Dynamical–Statistical Prediction of Week-2 Severe Weather for the United States

HUI WANG,[a] ARUN KUMAR,[a] ALIMA DIAWARA,[a,b] DAVID DEWITT,[a] AND JON GOTTSCHALCK[a]

[a] *NOAA/NWS/NCEP/Climate Prediction Center, College Park, Maryland*
[b] *Innovim LLC, Greenbelt, Maryland*

(Manuscript received 17 January 2020, in final form 16 November 2020)

ABSTRACT: A dynamical–statistical model is developed for forecasting week-2 severe weather (hail, tornadoes, and damaging winds) over the United States. The supercell composite parameter (SCP) is used as a predictor, which is derived from the 16-day dynamical forecasts of the National Centers for Environmental Prediction (NCEP) Global Ensemble Forecast System (GEFS) model and represents the large-scale convective environments influencing severe weather. The hybrid model forecast is based on the empirical relationship between GEFS hindcast SCP and observed weekly severe weather frequency during 1996–2012, the GEFS hindcast period. Cross validations suggest that the hybrid model has a low skill for week-2 severe weather when applying simple linear regression method at 0.5° × 0.5° (latitude × longitude) grid data. However, the forecast can be improved by using the 5° × 5° area-averaged data. The forecast skill can be further improved by using the empirical relationship depicted by the singular value decomposition method, which takes into account the spatial covariations of weekly severe weather. The hybrid model was tested operationally in spring 2019 and demonstrated skillful forecasts of week-2 severe weather frequency over the United States.

KEYWORDS: Hail; Tornadoes; Forecast verification/skill; Statistical forecasting; Ensembles

1. Introduction

Large hail (size ≥ 1 in.), tornadoes, and high winds are severe convective storm events, which cause significant property damages with billions of dollars in losses and personal injuries or deaths every year in the United States (NOAA 2019). Some of these events are characterized by small spatial scales and short lifetime (e.g., severe thunderstorms of 15–25 km in diameter and a lifetime of 20–30 min), and thus, pose challenges in forecasting severe weather. Skillful severe weather outlooks and warnings are of critical importance to the National Oceanic and Atmospheric Administration (NOAA)/National Weather Service's (NWS) mission to protect lives and property.

In general, severe weather forecasts can be divided into three different time scales, namely, operational weather forecasts, subseasonal forecasts, and seasonal forecasts. The operational weather forecasts for 1–7 days with a numerical model are mainly determined by atmospheric initial conditions. Currently, the NOAA's Storm Prediction Center (SPC) produces convective outlooks for days 1–8. The NOAA's Climate Prediction Center (CPC) also includes severe weather in its U.S. day 3–7 hazards outlook. In contrast, the seasonal forecasts are from one month to several seasons (e.g., Tippett et al. 2012; Lepore et al. 2017, 2018). At this time scale, slow evolving components, such as sea surface temperature (SST), sea ice, soil moisture, and low-frequency modes, provide the sources of predictability (e.g., Shepherd et al. 2009; Allen et al. 2015; Lee et al. 2016; Trapp and Hoogewind 2018). Between weather and seasonal forecasts reside the subseasonal forecasts from week 2 to week 4. At this time scale, the sources of predictability for severe weather are limited, which include the tropical Madden–Julian oscillation (MJO; Barrett and Gensini 2013; Thompson and Roundy 2013; Baggett et al. 2018) and the global wind oscillation (GWO, Gensini and A. Marinaro 2016). In addition, while the influence of initial condition diminishes, the time-average is small enough that the signal associated with the slow evolving components is not appreciable compared to noise.

Developing week-2 to week-4 severe weather outlooks is one of the NOAA's Climate Prediction Center projects under the Office of Science and Technology Policy (OSTP) initiative (NWS 2016). The goals of this project are 1) to develop and perform evaluation of week-2 severe weather model guidance, and 2) to explore the potential for developing experimental forecast tools for week-3 and week-4 severe weather. As the first step toward meeting these goals, the present work focuses on developing a week-2 severe weather outlook for the United States.

Given their small spatial scales and short lifetime, severe convective storms cannot be sufficiently resolved by current operational global forecasting models (Weisman et al. 1997; Gensini and Mote 2014). However, these global models can predict the large-scale environments that may affect severe weather, even beyond one week. Previous studies (Thompson et al. 2003, 2007) introduced a variable, the so-called supercell composite parameter (SCP), to characterize large-scale convective environments that favor right-moving supercells in the Great Plains. It should be noted that the use of the SCP implies severe weather is produced by supercell thunderstorms, primarily for hail and tornadoes, and is not necessarily related to quasi-linear convective systems (QLCSs). The latter is a significant source of damaging wind (e.g., Ashley et al. 2019). Carbin et al. (2016) use the SCP predicted by a dynamical model as environment guidance for extended-range forecasts of severe thunderstorms. More recently, Gensini and Tippett (2019) also use model predicted SCP to explicitly forecast day

Corresponding author: Hui Wang, hui.wang@noaa.gov

DOI: 10.1175/WAF-D-20-0009.1

1–15 tornado and hail frequencies over the United States. They choose different SCP thresholds for tornado (SCP ≥ 4) and hail (SCP ≥ 2). Skillful forecasts are found to day 9 for tornado and day 12 for hail, respectively. The forecast tool developed in this study will provide a general guide to the expected overall severe weather activity during week 2. It does not predict individual severe weather events for any particular day and time.

The present study complements Carbin et al. (2016) and Gensini and Tippett (2019) by taking another step to explicitly forecast week-2 severe weather based on dynamical model predicted week-2 SCP and the empirical relationship between model predicted SCP and observed weekly severe weather over a historical record, a dynamical–statistical approach (e.g., Wang et al. 2009; Harnos et al. 2019). Based on our experience with this methodology, the forecast skill of such a hybrid model will largely be determined by the strength of the statistical relationship between the SCP and severe weather. We will show that weekly severe weather averaged over a relatively large domain (e.g., 5° × 5° latitude × longitude box) is better related to the large-scale environments (SCP) than that over a small area (0.5° × 0.5° box). The same approach of regional average has been used to forecast subseasonal severe weather (e.g., Lepore et al. 2018), as well as lightning (Tippett and Koshak 2018). The forecast skill can also be improved by considering the spatial covariations of weekly severe weather with its surroundings.

The present study is aimed at developing a forecast tool for week-2 severe weather over the United States. The primary foci are (i) to characterize the seasonality and spatial coherence of severe weather in the United States, (ii) to establish empirical relationships between large-scale environments (SCP) and weekly severe weather, (iii) to develop and cross validate the hybrid dynamical–statistical forecast model, and (iv) to test the model in real-time forecasting during spring 2019.

2. Data and methodology

The data used in this study consist of both observational and reanalysis data, as well as model forecast/hindcast data. For observations, the National Weather Service (NWS) Local Storm Reports (LSRs, available at https://www.spc.noaa.gov/wcm/#data) are used. The National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR, Saha et al. 2010) are also employed. The LSRs include three types of severe weather, namely, hail, tornado, and damaging wind, as well as their location, date/time, and intensity. In this paper, we focus on developing the forecast tool for a total weekly (or 7-day) number of severe weather events, referred to as LSR3 hereafter, without distinguishing them for any specific types of severe weather. Daily LSR data are regridded to a 0.5° × 0.5° (latitude × longitude) grid by counting the number of LSRs within a 24-h period for each 0.5° × 0.5° box centered at each grid point. Weekly (or 7-day) LSR values are the sum of corresponding 7-day LSRs (e.g., from Monday to Sunday, from Tuesday to Monday).

Following Carbin et al. (2016), the SCP is expressed as

$$\mathrm{SCP} = (\mathrm{CAPE}/1000\,\mathrm{J\ kg^{-1}}) \times (\mathrm{SRH}/50\,\mathrm{m^{-2}s^{-2}}) \times (\mathrm{BWD}/20\,\mathrm{m\ s^{-1}}),$$

and

$$\mathrm{BWD} = [(u_1 - u_2)^2 + (v_1 - v_2)^2]^{0.5},$$

where CAPE is convective available potential energy, SRH is storm-relative helicity, and BWD is bulk wind difference between 500 hPa (u_1, v_1) and 10 m (u_2, v_2). Three constants are used to normalize SCP so that when SCP is greater than 1, severe weather is likely to occur. To derive SCP from both reanalysis data (CFSR; Saha et al. 2010) and model forecasts, 6-hourly lower-level CAPE in the layer of 180–0 hPa above the ground and SRH in the layer of 3000–0 m above the ground are directly taken from the reanalysis data and model forecast output. It should be noted that when the deep-layer wind shear BWD is derived using winds at 500 hPa and 10 m, it could be significantly underestimated on the high elevations where 500 hPa is closer to the model surface, as discussed by Gensini and Ashley (2011). A recent study by King and Kennedy (2019) find that CAPE, SRH, and BWD in CFSR have negative biases, as compared to sounding data, leading to a negative bias in SCP. Similar to Carbin et al. (2016), daily SCP is the average of 6-hourly SCP from 1200 to 1200 UTC (average of five instantaneous values at 1200, 1800, 0000, 0600, and 1200 UTC).

Given that the current operational global dynamical models cannot resolve the characteristic spatial and temporal scales of severe weather, the forecast tool for LSR3 developed in this study is a hybrid dynamical–statistical model (e.g., Wang et al. 2009; Mehta et al. 2014; Harnos et al. 2019). Briefly, it uses the dynamical model predicted SCP as a predictor, and then forecasts LSR3 based on a statistical relationship between *model predicted* SCP and *observed* LSR3 in the historical record. Although dynamical models can provide some useful guidance for depicting large-scale patterns that are associated with strongly forced synoptically evident events, they clearly have no skill with weakly forced or mesoscale-type severe weather events, especially at the week-2 time range. The merit of bringing in the statistical component is to help illuminate where the dynamical model guidance on the large scale is hinting at synoptic-scale severe weather potential. The dynamical model employed is the NCEP Global Ensemble Forecast System (GEFS; Wei et al. 2008), an atmospheric model with observed SST relaxed to climatology as a low boundary. It should be noted that using climatological SST as boundary forcing may reduce forecast quality, especially since moist return flow from the Gulf of Mexico is a critical parameter in severe weather events (e.g., Jung and Kirtman 2016). Therefore, improving skill of SST forecasts could provide an opportunity to further improve skill. The GEFS makes 16-day forecasts with horizontal resolutions of T254 (~55 km) for days 1–8 and T190 (~70 km) for days 9–16, and 42 vertical levels (Zhou et al. 2017).

