

Geophysical Research Letters

RESEARCH LETTER

10.1029/2020GL087253

Key Points:

- The large-scale weather regimes strongly modulate United States tornado occurrence
- A large fraction of tornado outbreak days are associated with a persisting regime
- Weekly tornado activity can be predicted using forecast model-predicted weather regimes

Supporting Information:

- Supporting Information S1

Correspondence to:

Z. Wang,
zhuowang@illinois.edu

Citation:

Miller, D. E., Wang, Z., Trapp, R. J., & Harnos, D. S. (2020). Hybrid prediction of weekly tornado activity out to Week 3: Utilizing weather regimes. *Geophysical Research Letters*, 47. <https://doi.org/10.1029/2020GL087253>

Received 26 JAN 2020

Accepted 10 APR 2020

Accepted article online 17 APR 2020

Hybrid Prediction of Weekly Tornado Activity Out to Week 3: Utilizing Weather Regimes

Douglas E. Miller¹ , Zhuo Wang¹ , Robert J. Trapp¹ , and Daniel S. Harnos² 

¹Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, ²Climate Prediction Center, NCEP/NWS/NOAA, College Park, MD, USA

Abstract Tornadoes are one of the high-impact weather phenomena that can induce life loss and property damage. Here, we investigate the relationship between large-scale weather regimes and tornado occurrence in boreal spring. Results show that weather regimes strongly modulate the probability of tornado occurrence in the United States due to changes in shear and convective available potential energy and that persisting weather regimes (lasting ≥ 3 days) contribute to greater than 70% of outbreak days (days with ≥ 10 tornadoes). A hybrid model based on the weather regime frequency predicted by a numerical model is developed to predict above/below normal weekly tornado activity and has skill better than climatology out to Week 3. The hybrid model can be applied to real-time forecasting and aid in mitigation of severe weather events.

Plain Language Summary Severe storms that are capable of producing tornadoes, strong winds, and hail may lead to a high number of fatalities and property damage. The relationship between tornadoes and the large-scale, recurrent weather patterns over the United States is investigated. It is shown that specific weather patterns may alter tornado activity and that persisting weather patterns contribute to greater than 70% of tornado outbreak days. Finally, a hybrid model is created to predict tornado activity, and the skill of the model is better than using a long-term average out to Week 3, which is an improvement of the current forecasting state.

1. Introduction

Severe convective storms produce strong winds (>25.9 m/s), large hail (25.4 mm), and/or tornadoes and have the potential to induce significant socioeconomic loss. In 2018 alone, tornadoes in the United States caused over 10 fatalities and greater than \$600 million in combined property and crop damage (National Weather Service (NWS) U.S. n.d., Natural Hazard Statistics). Skillful extended-range forecasts would allow emergency management agencies to better prepare for proactive and reactive measures.

Currently, the Storm Prediction Center (SPC) issues convective outlooks for lead times out to Day 8 (<https://www.spc.noaa.gov/misc/about.html>). Outlooks for Day 1 include probabilistic forecasts for tornadoes, severe wind, and severe hail, while Days 2 and 3 include probabilistic forecasts for the overall severe weather threat. At Days 4 through 8, probabilistic forecasts are made using either a 15% or 30% threshold. A majority of the time, no 15% probabilistic forecast is issued because the potential for severe weather is low or predictability is low. It is common to see the latter listed on the forecast as the weather model's deterministic limit and forecaster's intuition of severe weather decreases after Day 3.

Hitchens and Brooks (2014) evaluated the skill of the Days 1 to 3 categorical convective outlooks and found that the frequency of skillful Day 1 forecasts (defined as added value compared to practically perfect forecast) increased from ~ 0.2 in 1973 to ~ 0.78 in 2011. An increasing trend of skillful forecasts was also evident for Days 2 and 3, with skillful Day 3 forecasts occurring $\sim 50\%$ of the time in 2011. Herman et al. (2018) showed that, in general, the SPC's tornado, hail, and wind forecasts were more skillful where climatological severe weather occurrence is large. Day 1 tornado outlooks were the least skillful compared to hail and wind forecast, with brier skill scores (BSS) hardly exceeding 0.2 with a nationwide average of 0.049. Given the longer lead time, prediction of tornado activity on the subseasonal to seasonal (S2S) time scale is even more challenging.

Most tornado predictions make use of quantifications of the environmental conditions known to strongly control physical processes important for tornado formation (Brooks et al., 1994; Brooks et al., 2003; Grams

et al., 2012; Maddox, 1976; Rasmussen & Blanchard, 1998; Weisman & Klemp, 1982). Among the different environmental parameters, convective available potential energy (CAPE) and vertical wind shear are often used to represent whether the environment is conducive for tornado activity.

Multiple low-frequency modes modulate tornado activity on the S2S time scale via changes of the environmental parameters. Previous studies have shown that tornado activity tends to increase (decrease) in La Niña (El Niño) years (e.g., Allen et al., 2015; Cook & Schaefer, 2008; Lee et al., 2013). A robust relationship between the Madden-Julian Oscillation (MJO) and tornado activity has also been identified (Tippett, 2018), with increased tornado activity during Phase 2 of the MJO when convection is active over the Indian Ocean (Thompson & Roundy, 2013). Additionally, tornadoes are more likely to occur in Phases 1–4 of the global wind oscillation (GWO) when atmospheric angular momentum is low (Gensini & Marinaro, 2016; Moore, 2018). Statistical models have been developed based on these climate modes to predict severe storm activity on the S2S time scale (Baggett et al., 2018; Gensini et al., 2019; Lepore et al., 2017). However, the ENSO, MJO, and GWO explain limited variability of tornado activity. Enhanced tornado activity can still occur when these low-frequency climate modes suggest an overall inactive time period (Moore et al., 2018). It is likely that synoptic-scale events strongly modulate the environmental conditions on the shorter time scales and induce tornado outbreaks even when the climate modes suggest otherwise (Moore et al., 2018). If predictability exists for the statistics of such synoptic-scale events, effectively exploiting this source of predictability can improve S2S prediction of severe storm activity.

