Linking Weather Regimes to the Variability of Warm-Season Tornado Activity over the United States

Authors: Matthew Graber[1], Zhuo Wang[1], & Robert J. Trapp[1]

[1]Department of Climate, Meteorology, & Atmospheric Sciences, University of Illinois Urbana-Champaign, Urbana, 61820, United States

Correspondence to: Zhuo Wang (zhuowang@illinois.edu)

Abstract. The contiguous United States (CONUS) experiences considerable interannual variability in tornado activity. The high impacts of tornadoes motivate the need to better understand the link between seasonal tornado activity and large-scale atmospheric circulation, which may contribute to better seasonal prediction. We employed K-means clustering analysis of 500 hPa geopotential height (500H) daily anomalies from the ERA-5 reanalysis and identified five warm-season weather regimes (WRs). Certain WRs are shown to strongly affect tornado activity, especially outbreaks, due to their relationship with environmental parameters including convective available potential energy (CAPE) and vertical wind shear (VWS). In particular, WR-B, which is characterized by a three-cell wave-like pattern with an anomalous low over the central-CONUS, is associated with enhanced CAPE and VWS in tornado-prone regions and represents a tornado-favorable environment. Persistent WRs, those lasting for ≥5 consecutive days, are associated with 76% of all tornado outbreaks (days with >10 EF-1+ tornadoes) since 1960; persistent WR-B, in particular, accounts for about 30% of all tornado outbreaks. The impacts of WR persistence on tornado activity anomalies, however, are found to be asymmetric: compared to non-persistent WRs, persistent WRs amplify positive tornado activity anomalies but may not further enhance negative tornado activity anomalies. An empirical model using WR frequency and persistence captures the year-to-year variability of warm-season tornado days and outbreaks reasonably well, including some years with high-impact outbreaks. Our study highlights the potential application of WRs for better seasonal prediction of tornado activity.

1 Introduction

The contiguous United States (CONUS) experiences more tornadoes than anywhere else in the world, leading to significant economic and life losses (NCEI, 2024). Tornado outbreaks (TOs) are a primary contributor to these impacts, and the annual frequency has increased by 2.5 events over the past 63 years (Brooks et al. 2014; Graber et al. 2024), particularly over the Southeast U.S. (Gensini and Brooks, 2018; Graber et al., 2024; Moore, 2018; Moore and DeBoer, 2019). In contrast, tornado days (TDs) have decreased in frequency at a rate of ~10 per decade since 1960 (Brooks et al. 2014; Graber et al. 2024), especially from March to September, and over the southern Great Plains (Gensini and Brooks, 2018; Graber et al., 2024; Moore, 2018; Moore and DeBoer, 2019). Embedded within these trends is large interannual variability, as evidenced by the percent change, with respect to the previous year, in annual CONUS tornado reports over the recent five years (2019-2023): +34.7%, -28.7%, +21.4%, -13.0%, and +24.5%, as well as by the corresponding percent change in tornado fatalities: +320.0%, +80.9%, +36.8%, -77.9%, and +260.9% (Storm Prediction Center, 2024). Such variability affects the situational awareness and

vulnerability of the populations, especially those that are disadvantaged. It also complicates decision making and resource management by key stakeholders across multiple sectors. In addition, exposure to future tornadoes is increasing with growing urban areas (Ashley and Strader, 2016; Strader et al., 2017, 2024). These and other impacts motivate the efforts to better understand the variability of tornado activity over the seasonal and longer time scale, which would ultimately contribute to improved prediction of tornado activity.