A 17-yr GEFS hindcast dataset (Guan et al. 2015) is used to establish the statistical relationship between model SCP and observed LSR3. The GEFS hindcasts are five members initialized

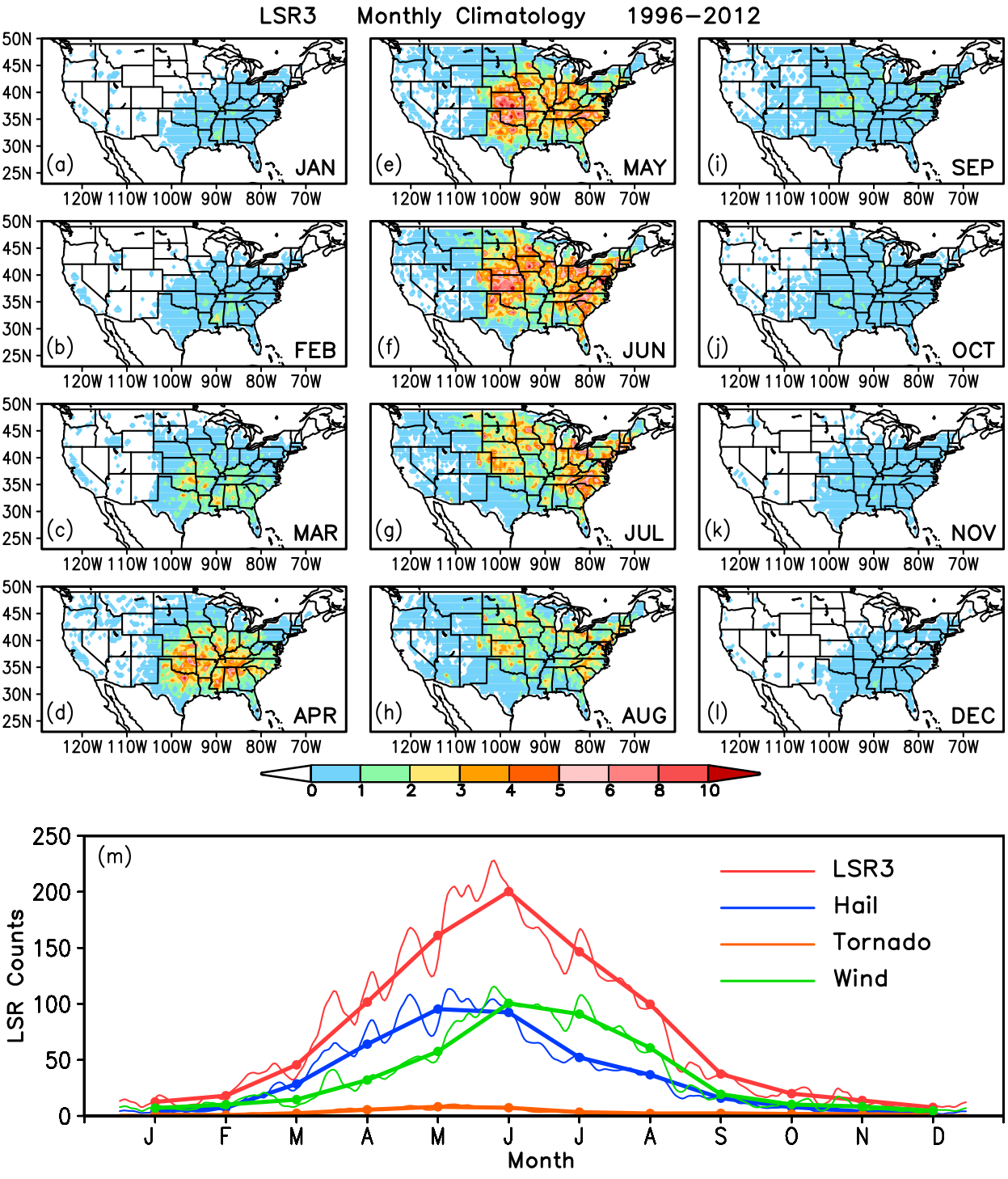

FIG. 1. Climatological monthly total number of local storm reports (LSR3), including hail events, tornadoes, and damaging wind events from (a) January to (l) December in the period of 1996–2012, and (m) climatological time series of monthly mean (thick curve with dots) and weekly mean (7-day running mean, thin curves) daily total LSR counts over the United States from January (J) to December (D), with red for LSR3, blue for hail, orange for tornado, and green for damaging wind.

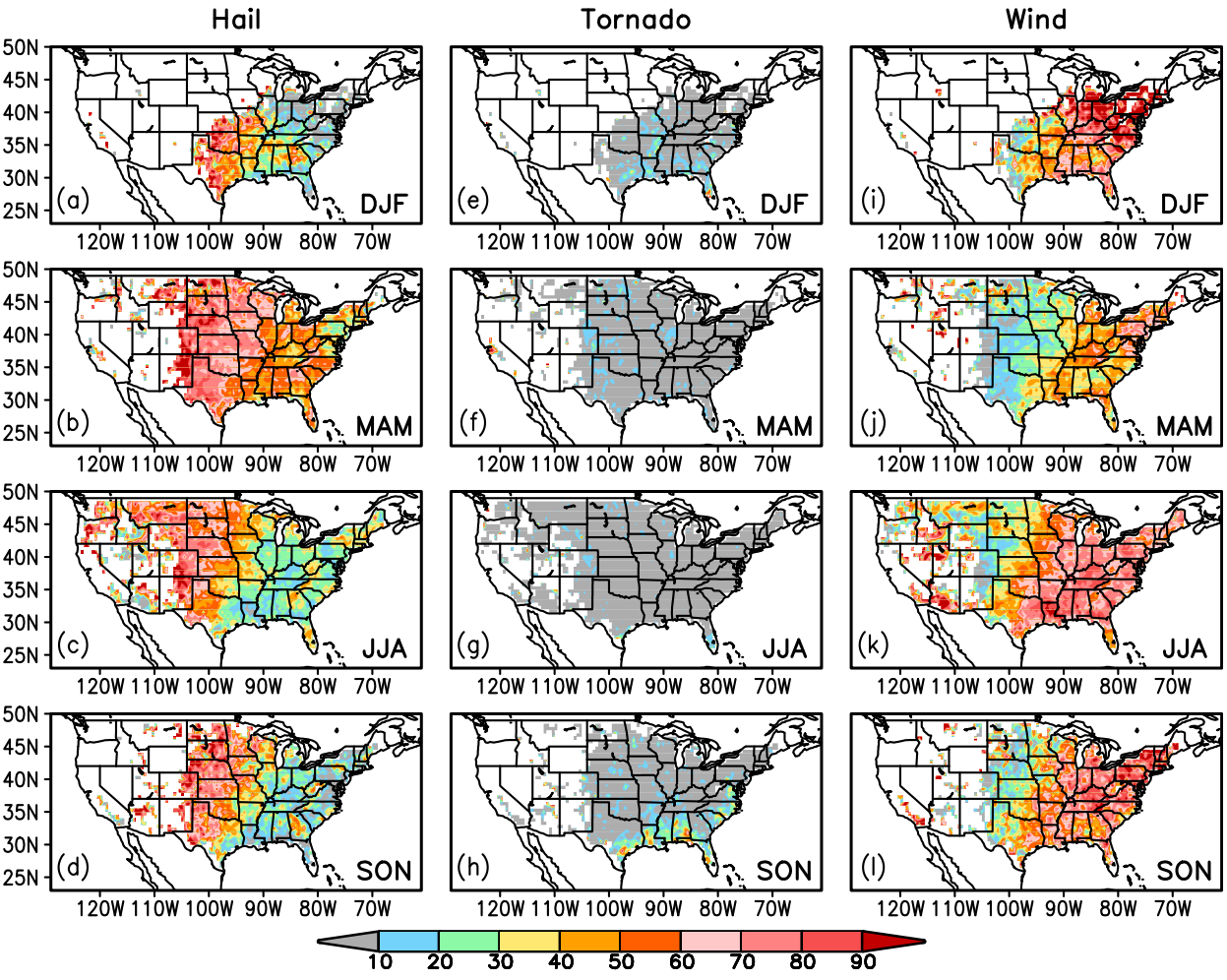

FIG. 2. Ratio (%) of climatological seasonal total number of individual type of severe weather to the climatological seasonal total number of LSR3 for (left) hail, (center) tornado, and (right) damaging wind, respectively, in (a),(e),(i) DJF; (b),(f),(j) MAM; (c),(g),(k) JJA; and (d),(h),(l) SON.

at 0000 UTC and 4 days apart with a total of 455 (91 × 5) 16-day forecasts each year from 1996 to 2012. The SCP derived from the GEFS hindcasts is the five-member ensemble mean value. For the real-time prediction of LSR3, the operational GEFS products are used, which consist of 20 runs initialized every 6 h with a daily total of 80 ensemble member forecasts. The model predicted week-1 (week-2) mean SCP is the average of 7-day SCP from the GEFS day-1 to day-7 (from day 8 to day 14) forecasts.

The statistical relationship between the GEFS predicted SCP and observed LSR3 is the basis for the dynamical–statistical prediction, and the forecast skill of the hybrid model for LSR3 largely depends on the strength of this relationship. There are at least two ways to establish the relationship. One is the simple linear regression method, in which the relation between SCP and LSR3 is found for each grid point, based on the SCP and LSR3 values only at that grid point over the GEFS hindcast period. Therefore, the SCP–LSR3 relation based on the linear regression is not directly affected by the SCP values and severe weather frequency at adjacent grid points.

There are significant changes in the statistics of the LSR3 dataset over time (e.g., Tippett et al. 2015). They were mostly caused by nonmeteorological variability, such as observer availability, including population and local forecast office practices, and changes in observation procedures, such as changing hail thresholds (Verbout et al. 2006; Anderson et al. 2007; Allen and Tippett 2015; Edwards et al. 2018; Potvin et al. 2019). The nonmeteorological changes affect the quality of LSR3, such as trends, especially in early years (1950–80). Since the LSR3 data used in this study are over the GEFS hindcast period from 1996 to 2012, the effect of the nonmeteorological factor on the data quality is expected to be small. Additionally, damaging wind reports may underrepresent or overrepresent the severity of damage caused by severe weather (Trapp et al. 2006). Therefore, caution should be taken when using the damaging wind data.

The second method is the singular value decomposition (SVD) technique (e.g., Bretherton et al. 1992; Ting and Wang 1997; Wang and Ting 2000), which can objectively identify the

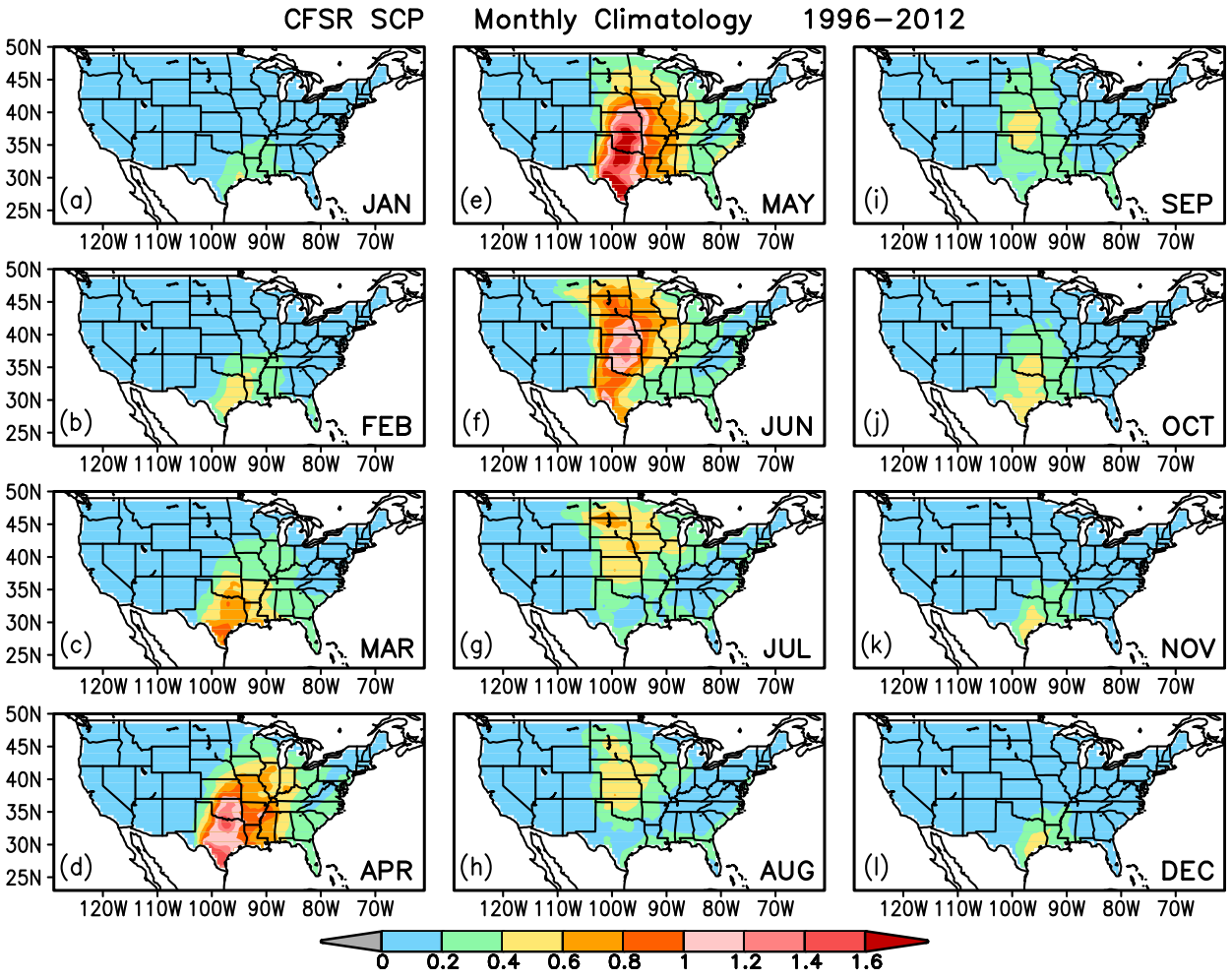

FIG. 3. Climatological monthly mean daily SCP from (a) January to (l) December derived from the CFSR data in the period of 1996–2012.

covarying spatial patterns of SCP and LSR3 with maximum temporal covariance between the two fields. In this method, an SVD analysis is performed using observed weekly mean LSR3 (predictand) and GEFS predicted week-2 SCP (predictor), based on the covariance matrix between the two data fields over the GEFS hindcast period (1996–2012). Each SVD mode consists of two spatial patterns and two time series for SCP and LSR3, respectively. The relationships between GEFS SCP and LSR3 are represented by the three leading SVD modes. Real-time week-2 LSR3 can be predicted based on the SCP–LSR3 relationships depicted by the SVD analysis over the GEFS hindcast period, as well as the real-time GEFS week-2 SCP forecast, a methodology similar to Wang et al. (1999) and Pan et al. (2018) for predicting precipitation. Specifically, the SVD-based hybrid model first projects the real-time week-2 GEFS forecast SCP onto the three SCP SVD modes (SCP spatial patterns). Associated with each SVD mode, the SCP projection coefficient is then multiplied by the correlation coefficient between the corresponding SCP and LSR3 SVD time series to obtain an LSR3 projection coefficient. Finally, the week-2 LSR3 anomalies associated with that SVD mode is obtained from the regressing LSR3 pattern against the SVD LSR3 time series, multiplied by the LSR3 projection coefficient.