To investigate the synoptic-scale events that may contribute to enhanced tornado activity, we turn our attention to weather regimes (WRs). The concept of WRs was introduced decades ago (Rex, 1950) and are defined as recurrent atmospheric patterns (Michelangeli et al., 1995). The underlying assumption is that the large-scale atmospheric circulation can be represented by a finite number of states, an assumption supported by theoretical work on the existence of multi-equilibria of the climate system (Charney & Devore, 1979). It is generally believed that the spatial patterns of WRs are determined by the internal dynamics of the atmosphere, while low-frequency climate modes, boundary forcing (e.g., SST), and external forcing (e.g., anthropogenic forcing) may modulate the frequency of occurrence of WRs (Michelangeli et al., 1995; Molteni & Palmer, 1993). Since a WR may last for weeks, its persistence may serve as a source of predictability on the S2S time scale, especially for severe weather, as persisting large-scale systems appear to contribute to multiday tornado events (Trapp, 2014).

This study aims to explore skillful tornado prediction on the weekly time scale through 4 weeks in the future by employing the concept of WRs. Data and methodology are described in section 2. We investigate how large-scale weather regimes modulate severe weather environments and tornado activity in section 3 and present a skillful hybrid prediction scheme for weekly tornado activity in section 4, followed by the discussion and conclusion in section 5.

2. Data and Methodology

2.1. Data

This study uses tornado reports taken from the Storm Prediction Center website (<https://www.spc.noaa.gov/wcm/#data>). Each report includes the Enhanced Fujita (EF)-scale rating (a damaged-based proxy for intensity), starting geographical location, and date of the tornado. Here, we focus on reports in 1990–2019. This period represents a compromise between data set length and an allowance for a significant fraction of the reports to have occurred during the Next-Generation Radar era and thus have undergone some quality control (Smith et al., 2012). Tornado activity is quantified in terms of tornado days (defined as a day with ≥ 1 tornado report, or stated otherwise) as well as total numbers of tornado reports within five regions: Southern Plains (SP), Northern Plains (NP), Midwest (MW), Southeast (SE), and Northeast (NE) (see Figure S1 in the supporting information). For composite analysis, reports are interpolated to the nearest grid point on a 1° grid and then averaged within a $5^\circ \times 5^\circ$ box. For the purpose of this study, only reports of $\geq \text{EF1}$ are included. May, the peak season for tornado activity, was chosen for the analysis herein (NOAA NCEI, n.d.), but the skillful prediction of tornado activity can also be achieved in other months (Figure S7 and section 4).

The 500-hPa geopotential heights (H500) during 1990–2019 from the ERA-Interim reanalysis (ERA-Interim; Dee et al., 2011) data set are used to construct weather regimes over the United States. Data are available on a T255L60 (~0.7° in horizontal) grid but are interpolated to a 1° latitude-longitude grid. Additional ERA-Interim variables analyzed are CAPE, 500-hPa winds, 900-hPa winds, and 10-m winds. Low-level shear (deep layer) is calculated as the vector difference between the 900-hPa (500-hPa) and 10-m winds. These two shear parameters, herein labeled S900 and S500, roughly correspond to 0- to 1-km and 0- to 6-km bulk shear, respectively, which have been shown to discriminate nontornadoic environments from environments of significant tornadoes (Thompson et al., 2012). All anomalies are created by removing the daily linear trend from 1990–2019 and removing the long-term daily mean on each calendar day.

To create the hybrid prediction model, the reforecasts from the European Centre for Medium-Range Weather Forecast (ECMWF) S2S model are used. The reforecasts are produced twice per week for 20 years (1998–2017), containing one control member and 10 perturbation runs (Vitart, 2017). H500 out to 28 days is used to predict the large-scale weather regimes. The original resolution is $0.25^\circ \times 0.25^\circ$ from Days 0–10 and $0.5^\circ \times 0.5^\circ$ after Day 10, but all data are interpolated to a $1.0^\circ \times 1.0^\circ$ grid to be consistent with the reanalysis.

2.2. K-Means Clustering

K-means clustering using Lloyd's algorithm (Lloyd, 1982) is performed over North America (24–55°N, 130–60°W) and utilizes daily H500 during May from 1990–2019. A latitude weight ($\cos(\phi)$) is applied for the clustering analysis. The optimal number of WRs, determined by the elbow method (Kodinariya & Makwana, 2013), is five. The same algorithm is applied to each ECMWF ensemble output to identify predicted WRs.

2.3. Hybrid Model

We create a hybrid model to predict the weekly tornado activity, which is defined as the number of tornado days over a 7-day period. First, we perform K-means clustering on each ensemble member. The weekly WR frequency is calculated for each regime from Days 1–7 to Days 21–27 from individual ensemble members. Next, the weekly tornado activity is predicted using the ensemble member's WR frequency and the observed tornado frequency distributions of the training data set. Tornado frequency distributions are functions of the weekly WR frequency (Figure S2). For example, if the model predicts Week 1 to contain 5 WR2 days and 2 WR4 days, the model predicts ~2.1 (1.8 + 0.3) tornado days below normal activity. The tornado predictions are carried out for each ensemble member, and the ensemble mean is used for the final tornado activity prediction. The prediction is evaluated against the observed weekly tornado day anomalies.

2.4. Model Evaluation and Statistical Significance

A leave-one-out method is employed to assess the prediction skill, such that the year evaluated is independent of the training data set. The Anomaly Correlation Coefficient (ACC; Wilks, 2011) is calculated between the observed and model predicted weekly WR frequency to demonstrate the model's ability to predict the correct WRs. To evaluate the hybrid model, the Heidke Skill Score (HSS) for 2-tier prediction is calculated. The two tiers, above or below average, are defined with respect to the mean of the distribution of weekly tornado activity, and the results do not change quantitatively by using the median. The HSS is calculated by $HSS = 100 * (H - E)(T - E)^{-1}$, where H is the total number of correct forecast, E is the number of correct forecast by random chance (50% for two-tier predictions), and T is the total number of forecasts.