Some variability of tornado activity can be attributed to low-frequency climate modes (Miller et al., 2022; Niloufar et al., 2021; Thompson and Roundy, 1998; Vigaud et al., 2018). For example, Cook and Schaefer, (2008) examined winter tornado outbreaks in relation to the phase of the El Niño – Southern Oscillation (ENSO) and found that a La Niña phase favored tornadoes in the Southeast and a neutral phase favored tornadoes in the Great Plains. Allen et al., (2015) further found that La Niña (El Niño) years typically coincide with more (fewer) tornadoes in the spring across the central CONUS, and that the winter ENSO phase can be used to predict tornado frequency during the spring. Additionally, a positive (negative) phase of the Arctic Oscillation (AO) combined with a La Niña (El Niño) phase may increase (decrease) tornado activity (Tippett et al., 2022). Tornado activity can also be modulated by anthropogenic climate change (ACC), either indirectly via changes in climate modes or/and directly via changes in relevant atmospheric conditions. For example, increasing greenhouse gas concentrations are projected to lead to a moister atmosphere, especially in the lower troposphere, which in turn contributes to higher convective available potential energy (CAPE) (Trapp et al., 2007). Diffenbaugh et al. (2013) showed that climate models project a robust increase in the number of days with high CAPE coinciding with high vertical wind shear (VWS) in the eastern U.S., two important parameters for tornado-favorable environments (Brooks et al., 2003; Mercer and Bates, 2019; Rasmussen and Blanchard, 1998; Thompson et al., 2012).

Variability of synoptic-scale circulations provides another means of explaining tornado activity. Conducive synoptic-scale circulation anomalies for TOs in the United States show a trough-ridge pattern over the central- to eastern-CONUS, while non-TOs usually feature more zonal flow and weaker relative vorticity (Mercer et al., 2012). In particular, Cwik et al. (2022) performed rotated EOF analysis of 500-hPa geopotential height associated with historic May TOs and identified three circulation patterns. The three circulation patterns are all characterized by a trough feature over the central to eastern U.S. While their study concludes that the synoptic patterns associated with TOs remain the same over time, there is partial variability in the locations of TOs on multidecadal scales. Additionally, mesoscale processes without strong links to synoptic-scale circulations also affect tornadoes, especially weak or isolated tornado events (e.g., tornadogenesis in non-supercellular storm modes associated with mesoscale boundaries; see Wakimoto and Wilson, 1989).

In this study, we will investigate the link between the synoptic-scale circulation and tornado activity using the concept of weather regimes (WRs), identified using the K-means clustering analysis. Previous studies suggest that WRs represent a finite number of equilibrium states of the

climate system (Charney and DeVore, 1979; Hannachi et al., 2017; Michelangeli et al., 1995). Their spatial patterns are determined by the internal dynamics of the atmosphere, while their frequencies and persistence may be modulated by climate modes or external forcings such as ACC (Corti et al., 1999). The WR framework thus has a strong dynamic basis and have been used to reliably detect changes in regional temperature and precipitation (Robertson and Ghil, 1999). Additionally, unlike Cwik et al. (2022)'s study, which focuses on circulation patterns conditioned on major TOs, our identification of WRs is independent of TOs. This approach allows us to examine WRs that facilitate or hinder tornado activity, providing more comprehensive information for potential forecasting applications. Furthermore, we will examine environmental conditions relevant to tornado development, such as CAPE and VWS (Brooks et al., 2003; Mercer and Bates, 2019; Rasmussen and Blanchard, 1998; Thompson et al., 2012), which will help us better understand the link between WRs and tornado activity.

WRs were used previously to investigate sub-seasonal variability of tornado activity, and a skillful WR-based, hybrid model was developed for the sub-seasonal prediction of tornado activity in the month of May (Miller et al. 2020). Lee et al., (2023) applied the year-round WR method (Grams et al., 2017) over North America and defined four year-round WRs. (Tippett et al., (2024) identified statistically significant relationships between these year-round WRs and tornado activity in all months except June through August, but Tippett et al. (2024) made no consideration of WR persistency. This helps motivate our focus herein on the warm-season tornado activity and its interannual variability. We will test the hypothesis that the interannual variability of warm-season tornado activity is modulated by WR frequency and persistence. A better understanding of the possible links between WRs and tornado activity may contribute to improved seasonal prediction of tornado activity.

2 Data and Methodology

2.1 ERA-5 reanalysis data

Using the ERA-5 reanalysis dataset (Hersbach et al., 2020), 500 hPa geopotential heights (500H) and parameters including most unstable CAPE (MUCAPE), 10m and 500 hPa winds, and convective precipitation (CP) were analyzed over the CONUS $[24 - 55^o$ N, $130 - 60^o$ W] using the native $0.25^o \times 0.25^o$ (latitude $\times$ longitude) resolution. Since the tornadic storms leading to TOs tend to initiate between 18 and 00 UTC after sufficient atmospheric destabilization from surface heating (Cwik et al., 2022), 500H at 21 UTC was used to represent the daily circulation patterns. Daily maximum values of MUCAPE and CP at each grid point were used to represent the daily peak instability and in turn amplify the signal for days with the potential for significant convective storms. The 0-6 km bulk wind shear (S06), or deep-layer shear, was estimated at each grid point and 3-h interval as the magnitude of the vector difference between the 500 hPa and 10 m winds. The daily mean S06 was then determined at each ERA-5 grid point. Daily anomalies of MUCAPE, S06, and CP were calculated by subtracting each calendar day's mean from every calendar day, following