The forecast skill for week-2 severe weather is cross validated over the GEFS hindcast period (1996–2012). In this procedure, the forecast target year is removed from the data for training the hybrid model with both the simple linear regression method and the SVD method. The forecasts for the target year are then made with the statistical model based on the training data taken from the rest of 16 years. The same procedure is repeated for each target year from 1996 to 2012. The forecast skill is evaluated by the anomaly correlation between the predicted and observed weekly LSR3. The statistical significance of the anomaly correlation is determined by the two-tailed t test (Snedecor and Cochran 1989). To derive a weekly (7-day) anomaly, 17-yr (1996–2012) weekly climatologies for observations and model forecasts with different leads are constructed using observational data and model hindcast data, respectively, for each 7-day period of a calendar year with a total of 365 weekly (7-day) climatologies. In this way, the annual cycle is removed from the weekly anomalies. For a

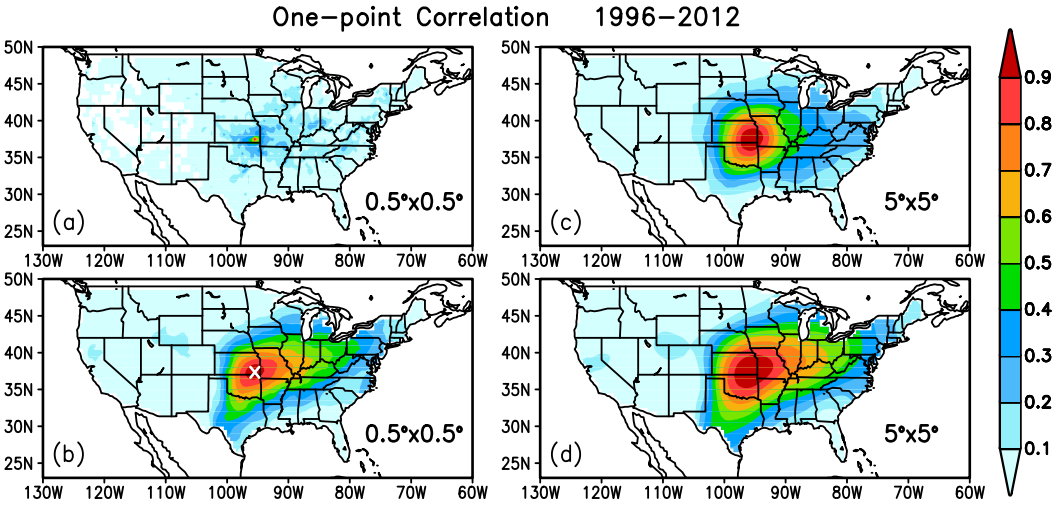

FIG. 4. One-point correlation maps for weekly (a) LSR3 and (b) SCP anomalies with data at 0.5° × 0.5° grid and weekly (c) LSR3 and (d) SCP anomalies with data averaged in the 5° × 5° box, respectively, over the period of 1996–2012. The base time series is selected at grid point (37.5°N, 95.5°W), which is marked with "×" in (b).

comparison purpose, the forecast skill for week-1 severe weather is also displayed.

3. Seasonality and spatial coherence of severe weather

The seasonality of severe weather over the United States is examined first. Figure 1 shows the observed climatological monthly total severe weather events (i.e., LSR3) from January to December, respectively. The seasonal variation of LSR3 is characterized by frequent severe weather activity in spring and summer (Figs. 1c–h) and less frequent activity in winter and fall (Figs. 1a,b,i–l). Regions of the highest activity are in the central and eastern United States during May and June (Figs. 1e,f) with maximum LSR3 greater than 10. The U.S. severe weather

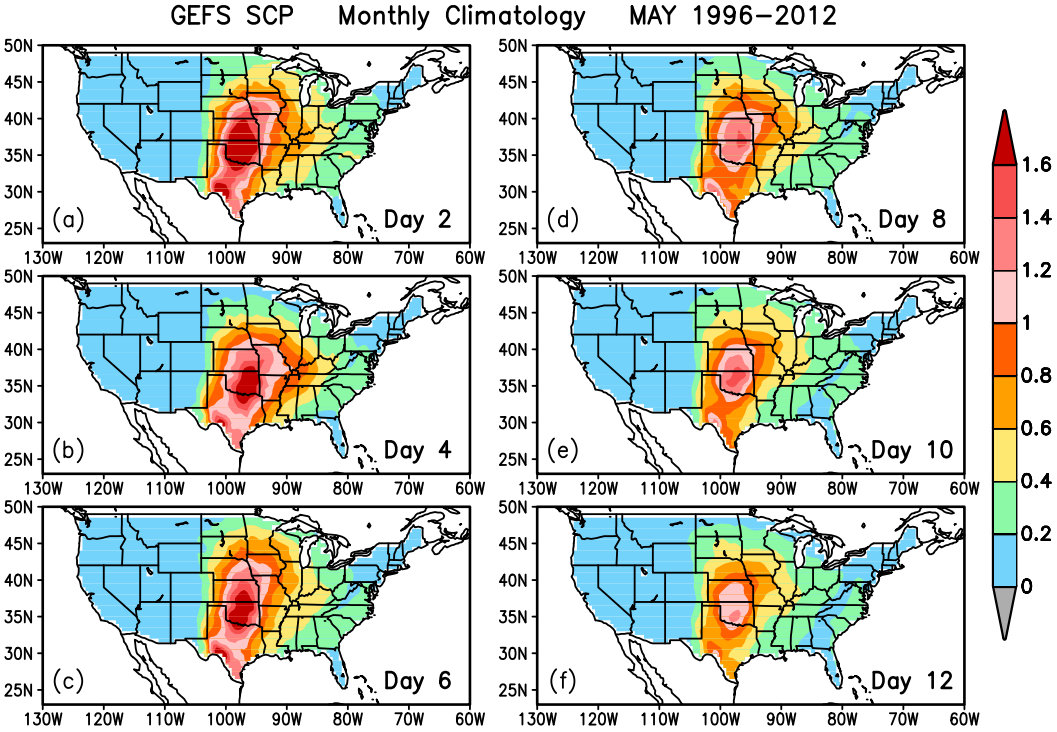

FIG. 5. Climatological monthly mean daily SCP in May derived from the GEFS hindcast data in the period of 1996–2012 for different lead times from (a) 2 to (f) 12 days.

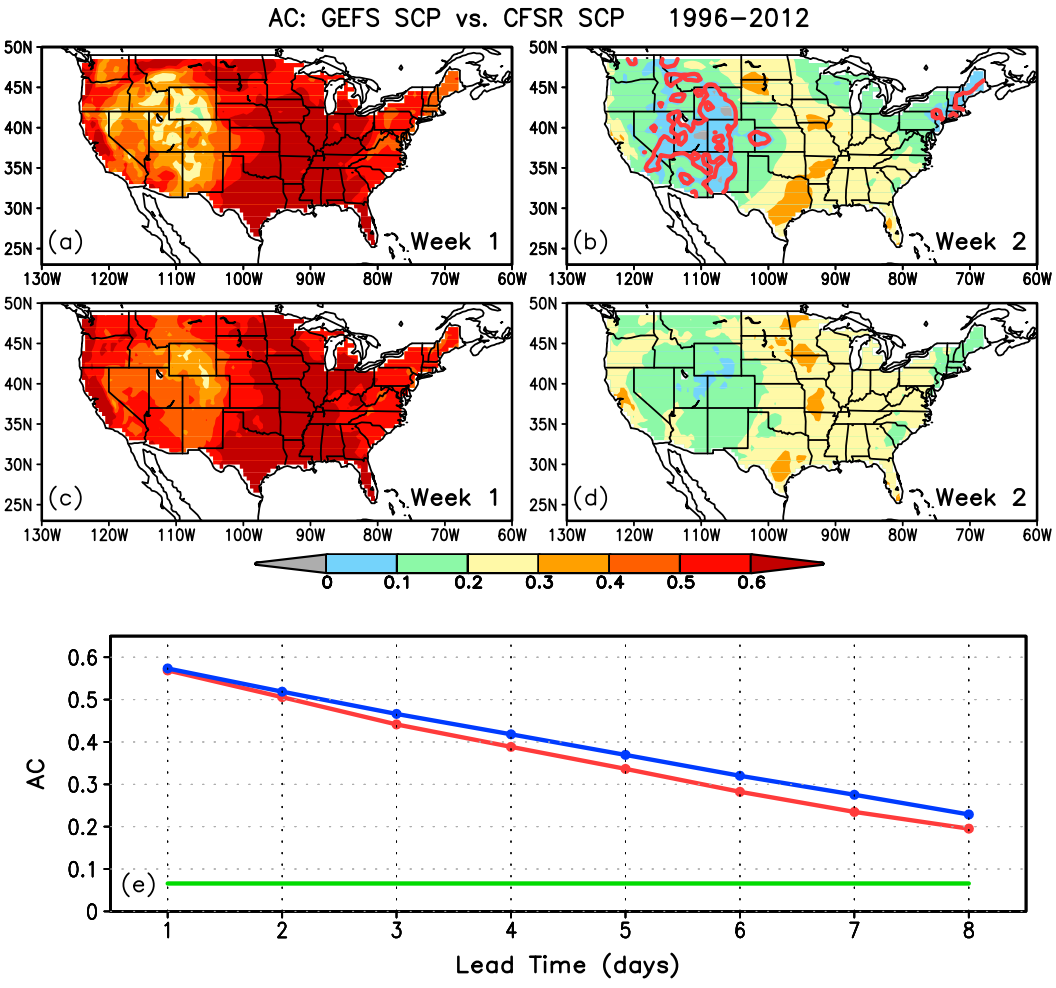

FIG. 6. Anomaly correlation (AC) between SCPs from CFSR and the GEFS hindcasts for (a) week 1 and (b) week 2 with the Pearson correlation coefficient and (c) week 1 and (d) week 2 with the Spearman rank correlation coefficient, respectively, over the 1996–2012 period, and (e) AC averaged over the central and eastern United States (east of 105°W) for weekly (7-day mean) SCP at different lead times from 1 day (week-1 forecast) to 8 days (week-2 forecast) with red for the Pearson correlations and blue for the Spearman rank correlations. Red curves in (b) and the green line in (e) denote the 99% significance level estimated by the two-tailed *t* test.

frequency thus displays strong seasonal and geographical dependence, consistent with the severe weather monthly climatology documented in previous studies (e.g., Cintineo et al. 2012; Tippett et al. 2015).