Composite anomalies of CAPE, S900, and S500 are constructed to examine the environmental condition contributing to the changes in tornado activity. We employ a two-tailed Student's t test, with the null hypothesis that the anomalies do not differ from zero, for composite anomalies of CAPE and S900/S500. A Mann-Whitney U test is used for composite anomalies of tornado frequency. Results are significant if the p value is less than 0.05.

3. Relationship Between Severe Environments and Large-Scale Weather Regimes

The five WRs identified using K-means clustering are displayed in Figure 1. WR1 occurs most frequently and is characterized by a zonally elongated low and a high center over the East Coast of the United States. WR2 and WR3 are characterized by a three-cell wavelike pattern spanning from the East Pacific to the West

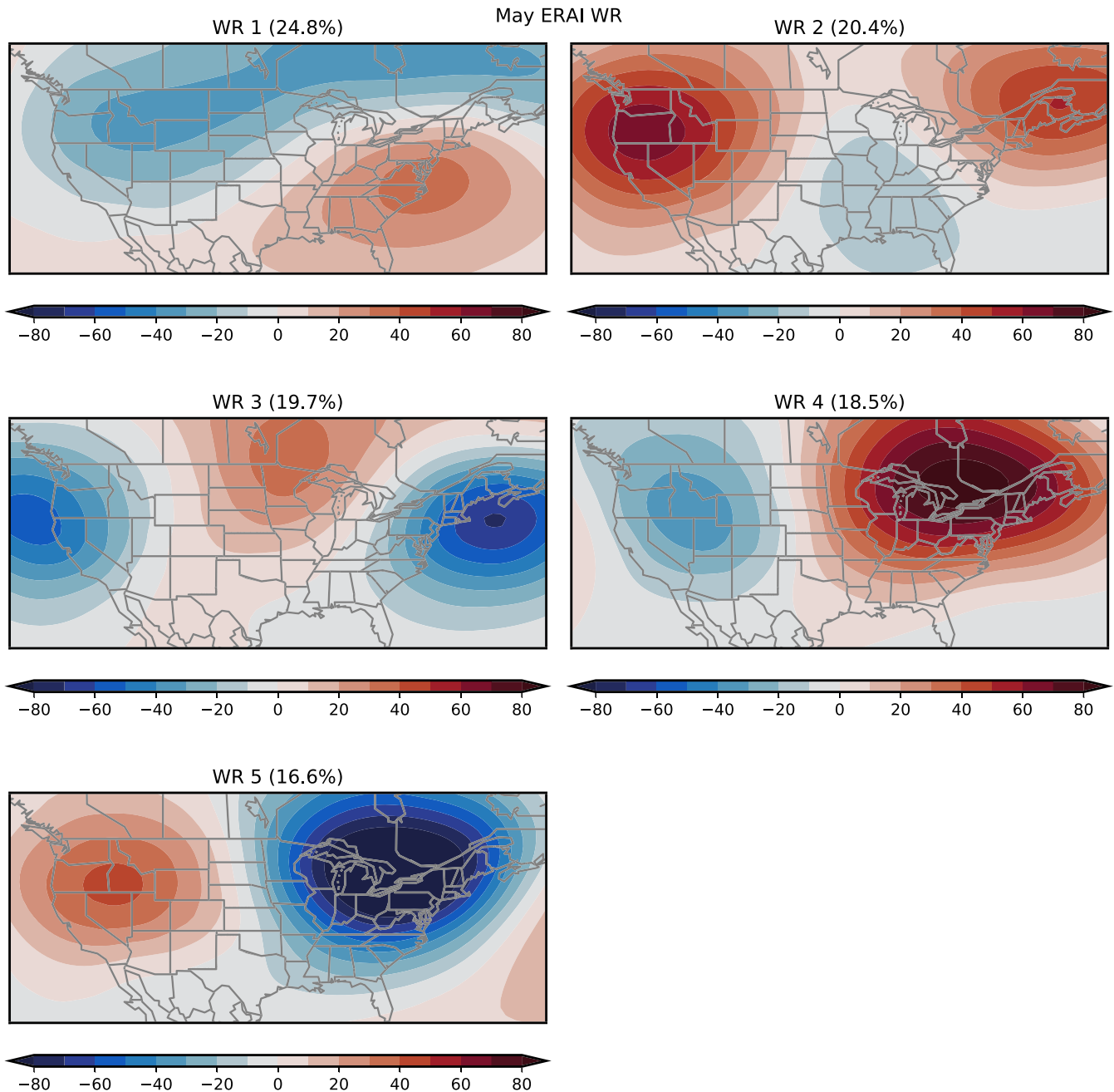


Figure 1. The H500 patterns for five WRs over North America during May ordered from most frequent (WR1) to least frequent (WR5). The frequencies of occurrence are listed at the top of each subfigure.

Atlantic, and WR4 and WR5 represent a dipole pattern of opposite polarities over North America, which modulates the meridional moisture transport from the Gulf of Mexico.

Tornado outbreaks are found to be associated with persistent occurrence of WRs (Figure 2a), especially WR1 and WR4. Here we defined a persistent WR as a WR lasting at least 3 days. For example, the tornado outbreaks during 17–29 May 2019, which were successfully predicted on the S2S time scale (Gensini et al., 2019), are associated with a persistent WR1. Additionally, the tornado outbreak in 5–10 May 2015 is associated with a persistent WR4. The frequency of occurrence for each WR changes annually and is potentially driven by the large-scale climate modes, such as the MJO and ENSO (Vigaud et al., 2018). About 75% of tornado days (≥ 1 tornado) are associated with a persistent WR (Figure 2b) while 26% of tornado days occur during a