$$\boldsymbol{H'(d,y) = H(d,y) - \bar{H}(d)} \quad (1)$$

where y is year, d is calendar day, H is the variable or parameter of consideration, and the overbar denotes the long-term mean.

Following Graber et al. (2024), all analyses were conducted over the period 1960-2022, and focused specifically on the warm season, defined as April to July, which is peak season of tornado activity.

2.2 Weather Regimes

To identify weather regimes, the seasonal cycle, defined as the long-term mean of 21 UTC 500H at each grid-point for each calendar day, was removed. The 500H data were then detrended by removing the linear trend of the seasonal mean (AMJJ) 500H averaged over the entire Northern Hemisphere (Fig. S1). The detrending approach removed the positive trend of hemispheric mean 500H caused by climate change while preserving the spatial patterns and potential changes of WR frequency or persistence. While many previous studies applied a low-pass filter or/and EOF dimension reduction prior to K-means clustering analysis (e.g., Grams et al., 2020; Lee et al., 2023; Lee and Messori, 2024; Robertson et al., 2020), (Falkena et al., (2020) cautioned against the use of either EOFs or time filtering on top of K-means clustering. Our analysis shows that the application of the 5-day running mean or EOF dimension reduction prior to K-means does not qualitatively affect the regime patterns or the regime frequencies (Figs. S2-S4). We thus chose to use the simplest procedures for regime classification. Additionally, unlike Tippett et al., (2024), 500H anomalies are not normalized as we focus on one season, so seasonality is not as much of a concern. K-means clustering analysis was applied to the 500H daily anomalies over the CONUS, and the number of clusters was determined as five using the elbow method (Miller et al. 2020; Kodinariya and Makwana 2013). A persistent WR was defined as a WR lasting for ≥5 consecutive days. It is worth mentioning that different from the EOF analysis by Cwik et al. (2022), WRs identified using K-means clustering analysis are not required to be orthogonal to each other. WRs can thus more flexibly represent various recurrent synoptic-scale circulation patterns.

2.3 Tornado Reports

Tornado reports for the period 1960-2022 were obtained from the NOAA Storm Prediction Center Severe Weather Database. These reports are georeferenced with time, date, and EF/F rating. TDs were defined as any day with ≥1 EF/F-1+ tornadoes, and TOs were defined as any day with >10 EF/F-1+ tornadoes. The >10 threshold provides a larger sample size than higher thresholds but has a similar trend as >20 or >30 thresholds (Graber et al. 2024). EF/F-0 reports were not included due to their reporting uncertainty (Brooks et al., 2014; Trapp, 2013). Nevertheless, there are remaining and well-known biases in this dataset, which we attempt to

manage with a focus on days with tornadoes rather than tornado counts (e.g., Brooks et al. 2014; Graber et al. 2024; Trapp, 2014).

The TD probability anomalies (P_a) were calculated at each grid-point for each WR as follows:

$$P_a = \frac{P_i - P_c}{P_c} \times 100 \quad (2)$$

where P_c is the climatological mean TD probability and was calculated as the total number of TDs divided by the total number of days in the warm-season from 1960-2022, and P_i represents the TD probability for WR-i (i.e., the number of TDs for WR-i divided by the total WR-i days). WRs that are (are not as) conducive for TDs would have probabilities above (below) the climatological mean and thus positive (negative) probability anomalies. The probability anomalies of TOs and the probability anomalies associated with persistent and non-persistent WRs were calculated similarly. A Monte Carlo simulation test with 10000 resamples was used to test for significance of the anomalies. The number of WR-i days was multiplied by the climatological mean TD probability to get an expected number of tornado days. The p-value was calculated based on the proportion of simulations that were more extreme than the observations.