Figure 1m shows the climatological monthly mean (thick red curve with dots) and weekly mean (7-day running mean, thin red curve) daily total LSR3 counts over the United States from January to December, as well as those for hail (blue), tornado (orange), and damaging wind (green), respectively. The monthly mean daily LSR3 counts peak in June. Both hail and damaging wind events significantly contribute to this peak. The monthly mean hail counts have a maximum in May, whereas the monthly mean damaging wind events have a maximum in June. Overall, the weekly (7-day mean, thin curves) severe weather frequency shows consistent seasonal variations as the monthly means. Some fluctuations from one 7-day window to the next 7-day window are also observed.

The contributions of individual types of severe weather to the observed climatological LSR3 are shown in Fig. 2 for winter [December–February (DJF)], spring [March–May (MAM)], summer [June–August (JJA)], and fall [September–November (SON)], respectively. In general, hail dominates the severe weather frequency (greater than 60%) over the central United States in all seasons (Figs. 2a–d), as well as over the eastern United States in spring (Fig. 2b). In contrast, damaging winds mainly dominate over the eastern United States (Figs. 2i–l). The tornado frequency accounts for less than 10% of total severe weather events over most of the United States (Figs. 2e–h). The tornado frequency is relatively high in the Gulf States and the Southeast during fall (Fig. 2h), likely associated with both synoptic-scale frontal systems and landfalling Atlantic tropical storms and hurricanes (e.g., Brooks et al. 2003a; Moore and Dixon 2011).

Figure 3 shows the climatological monthly mean daily SCP from January to December, derived from the CFSR data.

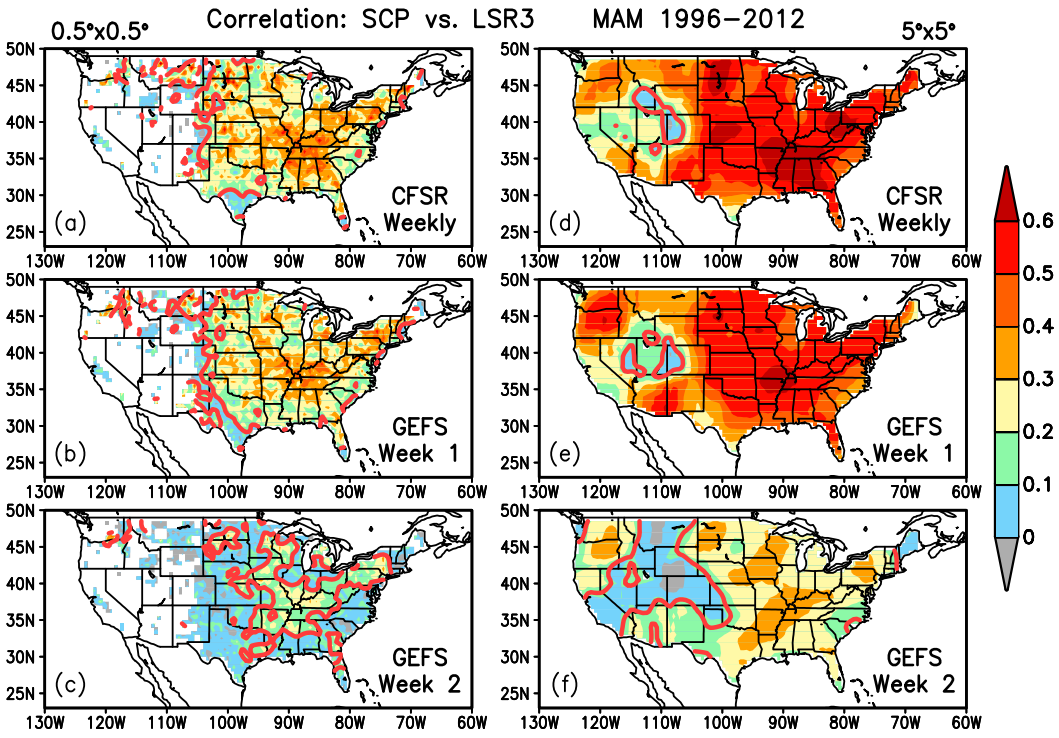

FIG. 7. Correlation between observed weekly LSR3 and weekly SCP derived from (a) CFSR and the GEFS hindcasts for (b) week 1 and (c) week 2 with (left) anomalies at the 0.5° × 0.5° grid, and also correlation between observed weekly LSR3 and weekly SCP derived from (d) CFSR and the GEFS hindcasts for (e) week 1 and (f) week2 with (right) anomalies area averaged over the 5° × 5° box, respectively, during MAM 1996–2012. The red curve denotes the 99% significance level estimated by the two-tailed *t* test.

The seasonal variation of SCP is characterized by relatively large values of SCP (~0.2–0.4) in the Gulf States during winter months (Figs. 3a,b,l). SCP then increases during spring (Figs. 3c–e) and peaks in May (Fig. 3e). In the meantime, the region of large SCP moves northward from the Southern Plains in spring to the Northern Plains in summer (Figs. 3c–g). From the summer to the following winter (Figs. 3h–l), the SCP value decreases and the center of the maximums moves back to the south. The seasonal progression of SCP in the central United States is similar to that of LSR3 (Fig. 1) and also consistent with that documented in the literature (e.g., Brooks et al. 2003b; Gensini and Ashley 2011; Tippett et al. 2015). Therefore, in terms of seasonal cycle, there is a good correspondence between severe weather and SCP in the central United States. However, during spring and summer, there is also frequent severe weather in the eastern United States (Figs. 1c–h) where SCP values are generally small (Figs. 3c–h). Since the majority of the severe weather events in the eastern United States are damaging winds (Fig. 2, right panels), which are tied to the predominant convective mode or QLCS (Ashley et al. 2019), the SCP more precisely indicates supercell thunderstorm environments, which produce mainly tornadoes and hails, and does not well identify QLCS environments, which are the major wind producers.

The scale of spatial coherence of severe weather over the United States is demonstrated by a one-point-correlation map, where the weekly LSR3 anomaly at a selected point (here 95.5°W, 37.5°N) is correlated with weekly LSR3 at every grid point over the United States in the period of 1996–2012, as shown in Fig. 4a. The correlations of weekly LSR3 at a 0.5° × 0.5° grid with those in the surrounding areas are generally small, except for the correlation with the base point itself. Therefore, weekly severe weather activities within a 0.5° × 0.5° area are largely isolated events with small spatial coherence. In contrast, the one-point-correlation map for SCP in CFSR at the 0.5° × 0.5° grid (Fig. 4b) shows higher correlations between the selected grid point and the surrounding grid points, indicating that SCP has a large-scale feature and high spatial coherence. It is interesting to note that when averaging LSR3 over a 5° × 5° box and then recalculating the one-point correlation, the result (Fig. 4c) shows much higher spatial coherence for LSR3. Its large-scale structure is comparable to that of the 5° × 5° area-averaged SCP (Fig. 4d). The increase in spatial coherence for LSR3 from the 0.5° × 0.5° grid to the 5° × 5° box (Figs. 4a,c) is much more significant than that of SCP (Figs. 4b,d), which already shows high spatial coherence at the 0.5° × 0.5° grid (Fig. 4b). Therefore, it is reasonable to expect that weekly severe weather over a larger domain may have higher predictability than that over a small area. It should be noted that the increase in correlation is at the expense of spatial specificity. The area-averaging procedure may enhance the predictability of the severe weather driven by the large-scale

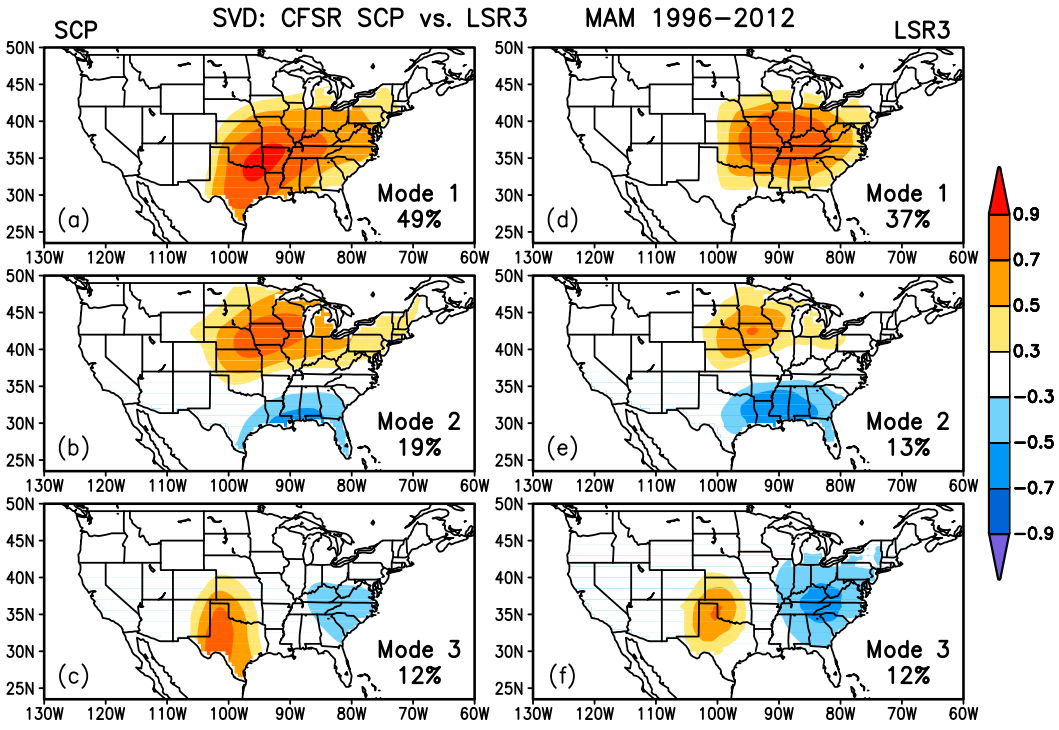

FIG. 8. Maps of homogeneous correlation for the first three SVD modes between 5° × 5° area-averaged weekly CFSR SCP and observed LSR3 over the United States during MAM 1996–2012. The percentage of the variance explained by each SVD mode is also provided at the bottom right of each panel for (left) SCP and (right) LSR3.

environment as in strongly forced, synoptically evident events by smoothing out weakly forced events or events driven by the mesoscale forcing.

4. GEFS forecast skill for SCP

Because the GEFS predicted SCP will be used as the predictor for LSR3, whether the model can skillfully forecast SCP is assessed. The climatological monthly mean daily SCP derived from the GEFS hindcasts displays very similar seasonal variations to the observations (Fig. 3) for lead times from 1 to 14 days (not shown). The amplitude of the GEFS-predicted mean SCP, however, decreases with lead time. Consistent with the decrease in monthly SCP climatology with lead time, the standard deviation of monthly SCP in the model also decreases with lead time (not shown). For example, Fig. 5 shows the long-term mean SCP in May, when the SCP is largest, at different lead times from 2 to 12 days based on the GEFS hindcasts. The 2-day forecasts of the mean SCP in May (Fig. 5a) is comparable to the observed (Fig. 3e) in terms of both spatial distribution and magnitude. The maximum value of the SCP decreases from above 1.6 at the 2-day lead (Fig. 5a) to 1.2 at the 12-day lead (Fig. 5f), about 25% reduction over the 10-day difference in lead time. Despite this, the GEFS model captures the monthly mean daily SCP reasonably well.

The GEFS forecast skill for SCP is assessed by point-to-point anomaly correlation (AC) between GEFS SCP and CFSR SCP. As expected, the forecast skill decreases with lead time from 1 day to 14 days (not shown). Figure 6 shows the anomaly correlation between weekly (7-day) mean daily SCPs from CFSR and GEFS hindcasts for week 1 (1-day lead) and week 2 (8-day lead), respectively, over the 1996–2012 period. Since SCP may be sporadic in time, the ACs are calculated using both the Pearson correlation coefficient (Figs. 6a,c) and the Spearman rank correlation coefficient (Figs. 6b,d). Overall, the correlations obtained with the two methods are close to each other. Consistent with the AC skill for daily SCP forecasts, the week-2 forecast skill (Figs. 6b,d) is much lower than the week-1 skill (Figs. 6a,c). However, the week-2 AC skill is still above the 99% significance level in the central and eastern United States, as estimated by the two-tailed t test (Snedecor and Cochran 1989).

