00277.1>.
- Lipperheide, M., J. L. Bosch, and J. Kleissl, 2015: Embedded nowcasting method using cloud speed persistence for a photovoltaic power plant. *Sol. Energy*, **112**, 232–238, <https://doi.org/10.1016/j.solener.2014.11.013>.
- Liu, Y., and R. H. Weisberg, 2011: A review of self-organizing map applications in meteorology and oceanography. *Self-Organizing Maps: Applications and Novel Algorithm Design*, J. I. Mwasiagi, Ed., IntechOpen, 253–272.

- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, [https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2).
- Lynch, K. J., D. J. Brayshaw, and A. Charlton-Perez, 2014: Verification of European subseasonal wind speed forecasts. *Mon. Wea. Rev.*, **142**, 2978–2990, <https://doi.org/10.1175/MWR-D-13-00341.1>.
- Mastrantonas, N., P. Herrera-Lormendez, L. Magnusson, F. Pappenberger, and J. Matschullat, 2021: Extreme precipitation events in the Mediterranean: Spatiotemporal characteristics and connection to large-scale atmospheric flow patterns. *Int. J. Climatol.*, **41**, 2710–2728, <https://doi.org/10.1002/joc.6985>.
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, <https://doi.org/10.1175/BAMS-87-3-343>.
- National Weather Service, 2022: Forecast terms. National Weather Service, accessed 21 July 2022, https://www.weather.gov/bgm/forecast_terms.
- Neal, R., D. Fereday, R. Crocker, and R. E. Comer, 2016: A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteor. Appl.*, **23**, 389–400, <https://doi.org/10.1002/met.1563>.
- , R. Dankers, A. Saulter, A. Lane, J. Millard, G. Robbins, and D. Price, 2018: Use of probabilistic medium- to long-range weather-pattern forecasts for identifying periods with an increased likelihood of coastal flooding around the UK. *Meteor. Appl.*, **25**, 534–547, <https://doi.org/10.1002/met.1719>.
- , J. Robbins, R. Dankers, A. Mitra, A. Jayakumar, E. N. Rajagopal, and G. Adamson, 2020: Deriving optimal weather pattern definitions for the representation of precipitation variability over India. *Int. J. Climatol.*, **40**, 342–360, <https://doi.org/10.1002/joc.6215>.
- 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>.
- Pinson, P., 2013: Wind energy: Forecasting challenges for its operational management. *Stat. Sci.*, **28**, 564–585, <https://doi.org/10.1214/13-STS445>.
- Pirhalla, D. E., C. C. Lee, S. C. Sheridan, and V. Ransibrahmanakul, 2022: Atlantic coastal sea level variability and synoptic-scale meteorological forcing. *J. Appl. Meteor. Climatol.*, **61**, 205–222, <https://doi.org/10.1175/JAMC-D-21-0046.1>.
- Prodhomme, C., F. Doblas-Reyes, O. Bellprat, and E. Dutra, 2016: Impact of land-surface initialization on sub-seasonal to seasonal forecasts over Europe. *Climate Dyn.*, **47**, 919–935, <https://doi.org/10.1007/s00382-015-2879-4>.
- Richardson, D., R. Neal, R. Dankers, K. Mylne, R. Cawling, H. Clements, and J. Millard, 2020a: Linking weather patterns to regional extreme precipitation for highlighting potential flood events in medium- to long-range forecasts. *Meteor. Appl.*, **27**, e1931, <https://doi.org/10.1002/met.1931>.
- , H. J. Fowler, C. G. Kilsby, R. Neal, and R. Dankers, 2020b: Improving sub-seasonal forecast skill of meteorological drought: A weather pattern approach. *Nat. Hazards Earth Syst. Sci.*, **20**, 107–124, <https://doi.org/10.5194/nhess-20-107-2020>.
- Riddle, E. E., A. H. Butler, J. C. Furtado, J. L. Cohen, and A. Kumar, 2013: CFSv2 ensemble prediction of the wintertime Arctic Oscillation. *Climate Dyn.*, **41**, 1099–1116, <https://doi.org/10.1007/s00382-013-1850-5>.