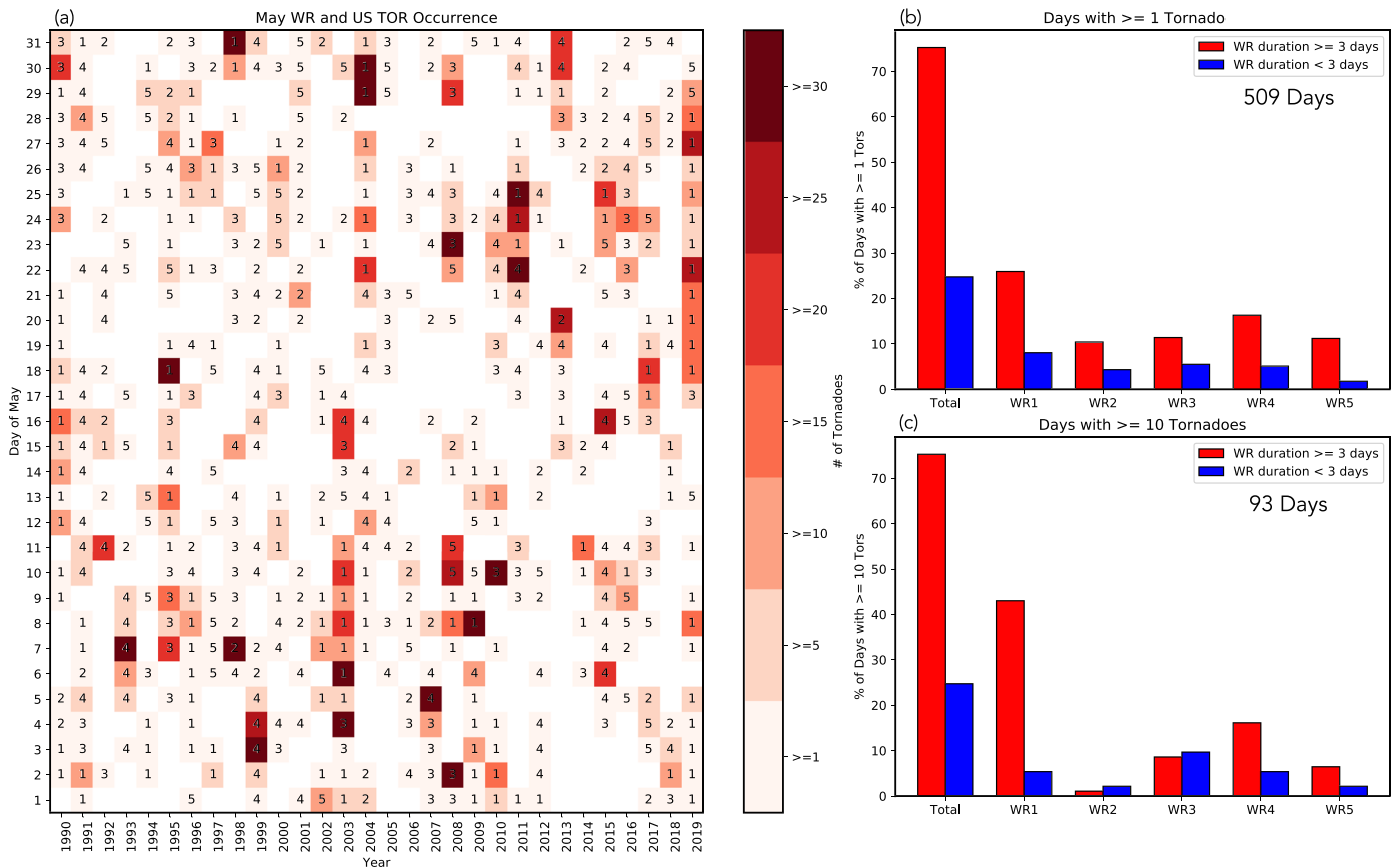


Figure 2. (a) U.S. tornado days (color shading indicates number of tornadoes per day) and WRs (numbers) during May 1990–2019. Non-tornado days are masked with white. Bar charts display the percentage of days with (b) ≥ 1 tornado and (c) ≥ 10 tornadoes for persisting (red) and nonpersisting (blue) WRs.

persistent WR1. Similarly, 75% of all outbreak days (≥ 10 tornadoes) occur during a persistent WR while 43% of all outbreak days occur during a persistent WR1 (Figure 2c). For context, the frequency of occurrence of persistent WR1 days is 18.9% (176 days). Of these 176 days, 76% (132 days) contain ≥ 1 tornado, while 22% of the 176 days (40 days) are considered outbreak days.

Figure 3 displays the probability of tornado occurrence for each regime (regardless of duration) relative to climatology for different regions. Table S1 lists the climatological probability of occurrence for different tornado day thresholds. The probability of occurrence of ≥ 1 tornado (≥ 5 tornadoes) on a WR1 day for the US is $>75\%$ ($>65\%$), or $\sim 20\%$ (13%) above climatology (Figure 3a). Variability occurs for the regions as well. The probability of occurrence of ≥ 1 tornado on a WR1 day is $\sim 40\%$ (or $>11\%$ above climatology) for the SP (Figure 3b), $>35\%$ (or $\sim 16\%$ above climatology) for the MW (Figure 3c), and $\sim 30\%$ (or $\sim 13\%$ above climatology) for the SE (Figure 3d). In summary, tornado day occurrence is enhanced during WR1 in all regions except the NP, reduced during WR2 in all regions except the NE, enhanced during WR3 in the SP and NP, enhanced in all regions during WR4 except the NE and SE, and reduced during WR5 in each region (negligible change over SE).

The modulation of WRs on the regional tornado activity can be explained by the changes in S900 and CAPE. Figures 4a–4e display the composite anomalies of tornado frequency, CAPE, and significant S900 vectors. The climatological values are shown in Figure 4f for reference. It is worth noting that Figure 3 examines tornado days, whereas Figure 4 examines the tornado frequency. A region may experience a decrease in tornado days but has above-normal tornado frequency due to a large number of reports on a given day.