The dependence of the AC skill for weekly mean SCP upon lead time is further examined in Fig. 6e with a mean AC averaged over the central and eastern United States (east of 105°W) at different lead times from 1 day to 8 days. The mean AC decreases from 0.57 at 1-day lead to 0.20 and 8-day lead, all above the 99% significance level (green line in Fig. 6e). When using the GEFS predicted SCP as a model guidance for the large-scale convective environment (e.g., Carbin et al. 2016), the performance of this approach will depend highly on the GEFS skill in predicting the SCP, which is modest for week 2, as shown in Fig. 6. We will show in the next section that the performance of the hybrid model approach largely depends on the correlation between the GEFS SCP and observed LSR3, rather than the GEFS skill for SCP. Thus, given the limited

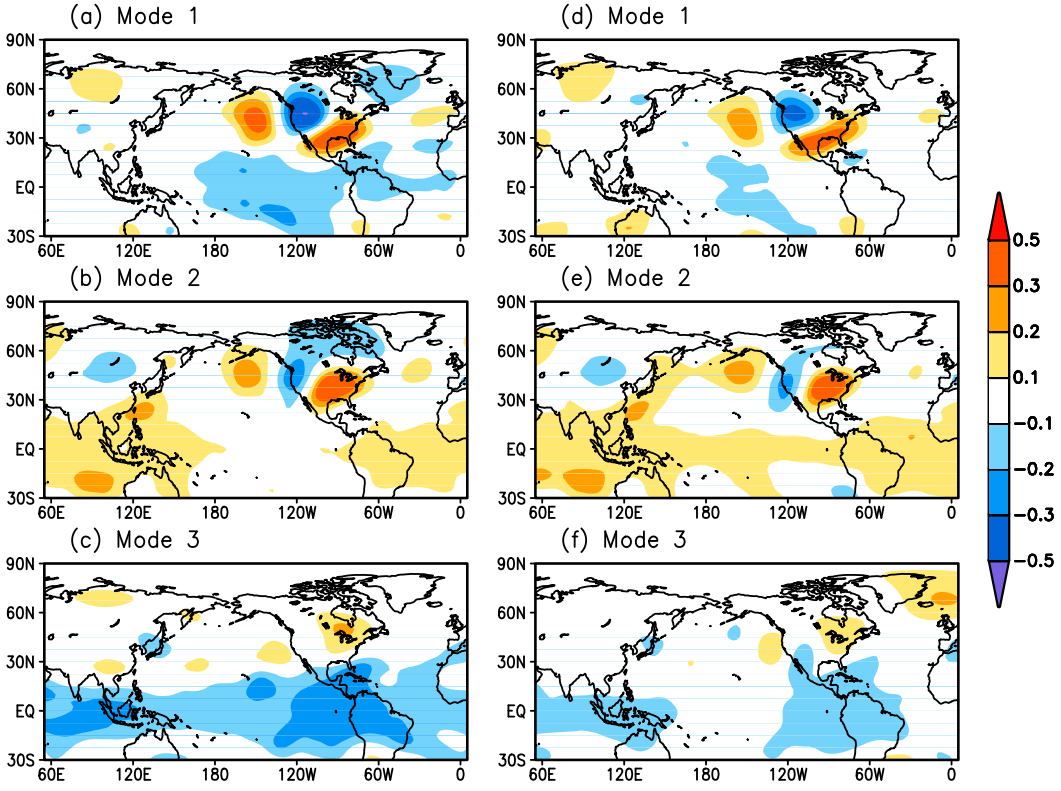

FIG. 9. Correlation of weekly 200-hPa height anomaly with (left) SVD CFSR SCP and (right) LSR3 time series, respectively, over MAM 1996–2012 for (a),(d) mode 1; (b),(e) mode 2; and (c),(f) mode 3.

skill of GEFS for week-2 SCP, a hybrid model is used to produce skillful forecast for week-2 LSR3.

5. Empirical relationship between SCP and LSR3

a. SCP–LSR3 relationship depicted by simple linear correlation

To develop a hybrid forecast model, a statistical relationship between GEFS predicted SCP and observed LSR3 needs to be first established. Given the strong seasonality of both SCP and SLR3 (Figs. 1 and 3; Brooks et al. 2003a, their Fig. 3) and seasonal dependence of the association between storms and convective parameters (Hart and Cohen 2016), a 3-month moving window is used in the analysis. For example, MAM data are used to quantify the relationship for the forecast target month of April.

Figures 7b and 7c show the correlations between observed weekly LSR3 and GEFS week-1 and week-2 forecasts of SCP, respectively, at each 0.5° × 0.5° grid point over MAM 1996–2012. For comparison, the SCP–LSR3 relationship in observations is also presented in Fig. 7a. The relationships between the observed weekly LSR3 and model predicted week-1 SCP (Fig. 7b) are slightly weaker than those in observations (Fig. 7a), with correlations ranging from 0.2 to 0.3 over most of the central and eastern United States and exceeding the 99% significance level estimated by the two-tailed t test. However, the correlations of LSR3 with the GEFS predicted week-2 SCP (Fig. 7c) is much weaker than the observations and the week-1 forecasts. Note that the SCP–LSR3 relationship is weak in south Texas as compared to other regions to the east of the Rockies. As discussed by Gensini and Ashley (2011), a low LSR3 frequency (Fig. 1) in this region may be responsible for the weak relationship to SCP.

Similar relationships between LSR3 and SCP are reestablished using the 5° × 5° area-averaged anomalies, also shown in Fig. 7 (right panels). The correlations in the right panels are much higher than the left panels for both week 1 (Fig. 7e) and week 2 forecasts (Fig. 7f), as well as in observations (Fig. 7d). The results indicate stronger correlations between LSR3 and the model predicted SCP when considering the severe weather over a larger domain, or relaxing the temporal and spatial constraints of both predictor and predictand (Gensini et al. 2020). It may also imply that the mean LSR3 in the 5° × 5° box is more controlled by the large-scale convective environments than the LSR3 in a 0.5° × 0.5° box. The correlations between LSR3 and SCP for other 3-month windows (not shown) are generally comparable to MAM presented in Fig. 7.

b. SCP–LSR3 relationship identified by leading SVD modes

In addition to the local relationship between SCP and LSR3 at each grid point, their empirical relationship can also be established by the SVD technique (Bretherton et al. 1992).

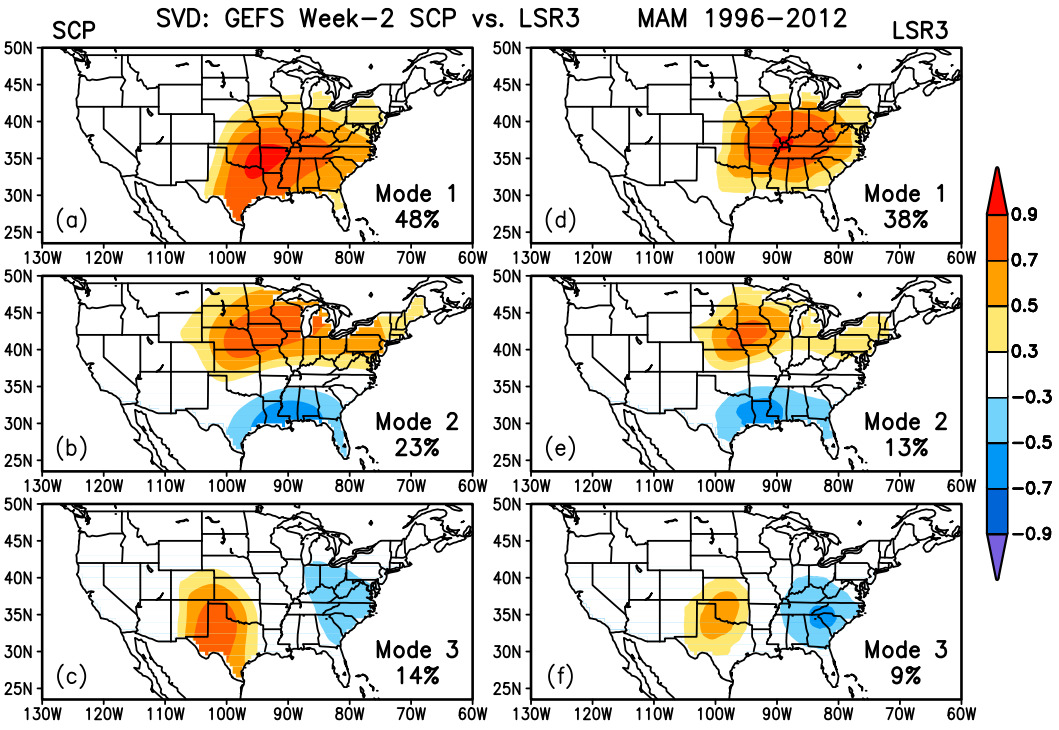

FIG. 10. As in Fig. 8, but for the three leading SVD modes between the GEFS week-2 SCP and observed LSR3.

This method can objectively identify pairs of modes (spatial patterns) for SCP and LSR3, both of which vary with maximum temporal covariance between the two fields. As shown in Fig. 4, the 5° × 5° area-averaged anomalies have higher spatial coherence than the 0.5° × 0.5° grid data, especially for LSR3. Therefore, both the 5° × 5° area-averaged SCP and LSR3 data over the United States are used in the SVD analysis to enhance their covariations spatially and temporally.

Figure 8 presents the spatial patterns of the three leading SVD modes for weekly SCP (left) and LSR3 (right), respectively, using the observational data in MAM 1996–2012. Each SVD mode is characterized by a distinctive pattern with consistent spatial distributions of SCP and LSR3. The first mode displays a monopole structure with above-normal (below-normal) SCP linked to above-normal (below-normal) severe weather frequency in the central and Southeast United States (Figs. 8a,d). In contrast, both the second and third modes exhibit a dipole structure in the meridional (Figs. 8b,e) and zonal (Figs. 8c,f) directions, respectively. The second mode suggests that positive SCP anomalies in the Midwest and negative anomalies in the Gulf States are associated with more severe weather in the Midwest and less severe weather in the Gulf States (Figs. 8b,e), and vice versa. The third mode indicates that positive SCP anomalies in the South and negative anomalies in the Southeast are related to above-normal and below-normal severe weather activities, respectively, in the two regions (Figs. 8c,f). The three modes account for 62% of total weekly LSR3 variance.

To better understand the physical mechanisms leading to the three modes, weekly 200-hPa height anomalies are correlated with the two time series of each SVD mode, and the correlation maps are shown in Fig. 9. Each mode shows distinctive correlations with local height anomalies, and with height anomalies in remote areas as well, such as in the tropics. In particular, there are well-defined wave train patterns in the 200-hPa height field over the Pacific/North American region associated with both mode 1 (Figs. 9a,d) and mode 2 (Figs. 9b,e). These wave trains originate from the tropics, indicating possible tropical forcing as their source. Superimposed onto the background weekly mean flow, the wave train alters the large-scale circulation, such as jet streams, and thus may modulate severe weather.

A similar SVD analysis using the GEFS week-2 forecasts of SCP reproduces the observed relationship between SCP and LSR3 well (Fig. 10). Together, the three modes account for 60% of total weekly LSR3 variance, comparable to that (62%) in observations. The results of the SVD analysis using the GEFS week-1 SCP and observed LSR3 (not shown) are similar to those in Fig. 10. Furthermore, the first three modes explain 38%, 13%, and 9% of total weekly LSR3 variance, respectively, similar to the observations.