- Robertson, A., and F. Vitart, Eds., 2018: *Sub-Seasonal to Seasonal Prediction: The Gap between Weather and Climate Forecasting*. Elsevier, 585 pp.
- , and —, Eds., 2019: *Sub-Seasonal to Seasonal Prediction: What Sets the Forecast Skill Horizon?* Elsevier, 569 pp.
- , A. Kumar, M. Peña, and F. Vitart, 2015: Improving and promoting subseasonal to seasonal prediction. *Bull. Amer. Meteor. Soc.*, **96**, ES49–ES53, <https://doi.org/10.1175/BAMS-D-14-00139.1>.
- Roman, J., 2017: METEO 825: Predictive analytic techniques for meteorological data. Department of Meteorology and Atmospheric Science, Pennsylvania State University, <https://www.e-education.psu.edu/meteo825/node/508>.
- Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, <https://doi.org/10.1175/JCLI-D-12-00823.1>.
- Sheridan, S. C., and C. C. Lee, 2011: The self-organizing map in synoptic climatological research. *Prog. Phys. Geogr. Earth Environ.*, **35**, 109–119, <https://doi.org/10.1177/0309133310397582>.
- , D. E. Pirhalla, C. C. Lee, and V. Ransibrahmanakul, 2013: Evaluating linkages of weather patterns and water quality responses in south Florida using a synoptic climatological approach. *J. Appl. Meteor. Climatol.*, **52**, 425–438, <https://doi.org/10.1175/JAMC-D-12-0126.1>.
- , C. C. Lee, D. E. Pirhalla, and V. Ransibrahmanakul, 2017: Atmospheric drivers of sea-level fluctuations and nuisance floods along the mid-Atlantic coast of the USA. *Reg. Environ. Change*, **17**, 1853–1861, <https://doi.org/10.1007/s10113-017-1156-y>.
- , —, R. E. Adams, E. T. Smith, D. E. Pirhalla, and V. Ransibrahmanakul, 2019: Temporal modeling of anomalous coastal sea-level values using synoptic climatological patterns. *J. Geophys. Res. Oceans*, **124**, 6531–6544, <https://doi.org/10.1029/2019JC015421>.
- Shin, C.-S., P. A. Dirmeyer, B. Huang, S. Halder, and A. Kumar, 2020: Impact of land initial states uncertainty on subseasonal surface air temperature prediction in CFSv2 reforecasts. *J. Hydrometeorol.*, **21**, 2101–2121, <https://doi.org/10.1175/JHM-D-20-0024.1>.
- Shutts, G., and A. C. Pallarès, 2014: Assessing parametrization uncertainty associated with horizontal resolution in numerical weather prediction models. *Philos. Trans. Roy. Soc.*, **A372**, 20130284, <https://doi.org/10.1098/rsta.2013.0284>.
- Smith, E. T., O. Obarein, S. C. Sheridan, and C. C. Lee, 2021: Assessing trends in atmospheric circulation patterns across North America. *Int. J. Climatol.*, **41**, 2679–2692, <https://doi.org/10.1002/joc.6983>.
- Solidoro, C., V. Bandelj, P. Barbieri, G. Cossarini, and S. Fonda Umami, 2007: Understanding dynamic of biogeochemical properties in the northern Adriatic Sea by using self-organizing maps and *k*-means clustering. *J. Geophys. Res.*, **112**, C07S90, <https://doi.org/10.1029/2006JC003553>.
- Soret, A., and Coauthors, 2019: Sub-seasonal to seasonal climate predictions for wind energy forecasting. *J. Phys.*, **1222**, 012009, <https://doi.org/10.1088/1742-6596/1222/1/012009>.
- Sousa, S. I. V., F. G. Martins, M. C. M. Alvim-Ferraz, and M. C. Pereira, 2007: Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environ. Modell. Software*, **22**, 97–103, <https://doi.org/10.1016/j.envsoft.2005.12.002>.