The significant increase in tornado frequency across the MW during WR1 days (Figure 4a) is associated with a significant increase in CAPE stretching north through Nebraska, Iowa, and Illinois. A significant increase

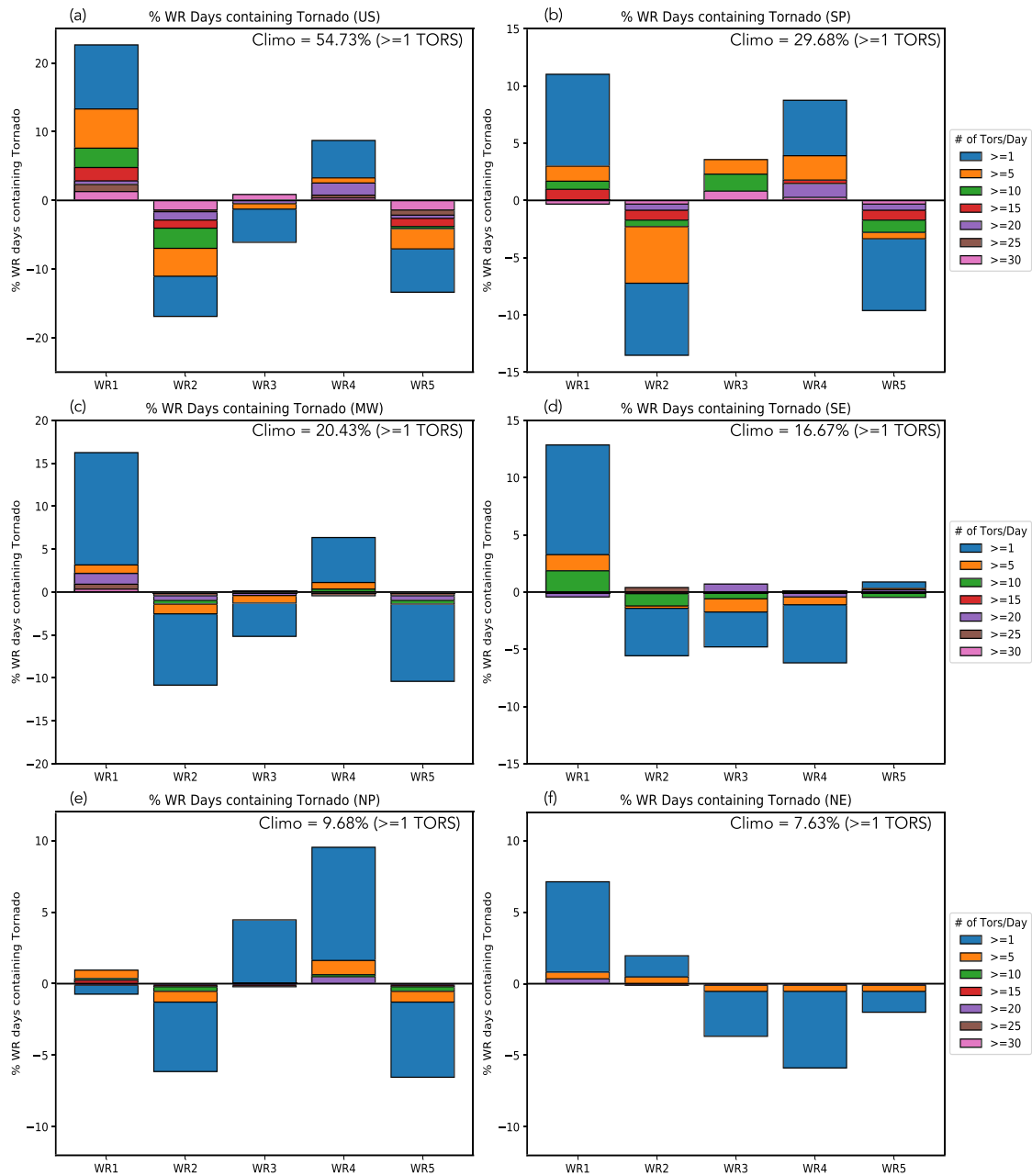


Figure 3. Percentage of WR days containing a tornado compared to climatology (percentage of total tornado days) for the (a) US, (b) SP, (c) MW, (d) SE, (e) NP, and (f) NE. Climatological percentages (≥ 1 tornado on a day) are listed below title. Different color bars indicate number of tornadoes required to count as a tornado day.

in S900 also contributes to a favorable environment for tornado occurrence (Thompson et al., 2012). Significant CAPE and S900 anomalies are evident during WR4 days contributing to the increase in tornado activity across the Plains (Figure 4d). An increase in moisture (dewpoint temperature) is also evident in collocations of CAPE during WR1 and WR4 (not shown). Consistent with Figure 3, Figure 4c shows enhanced tornado activity over the SP and NP due to an increase in CAPE and S900, while reduced (although nonsignificant) over the MW, SE, and NE during WR3 days. WR2 and WR5 are characterized by reduced tornado activity over Oklahoma and Kansas (Figures 4b and 4e), which is likely due to the significant decrease in CAPE and S900.