6. Dynamical–statistical forecast of week-2 severe weather

a. Forecast skill assessed through cross validation

A hybrid model is developed to forecast the number of week-2 severe weather events (LSR3) using the GEFS forecast week-2 SCP as a predictor and based on their empirical relationships depicted in either Fig. 7 or Fig. 10. The former applies a linear regression model to forecast LSR3 at each grid point, whereas the latter projects the week-2 GEFS SCP onto

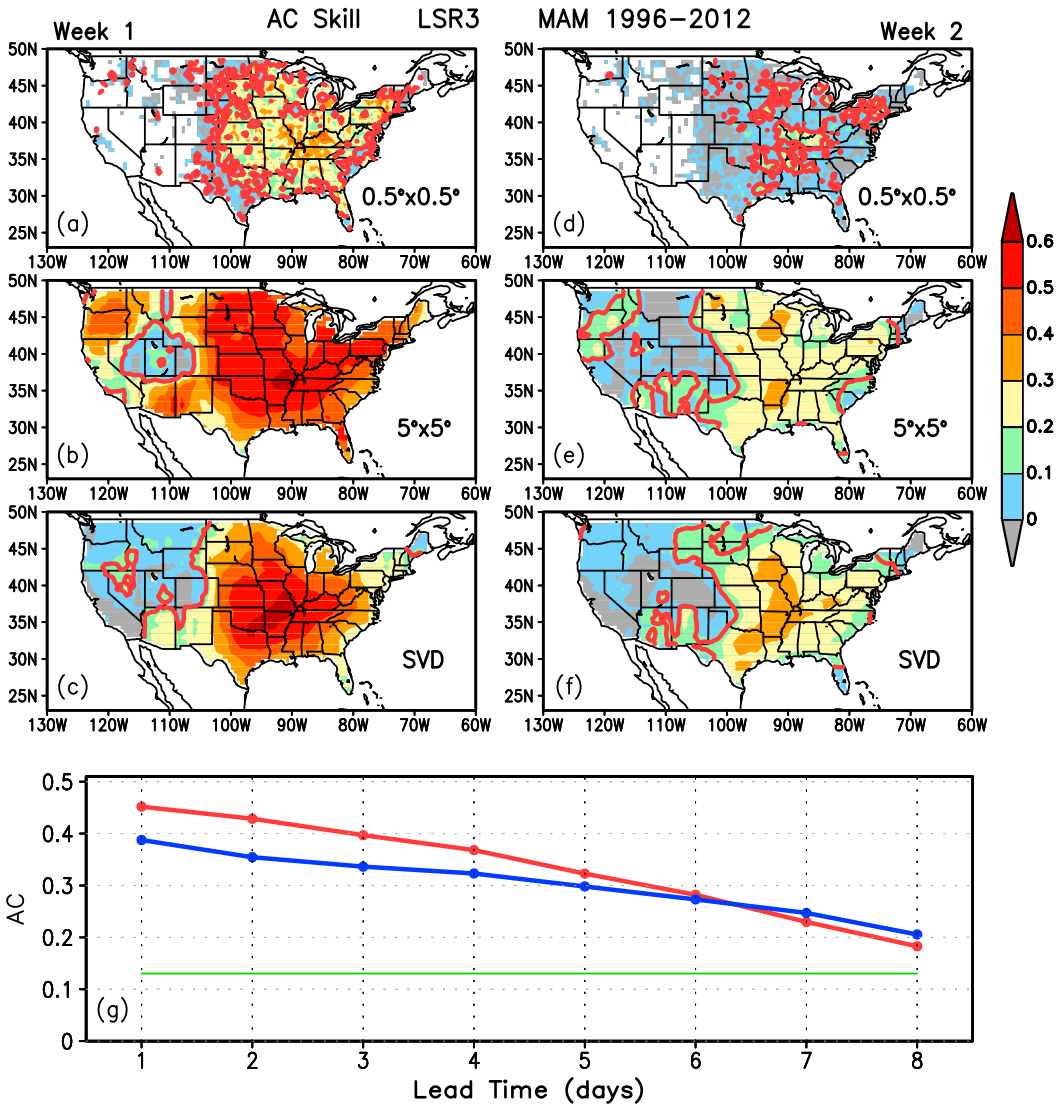

FIG. 11. Forecast skill measured by the anomaly correlation (AC) between the observed and predicted weekly LSR3 for (left) week 1 and (right) week 2 using the simple linear regression model with (a),(d) the anomalies in the 0.5° × 0.5° grid box and (b),(e) the anomalies averaged over the 5° × 5° box; (c),(f) using the SVD-based forecast model also with the anomalies averaged over the 5° × 5° box; and (g) AC averaged over the central and eastern United States (east of 105°W) for weekly (7-day mean) LSR3 at different lead times from 1 day (week-1 forecast) to 8 days (week-2 forecast) with the red curve for the simple linear regression model and the blue curve for the SVD-based model. The AC skill is assessed based on cross validations over the MAM of 1996–2012 with a total of 391 weeks. Red curves in (a)–(f) and the green line in (g) denote the 99% significance level estimated by the two-tailed *t* test.

the first three SCP SVD modes and then predicts LSR3 based on the SCP–LSR3 relationships depicted by the SVD analysis (Fig. 10). The forecast skill for week-2 severe weather is cross validated over the GEFS hindcast period (1996–2012).

Figure 11 shows the AC skill of both week-1 (left) and week-2 (right) LSR3 forecasts for MAM 1996–2012 with the two methods. When applying the linear regression model using the 0.5° × 0.5° grid data, the AC skill ranges from 0.2 to 0.4 and from 0 to 0.2 for week-1 (Fig. 11a) and week-2 (Fig. 11d) forecasts, respectively, over most of the central and eastern United States. By averaging the data over the 5° × 5° box and then using the linear regression model, the AC skill is improved appreciably for both week 1 (0.3–0.6, Fig. 11b) and week 2 (0.1–0.3, Fig. 11e) over these regions. The AC skill of the linear regression model is mainly determined by the strength of the empirical relationship between SCP and LSR3, and is thus similar to the corresponding correlation map between GEFS SCP and LSR3 (Fig. 7b versus Fig. 11a, Fig. 7c versus Fig. 11d, Fig. 7e versus Fig. 11b, and Fig. 7f versus Fig. 11e).

The AC skill is further improved (Figs. 11c,f) when using the SVD-based SCP–LSR3 relationships. Clearly, there is spatial

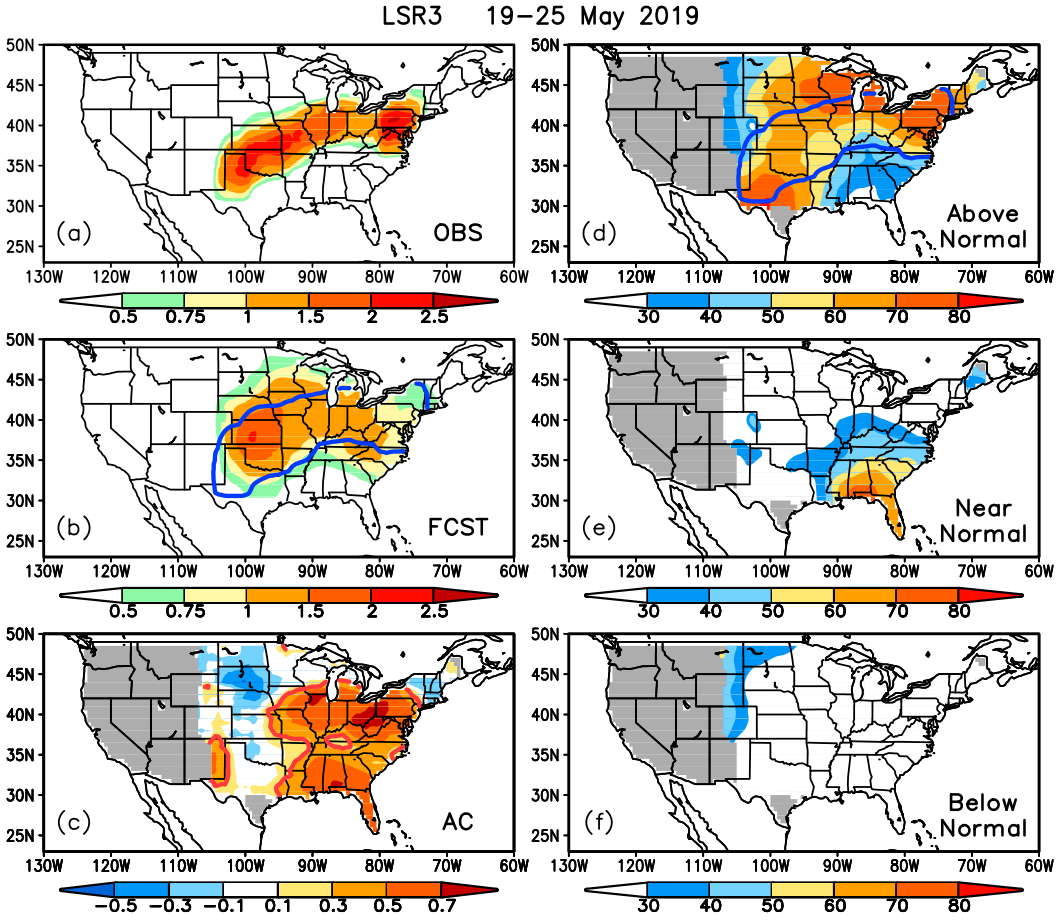

FIG. 12. (a) Observed and (b) linear regression model predicted LSR3 for the week of 19–25 May 2019, (c) anomaly correlation (AC) skill of the real-time forecasts for weekly LSR3 during MAM 2019, and probability forecasts of LSR3 for (d) above-normal, (e) near-normal, and (f) below-normal categories for the week of 19–25 May 2019, based on 80 ensemble members of the GEFS SCP forecasts. The red curve in (c) denotes the 99% significance level estimated by the two-tailed t test. The blue contour in (b) and (d) is 0.5 of the observed weekly LSR3 shown in (a). Regions with observed LSR3 climatology less than 0.05 are masked out by gray shading in (c)–(f).

variability in skill between the two statistical techniques. For the week-1 forecast, the SVD-based forecast skill (Fig. 11c) is better than the linear regression model (Fig. 11b) in the central United States, but worse in the Upper Midwest, East Coast, and Northeast. For the week-2 forecast, the SVD-based forecast skill (Fig. 11f) is better than the linear regression model (Fig. 11e) in the Lower Midwest, South, and Southeast regions. These improvements are likely due to the inclusion of the spatial covariations of both SCP and LSR3 with their surrounding areas in the SVD relationship, whereas the linear regression model is only based on the relationship determined by the SCP and LSR3 values at individual grid points.

To examine the dependence of the AC skill on lead time, the forecasts of weekly (7-day) LSR3 at lead times from 2 to 7 days are also made with both the linear regression model and the SVD-based model. Figure 11g shows the AC skill averaged over the central and eastern United States (east of 105°W) as a function of lead time, including both the week-1 (1-day lead) and week-2 (8-day lead) forecasts. The linear regression model (red curve) performs better at short leads (1–6 days), whereas the SVD-based model (blue curve) performs better at long leads (7–8 days). Both models show a mean AC skill of week-2 (8–14 days) severe weather forecast above the 99% significance level (green line). In contrast, Gensini and Tippett (2019) found that *daily* hail and tornado activities can be skillfully predicted at 9–12-day leads.

b. *Real-time forecast for spring 2019*

The hybrid dynamical–statistical tool with both the linear regression model and the SVD-based model, and also using the 5° × 5° area-averaged data has been tested and implemented for experimental real-time forecast of week-2 severe weather in 2019. The forecast is updated at 1000 eastern standard time on a daily basis. The week-2 SCP is derived from the 80-member GEFS operational 16-day forecasts with 20 runs initialized at 0000 and 0600 UTC of the present day and 1200 and 1800 UTC of previous day, respectively. The real-time week-2 severe weather forecasts consist of both deterministic