- Tochimoto, E., 2022: Environmental controls on tornadoes and tornado outbreaks. *Atmos.–Ocean*, **60**, 399–421, <https://doi.org/10.1080/07055900.2022.2079472>.
- Toth, Z., and R. Buizza, 2019: Weather forecasting: What sets the forecast skill horizon? *Sub-Seasonal to Seasonal Prediction*, A. W. Robertson and F. Vitart, Eds., Elsevier, 17–45.
- Vannitsem, S., D. Wilks, and J. Messner, Eds., 2018: *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, 362 pp.

- Vitart, F., 2014: Evolution of ECMWF sub-seasonal forecast skill scores. *Quart. J. Roy. Meteor. Soc.*, **140**, 1889–1899, <https://doi.org/10.1002/qj.2256>.
- , A. W. Robertson, and S2S Steering Group, 2015: Sub-seasonal to seasonal prediction: Linking weather and climate. *Seamless Prediction of the Earth System: From Minutes to Months*, World Meteorological Organization, 385–401.
- , and Coauthors, 2017: The Subseasonal to Seasonal (S2S) Prediction Project database. *Bull. Amer. Meteor. Soc.*, **98**, 163–173, <https://doi.org/10.1175/BAMS-D-16-0017.1>.
- Walsh, J. E., 1984: Forecasts of monthly 700 mb height: Verification and specification experiments. *Mon. Wea. Rev.*, **112**, 2135–2147, [https://doi.org/10.1175/1520-0493\(1984\)112<2135:FOMMHV>2.0.CO;2](https://doi.org/10.1175/1520-0493(1984)112<2135:FOMMHV>2.0.CO;2).
- Wedam, G. B., L. A. McMurdie, and C. F. Mass, 2009: Comparison of model forecast skill of sea level pressure along the East and West Coasts of the United States. *Wea. Forecasting*, **24**, 843–854, <https://doi.org/10.1175/2008WAF2222161.1>.
- Weyn, J. A., D. R. Durran, R. Caruana, and N. Cresswell-Clay, 2021: Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models. *J. Adv. Model. Earth Syst.*, **13**, e2021MS002502, <https://doi.org/10.1029/2021MS002502>.
- 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>.
- Wilkinson, J. M., and R. Neal, 2021: Exploring relationships between weather patterns and observed lightning activity for Britain and Ireland. *Quart. J. Roy. Meteor. Soc.*, **147**, 2772–2795, <https://doi.org/10.1002/qj.4099>.
- Wilks, D. S., 2020: *Statistical Methods in the Atmospheric Sciences*. Vol. 100, Academic Press, 704 pp.
- , and S. Vannitsem, 2018: Uncertain forecasts from deterministic dynamics. *Statistical Postprocessing of Ensemble Forecasts*, S. Vannitsem, D. Wilks, and J. Messner, Eds., Elsevier, 1–13.
- Yarnal, B., 1993: *Synoptic Climatology in Environmental Analysis: A Primer*. Belhaven Press, 195 pp.
- , A. C. Comrie, B. Frakes, and D. P. Brown, 2001: Developments and prospects in synoptic climatology. *Int. J. Climatol.*, **21**, 1923–1950, <https://doi.org/10.1002/joc.675>.
- Zhou, C., G. Wei, J. Xiang, K. Zhang, C. Li, and J. Zhang, 2018: Effects of synoptic circulation patterns on air quality in Nanjing and its surrounding areas during 2013–2015. *Atmos. Pollut. Res.*, **9**, 723–734, <https://doi.org/10.1016/j.apr.2018.01.015>.
- Zuo, Z., S. Yang, Z.-Z. Hu, R. Zhang, W. Wang, B. Huang, and F. Wang, 2013: Predictable patterns and predictive skills of monsoon precipitation in Northern Hemisphere summer in NCEP CFSv2 reforecasts. *Climate Dyn.*, **40**, 3071–3088, <https://doi.org/10.1007/s00382-013-1772-2>.