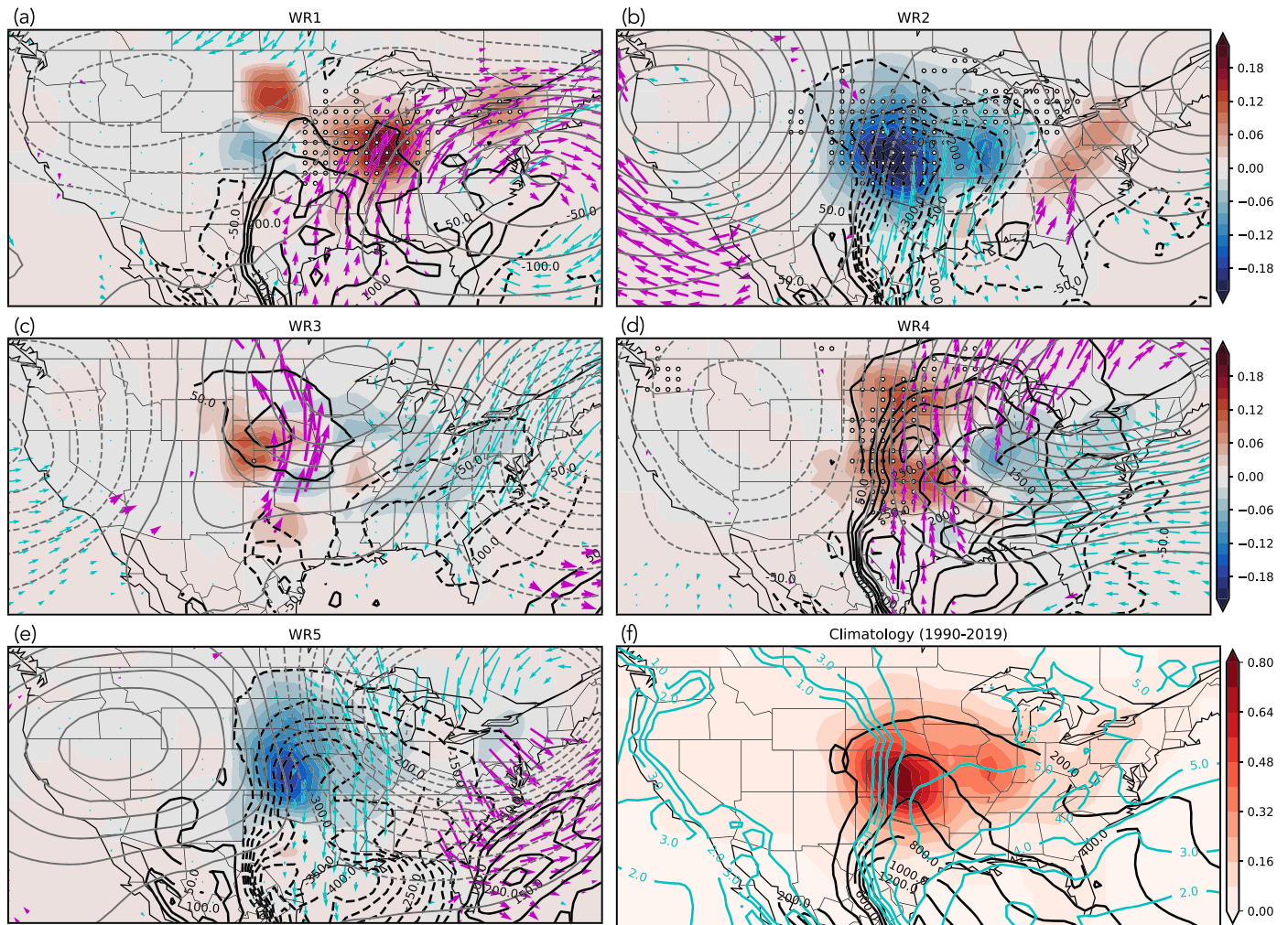


Figure 4. Composite anomalies of tornado frequency (color shading, units: tornadoes/day), H500 (gray contours: -100 to 100 gpm every 10), CAPE (black contours; units: J/kg), and S900 (vectors) for (a) WR1, (b) WR2, (c) WR3, (d) WR4, and (e) WR5. Climatology is shown (f) for reference with S900 magnitude (teal contours, units: m/s) in place of vectors. Dashed contours represent negative values. Only significant CAPE anomalies ($\alpha = 0.05$) are displayed in (a)–(e) while hatching indicates significant tornado frequency anomalies. Teal (magenta) vectors are where S900 magnitude is significantly lower (higher) than climatology.

In addition to S900, we also examined S500 (Figure S3). Moderate-to-large deep-layered shear is a necessary but not sufficient condition for tornado occurrence (Thompson et al., 2012). The spatial distribution of S500 anomalies is largely consistent with S900 and tornado activity anomalies (Figure 4).

The composites shown in Figure 4 are further separated into persisting (lasting ≥ 3 days) and nonpersisting WRs (Figure S4). The contrast is large, especially for WR1. No significant positive anomalies of tornado frequency are evident during nonpersisting WR1 days, as strong CAPE anomalies are confined to Texas, with significant S900 anomalies east of the CAPE maxima. The persistent events contain significant enhancement of tornado frequency associated with CAPE anomalies stretching as far northeast as New York, and an increase of S900 across the eastern SP, SE, and MW. A persisting regime would therefore support multiple days of tornado activity (Trapp, 2014) and may serve as a source of predictability for longer-range forecasts.

4. Hybrid Prediction of Severe Storm Activity

Next, we attempt to predict weekly tornado activity. The hybrid prediction scheme, where the ECMWF model's predicted WR frequencies are used to predict the weekly tornado activity, hinges on the model's ability to