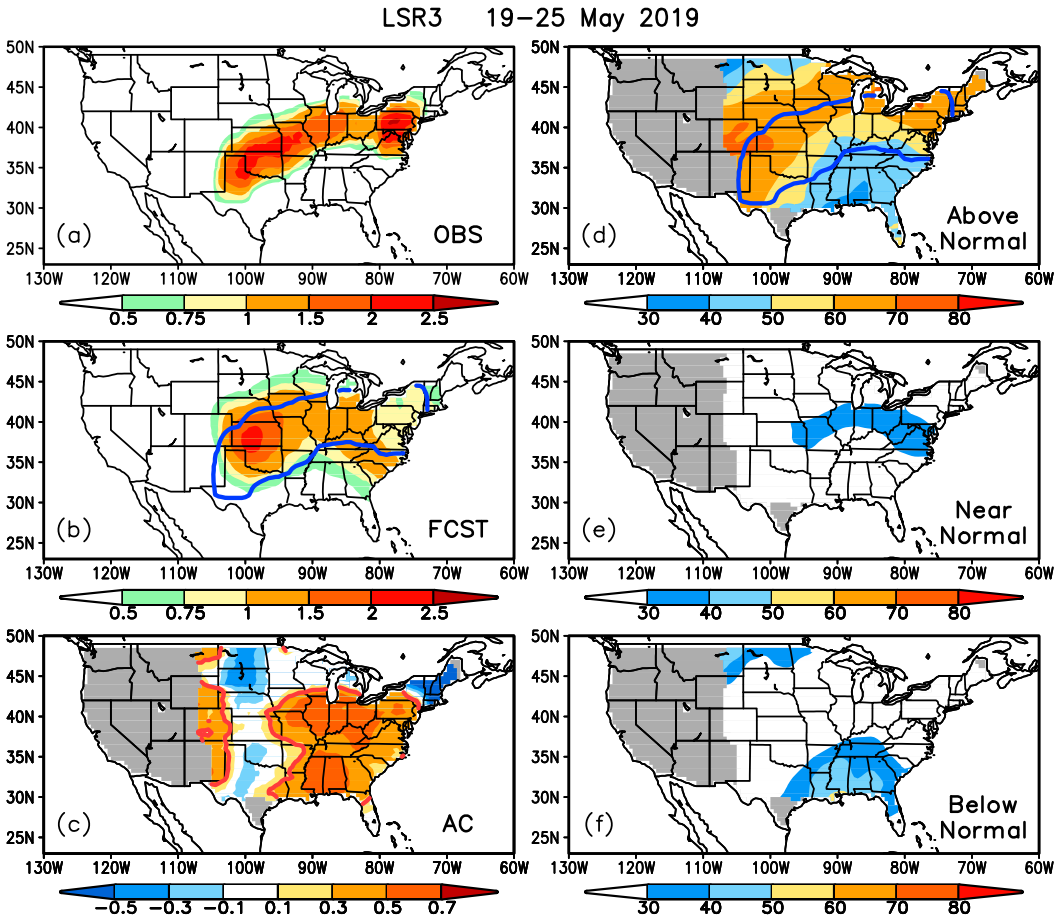

FIG. 13. As in Fig. 12, but for the SVD-based model forecast.

and probabilistic formats. The former is the 80-member ensemble mean forecast of week-2 LSR3. The latter is the percentage distribution of the 80-member forecasts in three categories, namely, above normal, near normal, and below normal. The thresholds for the three categories are determined by the week-2 LSR3 forecast values in the cross validations over the 1996–2012 period, so that each category accounts for 33.3% of the total forecast events.

Figure 12 shows an example of the real-time week-2 severe weather forecast for the week of 19–25 May 2019 issued on 12 May 2019, using the linear regression model. During that week, there was a severe weather outbreak sequence with 79 EF1–EF3 tornadoes across the central United States and the mid-Atlantic (Gensini et al. 2019). Figure 12a presents the observed total LSR3 during the 7-day period. Compared to the observations (Fig. 12a), the 80-member ensemble mean week-2 forecast indicates potential for broader but less intense severe weather activities (Fig. 12b). The predicted weekly LSR3 is slightly lower but comparable to the observed in the central United States, and much lower than the observed in the mid-Atlantic. The probabilistic forecast is shown in Figs. 12d–f, respectively, for the three categories. Most of the central and eastern United States are in the above-normal category (Fig. 12d), except the Southeast United States in the near-normal category (Fig. 12e). In particular, almost all regions of the observed LSR3 (circled by blue curve in Fig. 12d) were predicted with more than 50% chance of above-normal severe weather frequency. Gensini et al. (2019) demonstrated that the sources of the predictability for this severe weather event partially came from anomalous convection over the tropical Indian and Pacific Oceans through an atmospheric teleconnection.

Figure 12c shows the forecast skill measured by the anomaly correlation between the observed and predicted 7-day LSR3 for MAM 2019 with a total of 92 week-2 forecasts. The AC skill is generally greater than 0.3 in the central and eastern United States, which is above the 99% significance level (0.27). However, the forecast skill is low in the Great Plains, especially in the northern plains with negative AC coefficients, as the northern plains are not a climatologically favored area for LSR3 in MAM and the predictable signal is low (Fig. 1). Overall, Fig. 12c indicates considerable skill of the hybrid model in forecasting week-2 severe weather over the central and eastern United States for spring 2019.

Figure 13 presents the week-2 forecast for the same week (19–25 May 2019), but using the SVD-based model. Compared to Fig. 12, both the forecasts of week-2 LSR3 values (Figs. 12b and 13b) and the above-normal probability (Figs. 12d and 13d) are similar to each other between the two methods in this case. The AC over the entire season (MAM 2019) shows better skills

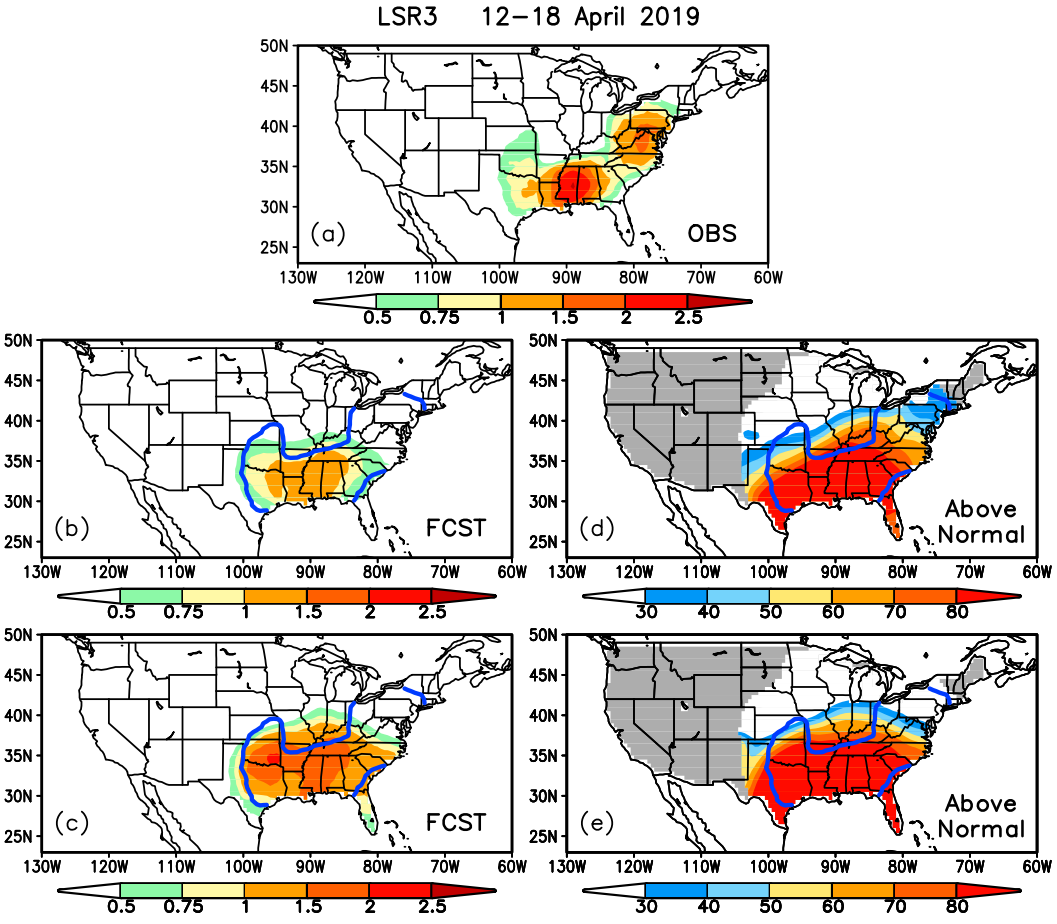

FIG. 14. (a) Observed, (b) linear regression model, and (c) SVD-based model predicted LSR3 for the week of 12–18 Apr 2019, and the above-normal probability forecast for the week of 12–18 Apr 2019 using (d) linear regression model and (e) SVD-based model. The blue contour in (b)–(e) is 0.5 of the observed weekly LSR3 shown in (a). Regions with observed LSR3 climatology less than 0.05 are masked out by gray shading in (d) and (e).

with the SVD method in west Central and Northern Plains (Fig. 13c). However, the AC skill is higher in the regions south of the Great Lakes and in some Gulf States with the linear regression model (Fig. 12c). Overall, the AC skills are comparable between the two methods for spring 2019 (Figs. 12c and 13c).

Figure 14 shows another example of week-2 LSR3 forecasts for week 12–18 April 2019 issued on 5 April 2019. During this week, widespread severe weather with tornado outbreaks stretched from the south-central United States to the East Coast (Fig. 14a). The forecasts with both methods captured the severe weather in the Gulf States, but missed that in the mid-Atlantic states (Figs. 14b,c). The forecast of weekly LSR3 is closer to the observations with the SVD technique (Fig. 14c) then with the linear regression method (Fig. 14b). However, the latter shows a better probability forecast in the mid-Atlantic region (Fig. 14d). Overall, in this case, the forecasts with the two methods are also comparable.

7. Conclusions

The development and evaluation of a hybrid dynamical–statistical model for forecasting week-2 severe weather over the United States was presented. Following the work of Carbin et al. (2016), the supercell composite parameter (SCP) was used as a predictor to represent the large-scale environments and link the dynamical model forecast to actual severe weather, an approach similar to Gensini and Tippett (2019). The performance of the hybrid model has been cross validated over the 1996–2012 period and also tested in real time for spring 2019.

The hybrid model forecasts suggest a low skill for week-2 severe weather when applying the linear regression model to the 0.5° × 0.5° grid data. Weekly severe weather over such a small area is largely isolated from its surroundings (Fig. 4a). Thus, it is mainly determined by small-scale and local conditions, but less related to large-scale convective environments. However, the forecast skill can be improved by applying the linear regression model to the 5° × 5° area-averaged anomalies. The weekly severe weather over such an extended area displays some large-scale features with higher spatial coherence, and thus, is also more closely related to the large-scale environments. It is also demonstrated that the forecast skill can be further improved by using the SVD-based statistical

relationship. The SVD method objectively picks out the dominant spatial patterns of weekly severe weather that covary with the large-scale SCP patterns. The SVD-based SCP–LSR3 relationship thus accounts for more variance of weekly severe weather than the relationships established by the simple linear regression at individual grid points.

Experimental week-2 severe weather outlooks have been tested in real time for spring 2019, using both the linear regression and the SVD approaches, and the $5° \times 5°$ area-averaged data. Both the forecasts for the weeks of 12–18 April and 19–25 May 2019, when there was a severe weather outbreak sequence affecting the central and eastern United States, and the anomaly correlation between forecasts and observations across the entire season suggest considerable skill for week-2 severe weather over the United States. It is expected that the dynamical–statistical tool developed in this study will be implemented into operations in the near future.

Acknowledgments. The authors thank Vittorio A. Gensini, Ray Wolf, an anonymous reviewer, and Editor Matthew J. Bunkers for their insightful and constructive comments and suggestions. The authors would also like to thank Stephen Baxter, Mike Charles, Mingyue Chen, Daniel Harnos, Zeng-Zhen Hu, Melissa Ou, Brad Pugh, Matthew Rosencrans, and Wanqiu Wang at the NCEP Climate Prediction Center (CPC) for helpful discussions and also Hong Guan and Yuejian Zhu at the NCEP Environmental Modeling Center (EMC) for providing GEFS model forecast data. The Local Storm Report data are provided by the NOAA/SPC, Norman, Oklahoma, from their web site at https://www.spc.noaa.gov/wcm/#data. The CFSR data are available from the NOAA/NCEI, Asheville, North Carolina, at https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/reanalysis.

REFERENCES

Allen, J. T., and M. K. Tippett, 2015: The characteristics of United States hail reports: 1955–2014. *Electron. J. Severe Storms Meteor.*, **10** (3), https://ejssm.org/ojs/index.php/ejssm/article/viewArticle/149.

——, ——, and A. H. Sobel, 2015: Influence of the El Niño/Southern Oscillation on tornado and hail frequency in the United States. *Nat. Geosci.*, **8**, 278–283, https://doi.org/10.1038/ngeo2385.

Anderson, C. J., C. K. Wikle, Q. Zhou, and J. A. Royle, 2007: Population influences on tornado reports in the United States. *Wea. Forecasting*, **22**, 571–579, https://doi.org/10.1175/WAF997.1.

Ashley, W. S., A. M. Haberlie, and J. Strohm, 2019: A climatology of quasi-linear convective systems and their hazards in the United States. *Wea. Forecasting*, **34**, 1605–1631, https://doi.org/10.1175/WAF-D-19-0014.1.

Baggett, C. F., K. M. Nardi, S. J. Childs, S. N. Zito, E. A. Barnes, and E. D. Maloney, 2018: Skillful subseasonal forecasts of weekly tornado and hail activity using the Madden–Julian Oscillation. *J. Geophys. Res. Atmos.*, **123**, 12 661–12 675, https://doi.org/10.1029/2018JD029059.

Barrett, B. S., and V. A. Gensini, 2013: Variability of central United States April–May tornado day likelihood by phase of the Madden–Julian oscillation. *Geophys. Res. Lett.*, **40**, 2790–2795, https://doi.org/10.1002/grl.50522.

Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate*, **5**, 541–560, https://doi.org/10.1175/1520-0442(1992)005<0541:AIOMFF>2.0.CO;2.

Brooks, H. E., C. A. Doswell III, and M. P. Kay, 2003a: Climatological estimates of local daily tornado probability for the United States. *Wea. Forecasting*, **18**, 626–640, https://doi.org/10.1175/1520-0434(2003)018<0626:CEOLDT>2.0.CO;2.

——, J. W. Lee, and J. P. Craven, 2003b: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, **67–68**, 73–94, https://doi.org/10.1016/S0169-8095(03)00045-0.

Carbin, G. W., M. K. Tippett, S. P. Lillo, and H. E. Brooks, 2016: Visualizing long-range severe thunderstorm environment guidance from CFSv2. *Bull. Amer. Meteor. Soc.*, **97**, 1021–1031, https://doi.org/10.1175/BAMS-D-14-00136.1.

Cintineo, J. L., T. M. Smith, V. Lakshmanan, H. E. Brooks, and K. L. Ortega, 2012: An objective high-resolution hail climatology of the contiguous United States. *Wea. Forecasting*, **27**, 1235–1248, https://doi.org/10.1175/WAF-D-11-00151.1.

Edwards, R., J. T. Allen, and G. Carbin, 2018: Reliability and climatological impacts of convective wind estimations. *J. Appl. Meteor. Climatol.*, **57**, 1825–1845, https://doi.org/10.1175/JAMC-D-17-0306.1.

Gensini, V. A., and W. S. Ashley, 2011: Climatology of potentially severe convective environments from North American regional reanalysis. *Electron. J. Severe Storms Meteor.*, **6** (8), https://ejssm.org/ojs/index.php/ejssm/article/view/85/68.

——, and T. L. Mote, 2014: Estimations of hazardous convective weather in the United States using dynamical downing. *J. Climate*, **27**, 6581–6589, https://doi.org/10.1175/JCLI-D-13-00777.1.

——, and A. Marinaro, 2016: Tornado frequency in the United States related to global relative angular momentum. *Mon. Wea. Rev.*, **144**, 801–810, https://doi.org/10.1175/MWR-D-15-0289.1.

——, and M. Tippett, 2019: Global Ensemble Forecast System (GEFS) predictions of days 1–15 U.S. tornado and hail frequencies. *Geophys. Res. Lett.*, **46**, 2922–2930, https://doi.org/10.1029/2018GL081724.

——, D. Gold, J. T. Allen, and B. S. Barrett, 2019: Extended U.S. tornado outbreak during late May 2019: A forecast opportunity. *Geophys. Res. Lett.*, **46**, 10 150–10 158, https://doi.org/10.1029/2019GL084470.

——, B. S. Barrett, J. T. Allen, D. Gold, and P. Sirvatka, 2020: The Extended-Range Tornado Activity Forecast (ERTAF) Project. *Bull. Amer. Meteor. Soc.*, **101** (6), E700–E709, https://doi.org/10.1175/BAMS-D-19-0188.1.

Guan, H., B. Cui, and Y. Zhu, 2015: Improvement of statistical postprocessing using GEFS reforecast information. *Wea. Forecasting*, **30**, 841–854, https://doi.org/10.1175/WAF-D-14-00126.1.

Harnos, D. S., J.-K. E. Schemm, H. Wang, and C. A. Finan, 2019: NMME-based hybrid prediction of Atlantic hurricane season activity. *Climate Dyn.*, **53**, 7267–7285, https://doi.org/10.1007/s00382-017-3891-7.

Hart, J. A., and A. E. Cohen, 2016: The challenge of forecasting significant tornadoes from June to October using convective parameters. *Wea. Forecasting*, **31**, 2075–2084, https://doi.org/10.1175/WAF-D-16-0005.1.

Jung, E., and B. P. Kirtman, 2016: Can we predict seasonal changes in high impact weather in the United States? *Environ. Res. Lett.*, **11**, 074018, https://doi.org/10.1088/1748-9326/11/7/074018.

King, A. T., and A. D. Kennedy, 2019: North American supercell environments in atmospheric reanalyses and RUC-2. *J. Appl. Meteor. Climatol.*, **58**, 71–92, https://doi.org/10.1175/JAMC-D-18-0015.1.

Lee, S.-K., A. T. Wittenberg, D. B. Enfield, S. J. Weaver, C. Wang, and R. Atlas, 2016: U.S. regional tornado outbreaks and their links to spring ENSO phases and North Atlantic SST variability. *Environ. Res. Lett.*, **11**, 044008, https://doi.org/10.1088/1748-9326/11/4/044008.

Lepore, C., M. K. Tippett, and J. T. Allen, 2017: ENSO-based probabilistic forecasts of March–May U.S. tornado and hail activity. *Geophys. Res. Lett.*, **44**, 9093–9101, https://doi.org/10.1002/2017GL074781.

——, ——, and ——, 2018: CFSv2 monthly forecasts of tornado and hail activity. *Wea. Forecasting*, **33**, 1283–1297, https://doi.org/10.1175/WAF-D-18-0054.1.

Mehta, V. M., H. Wang, K. Mendoza, and N. J. Rosenberg, 2014: Predictability and prediction of decadal hydrologic cycles: A case study in Southern Africa. *Wea. Climate Extremes*, **3**, 47–53, https://doi.org/10.1016/j.wace.2014.04.002.

Moore, T. W., and R. W. Dixon, 2011: Climatology of tornadoes associated with Gulf Coast-landfalling hurricanes. *Geogr. Rev.*, **101**, 371–395, https://doi.org/10.1111/j.1931-0846.2011.00102.x.

NOAA, 2019: Billion-dollar weather and climate disasters: Overview. NOAA, accessed 27 September 2019, https://www.ncdc.noaa.gov/billions/.

NWS, 2016: Weeks 3-4 improving mid-range weather outlooks initiative. NOAA/NWS, accessed 16 October 2018, https://www.weather.gov/sti/stimodeling_nggps_weeks3-4.

Pan, Y., N. Zeng, A. Mariotti, H. Wang, A. Kumar, R. L. Sanchez, and B. Jha, 2018: Covariability of Central America/Mexico winter precipitation and tropical sea surface temperatures. *Climate Dyn.*, **50**, 4335–4346, https://doi.org/10.1007/s00382-017-3878-4.

Potvin, C. K., C. Broyles, P. S. Skinner, H. E. Brooks, and E. Rasmussen, 2019: A Bayesian hierarchical modeling framework for correcting reporting bias in the U.S. tornado database. *Wea. Forecasting*, **34**, 15–30, https://doi.org/10.1175/WAF-D-18-0137.1.

Saha, S., and Coauthors, 2010: The NCEP climate forecast system reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1058, https://doi.org/10.1175/2010BAMS3001.1.

Shepherd, M., D. Niyogi, and T. L. Mote, 2009: A seasonal-scale climatological analysis correlating spring tornadic activity with antecedent fall-winter drought in the southeastern United States. *Environ. Res. Lett.*, **4**, 024012, https://doi.org/10.1088/1748-9326/4/2/024012.

Snedecor, G. W., and W. G. Cochran, 1989: *Statistical Methods*. 8th ed. Iowa State University Press, 503 pp.

Thompson, D. B., and P. E. Roundy, 2013: The relationship between the Madden–Julian oscillation and U.S. violent tornado outbreaks in the spring. *Mon. Wea. Rev.*, **141**, 2087–2095, https://doi.org/10.1175/MWR-D-12-00173.1.

Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. M. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243–1261, https://doi.org/10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2.

——, C. M. Mead, and R. Edwards, 2007: Effective storm-relative helicity and bulk shear in supercell thunderstorm environments. *Wea. Forecasting*, **22**, 102–115, https://doi.org/10.1175/WAF969.1.

Ting, M., and H. Wang, 1997: Summertime U.S. precipitation variability and its relation to Pacific sea surface temperature. *J. Climate*, **10**, 1853–1873, https://doi.org/10.1175/1520-0442(1997)010<1853:SUSPVA>2.0.CO;2.

Tippett, M. K., and W. J. Koshak, 2018: A baseline for the predictability of U.S. cloud-to-ground lightning. *Geophys. Res. Lett.*, **45**, 10 719–10 728, https://doi.org/10.1029/2018GL079750.

——, A. H. Sobel, and S. J. Camargo, 2012: Association of U.S. tornado occurrence with monthly environmental parameters. *Geophys. Res. Lett.*, **39**, L02801, https://doi.org/10.1029/2011GL050368.

——, J. T. Allen, V. A. Gensini, and H. E. Brooks, 2015: Climate and hazardous convective weather. *Curr. Climate Change Rep.*, **1**, 60–73, https://doi.org/10.1007/s40641-015-0006-6.

Trapp, R. J., and K. A. Hoogewind, 2018: Exploring a possible connection between U.S. tornado activity and Arctic sea ice. *npj Climate Atmos. Sci.*, **1**, 14, https://doi.org/10.1038/s41612-018-0025-9.

——, D. M. Wheatley, N. T. Atkins, R. W. Przybylinski, and R. Wolf, 2006: Buyer beware: Some words of caution on the use of severe wind reports in postevent assessment and research. *Wea. Forecasting*, **21**, 408–415, https://doi.org/10.1175/WAF925.1.

Verbout, S. M., H. E. Brooks, L. M. Leslie, and D. M. Schultz, 2006: Evolution of the U.S. tornado database: 1954–2003. *Wea. Forecasting*, **21**, 86–93, https://doi.org/10.1175/WAF910.1.

Wang, H., and M. Ting, 2000: Covariabilities of winiter U.S. precipitation and Pacific sea surface temperature. *J. Climate*, **13**, 3711–3719, https://doi.org/10.1175/1520-0442(2000)013<3711:COWUSP>2.0.CO;2.

——, ——, and M. Ji, 1999: Prediction of seasonal mean United States precipitation based on El Niño sea surface temperatures. *Geophys. Res. Lett.*, **26**, 1341–1344, https://doi.org/10.1029/1999GL900230.

——, J.-K. E. Schemm, A. Kumar, W. Wang, L. Long, M. Chellian, G. D. Bell, and P. Peng, 2009: A statistical forecast model for Atlantic seasonal hurricane activity based on the NCEP dynamical seasonal forecast. *J. Climate*, **22**, 4481–4500, https://doi.org/10.1175/2009JCLI2753.1.

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus*, **60A**, 62–79, https://doi.org/10.1111/j.1600-0870.2007.00273.x.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective system. *Mon. Wea. Rev.*, **125**, 527–548, https://doi.org/10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2.

Zhou, X., Y. Zhu, D. Hou, Y. Luo, J. Peng, and R. Wobus, 2017: Performance of the new NCEP global ensemble forecast system in a parallel experiment. *Wea. Forecasting*, **32**, 1989–2004, https://doi.org/10.1175/WAF-D-17-0023.1.