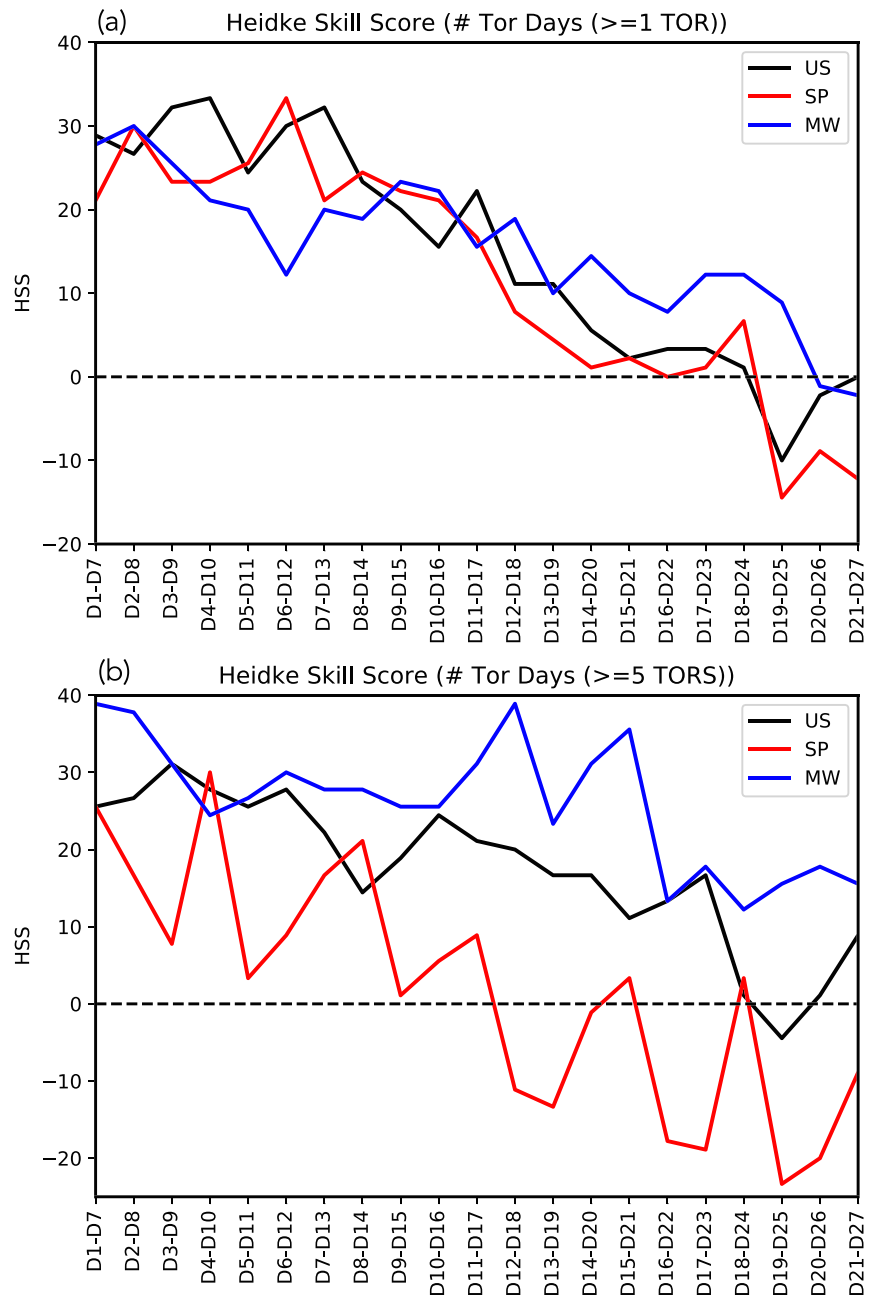


Figure 5. HSSs (units: %) for the hybrid prediction model predicting number of days per week with (a) ≥ 1 tornado and (b) ≥ 5 tornadoes for the United States (black), Southern Plains (red), and Midwest (blue).

represent and predict the weather regimes. After performing K-means clustering on the model data, the five regimes found in the reanalysis all emerge (not shown). The ACC between the observed and predicted ensemble mean weekly WR frequencies is examined (Figure S5). The mean ACC for the five WRs remains above 0.6 (arbitrary predictability limit) out to Days 7–13; WR4 has the largest skill at shorter lead times, and the ACC remains above 0.6 out to Days 8–14, while the ACC of WR1 drops below 0.6 beyond Days 6–12.

Figure 5a shows the HSS of the hybrid model in predicting number of tornado days per week. Only the US, SP, and MW are shown for brevity. The HSSs are between 20% and 30% for shorter lead times and remain

positive well into Week 3, especially for the US and MW, implying that the empirical model adds knowledge relative to climatology. The same model can be used to predict the number of days with ≥ 5 tornadoes (Figure 5b). The predictions are better than climatology well into Weeks 3 and 4, with the MW showcasing the largest skill. However, the HSSs have large fluctuations due to the smaller sample size of the days with ≥ 5 tornadoes.

A multiple linear regression (MLR) model was also tested using the predicted ensemble mean weekly WR frequencies as predictors. The HSSs also persist into Week 3 (Figure S6), but the skills are overall lower than those in Figure 5. This is probably because the MLR model does not consider the nonlinear impacts of persistent WRs on tornado activity. In other words, the tornado activity associated with five consecutive WR1 days is more than five times the tornado activity anomalies associated with one isolated WR1 day (Figure S2).

Although only the month of May is shown in sections 3 and 4 for brevity, the strong modulation of tornado activity by WRs is also found in March and April, and the hybrid approach produces skillful predictions in these months as well (Figure S7).

5. Summary and Conclusions

The relationship between large-scale WRs and tornado frequency is investigated during May (1990–2019). Results show strong variability of regional tornado occurrence during different WRs. The probability of occurrence of a U.S. tornado day during WR1 is about 75%, and the enhanced tornado activity during WR1, WR3, and WR4 can be explained by increases in CAPE and low-level shear (S900), while decreases in CAPE and S900 explain weakened activity during WR2 and WR5. Persisting weather regimes are responsible for a large fraction of outbreak days (70 of 93 days). Interestingly, a persistent WR1 produces a large increase in tornado activity across the southeast NP, northeast SP, and MW, while a nonpersisting WR1 produces an overall decrease in tornado occurrence. Model predicted weather regimes are then used to develop a hybrid model to predict weekly tornado activity. Skillful predictions are evident out to Week 3 for the United States and the subregions analyzed. Our study suggests that employing persistent weather regimes, along with the forecast of opportunity associated with low-frequency climate modes (Baggett et al., 2018; Gensini et al., 2019), has the potential to improve the extended-range prediction of severe storm activity.

Acknowledgments

This work is supported by the National Oceanic and Atmospheric Administration (NOAA) Grants NA15NWS4680007 and NA16OAR4310080. We acknowledge the NCAR Computational and Information Systems Laboratory (CISL) for providing computing resources. All ERAI data were provided and are available through the Research data Archive (RDA) (<https://doi.org/10.5065/D6CR5RD9>). Tornado report data are available at the Storm Prediction Center website (<https://www.spc.noaa.gov/wcm/#data>). The ECMWF S2S model data are available online (at <https://apps.ecmwf.int/datasets/>). We thank two anonymous reviewers for their helpful comments on this manuscript.

References

- Allen, J. T., Tippett, M. K., & Sobel, A. H. (2015). Influence of the El Niño/Southern Oscillation on tornado and hail frequency in the United States. *Nature Geoscience*, *8*, 278–283.
- Baggett, C. F., Nardi, K. M., Childs, S. J., Zito, S. N., Barnes, E. A., & Maloney, E. D. (2018). Skillful subseasonal forecasts of weekly tornado and hail activity using the Madden-Julian Oscillation. *Journal of Geophysical Research: Atmospheres*, *123*, 12,661–12,675. <https://doi.org/10.1029/2018D029059>
- Brooks, H. E., Doswell, C. A. III, & Cooper, J. (1994). On the environments of tornadic and nontornadic mesocyclones. *Weather and forecasting*, *9*, 606–618.
- Brooks, H. E., Lee, J. W., & Craven, J. P. (2003). The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmospheric Research*, *67–68*, 73–94.
- Charney, J. G., & Devore, J. G. (1979). Multiple flow equilibria in the atmosphere and blocking. *Journal of the Atmospheric Sciences*, *36*, 1205–1216.
- Cook, A. R., & Schaefer, J. T. (2008). The relation of El Niño-Southern Oscillation (ENSO) to winter tornado outbreaks. *Monthly Weather Review*, *136*, 3121–3137.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. <https://doi.org/10.1002/qj.828>
- Gensini, V. A., Gold, D., Allen, J. T., & Barrett, B. S. (2019). Extended U.S. tornado outbreak during late May 2019: A forecast of opportunity. *Geophysical Research Letters*, *46*, 10,150–10,158.
- Gensini, V. A., & Marinaro, A. (2016). Tornado frequency in the United States related to global relative angular momentum. *Monthly Weather Review*, *144*, 801–810.
- Grams, J. S., Thompson, R. L., Snively, D. V., Prentice, J. A., Hodges, G. M., & Reames, L. J. (2012). A climatology and comparison of parameters for significant tornado events in the United States. *Weather and Forecasting*, *27*(1), 106–123.
- Herman, G. R., Nielsen, E. R., & Schumacher, R. S. (2018). Probabilistic verification of Storm Prediction Center convective outlooks. *Weather and Forecasting*, *33*(1), 161–184.
- Hitchens, N. M., & Brooks, H. E. (2014). Evaluation of the Storm Prediction Center's convective outlooks from Day 3 through Day 1. *Weather Forecasting*, *29*, 1134–1142.
- Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of cluster in K-means clustering. *IJARCSMS*, *1*(6), 2321–7782.
- Lee, S.-K., Atlas, R., Enfield, D., Wang, C., & Liu, H. (2013). Is there an optimal ENSO pattern that enhances large-scale atmospheric processes conducive to tornado outbreaks in the United States. *Journal of Climate*, *26*, 1626–1642.

- Lepore, C., Tippett, M. K., & Allen, J. T. (2017). ENSO-based probabilistic forecasts of March–May U.S. tornado and hail activity. *Geophysical Research Letters*, *44*, 9093–9101.
- Lloyd, S. P. (1982). Least squares quantization in PCM. *Inf. Theory IEEE Trans*, *28*, 129–137.
- Maddox, R. A. (1976). An evaluation of tornado proximity wind and stability data. *Monthly Weather Review*, *104*, 133–142.
- Michelangeli, P.-A., Vautard, R., & Legras, B. (1995). Weather regimes: Recurrence and quasi stationarity. *Journal of the Atmospheric Sciences*, *52*(8), 1237–1256.
- Molteni, F., & Palmer, T. N. (1993). Predictability and finite-time instability of the northern winter circulation. *Quarterly Journal of the Royal Meteorological Society*, *119*, 269–298.
- Moore, T. W. (2018). Annual and seasonal tornado activity in the United States and the global wind oscillation. *Climate Dynamics*, *50*, 4323–4334.
- Moore, T. W., St. Clair, J. M., DeBoer, T. A. (2018) An analysis of anomalous winter and spring tornado frequency by phase of the El Niño/Southern Oscillation, the Global Wind Oscillation, and the Madden–Julian Oscillation. *Advances in Meteorology*.
- National Oceanic and Atmospheric Administration (NOAA) (n.d.) National Centers for Environmental Information (NCEI) Historical Records and Trends. Retrieved from <https://www.ncdc.noaa.gov/climate-information/extreme-events/us-tornado-climatology/trends>
- National Weather Service (NWS) U.S. (n.d.) Natural hazard statistics. Retrieved from <https://www.weather.gov/media/hazstat/sum18.pdf>
- Rasmussen, E. N., & Blanchard, D. O. (1998). A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Weather and Forecasting*, *13*(4), 1148–1164.
- Rex, D. (1950). Blocking action in the middle troposphere and its effect on regional climate. *I: An aerological study of blocking*. *Tellus*, *2*, 169–211.
- Smith, B. T., Thompson, R. L., Grams, J. S., Broyles, C., & Brooks, H. E. (2012). Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm Classification and Climatology. *Weather and Forecasting*, *27*, 1114–1135.
- Thompson, D. B., & Roundy, P. E. (2013). The relationship between the Madden-Julian Oscillation and US violent tornado outbreaks in the spring. *Monthly Weather Review*, *141*(6), 2087–2095.
- Thompson, R. L., Smith, B. T., Grams, J. S., Dean, A. R., & Broyles, C. (2012). Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Weather and Forecasting*, *27*, 1136–1154.
- Tippett, M. K. (2018). Robustness of relations between the MJO and U.S. tornado occurrence. *Monthly Weather Review*, *146*(11), 3873–3884.
- Trapp, R. J. (2014). On the significance of multiple consecutive days of tornado activity. *Monthly Weather Review*, *142*(4), 1452–1459.
- Vigaud, N., Robertson, A. W., & Tippett, M. K. (2018). Predictability of recurrent weather regimes over North America during winter from submonthly reforecasts. *Monthly Weather Review*, *146*(8), 2559–2577.
- Vitart, F., and Coauthors, 2017: The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, *98*, 163–173.
- Weisman, M. L., & Klemp, J. B. (1982). The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Monthly Weather Review*, *110*, 504–520.
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*. 3rd ed. Oxford; Waltham, MA: Academic Press.