2.4 Empirical model for tornado activity

Using WR frequency and tornado probabilities for both persistent (subscript *p)* and non-persistent (subscript *np*) WRs, we developed an empirical model to assess the relationship between the variability of seasonal tornado activity and WRs:

$$\boldsymbol{TI(t)} = \sum_{i=1}^{5} \boldsymbol{f(i,t)_p \times P_{i,p}} + \sum_{i=1}^{5} \boldsymbol{f(i,t)_{np} \times P_{i,np}} \quad \textbf{(3)}$$

where TI(t) denotes a tornado index for year t. The model takes the count of WR-i days (**f (i, t)**) for year t and multiplies it by the tornado probability corresponding to WR-i (**P(i)**). The WR count is a function of regime (i) and year (t). The WR tornado probability is only a function of regime (i) and represents the likelihood that a TD will occur. Probabilities are assessed for persistent and non-persistent WRs separately, under the hypothesis that persistent WRs contribute to stronger TD or TO anomalies (Miller et al., 2020; Trapp, 2014). Spearman Rank correlation is used to compare the modeled and observed tornado indices.

3 Weather Regimes and Tornado Activity

The composite mean 500H anomalies for each WR are shown in Fig. 1, ordered with decreasing frequency of occurrence. WR-A is the most frequent regime and is characterized by an anomalous high over the west-central-CONUS and a weak anomalous low over the Southeast. WR-B and WR-C are both characterized by a three-cell wave pattern, with negative and positive 500H anomalies over the central-CONUS, respectively. WR-D and WR-E are west-east dipole patterns that nearly mirror each other. Some WRs are similar to the year-round WRs in Lee et al.,

(2023), which were subsequently used by Tippett et al., (2024). More specifically, WR-A features spatial similarities to a Pacific Trough, WR-B and WR-D show warm and cool phases of a Pacific Ridge associated with ENSO, and WR-E is characterized by an Alaskan Ridge. WR-C features spatial similarities to a Greenland High as well. It is worth mentioning that our study focuses on a different region, a specific season and chooses a different k value, and there are thus noticeable differences. WR-A features two anomalous highs over the two coasts as opposed to one anomalous high over the central-CONUS. The anomalous low in WR-B is more pronounced than in Lee et al., (2023). The anomalous high in WR-C is wavelike unlike the Greenland high in Lee et al., (2023). The dipoles in WR-E are further south than they are in the Alaskan Ridge in Lee et al., (2023).

The potential links between these WRs and tornado activity are indicated by MUCAPE and S06 (Fig. 1a-e). Composite anomalies of these parameters were calculated by subtracting the corresponding climatological mean (Fig. 1f) from the composite mean associated with each WR. The high values of the climatological MUCAPE across the central- and southeastern CONUS are connected to the physical geography of North America (Brooks et al., 2003; Trapp, 2013) and the warm-season climatological mean 850-hPa circulation, with southerly winds transporting heat and moisture into the central-CONUS (Mercer and Bates, 2019). The climatological S06 is characterized by high values over the eastern-CONUS, which are tied to the midlatitude jet stream. With an anomalous high over the west-central-CONUS and an anomalous low over the Southeast, WR-A favors anomalously low MUCAPE and S06 relative to climatological means. In contrast, the anomalous low over central North America and the anomalous high over the southeastern U.S. in WR-B imply enhanced westerly flow and increased moisture and warm-air transport from the Gulf, leading to positive S06 and MUCAPE anomalies in southeastern U.S. The favorable anomalies presented in WR-B agree with the Pacific Ridge findings in Tippett et al. (2024). In WR-C, the anomalous low over western North America and the anomalous high over central North America imply enhanced southerly flow and increased moisture and heat transport leading to positive MUCAPE anomalies in the central U.S., which overlap with reduced S06 south of the anomalous high. For WR-D, the anomalous high over the eastern CONUS and the anomalous low over the western CONUS imply enhanced southerly flow and increased moisture and heat transport, consistent with positive MUCAPE anomalies in the central U.S., while S06 decreases in the south of the anomalous high and increases in the north. WR-E, in contrast to WR-D, implies reduced southerly flow and decreased moisture and warm-air transport and is associated with negative MUCAPE anomalies in the central U.S., but S06 increases substantially over the Southeast.

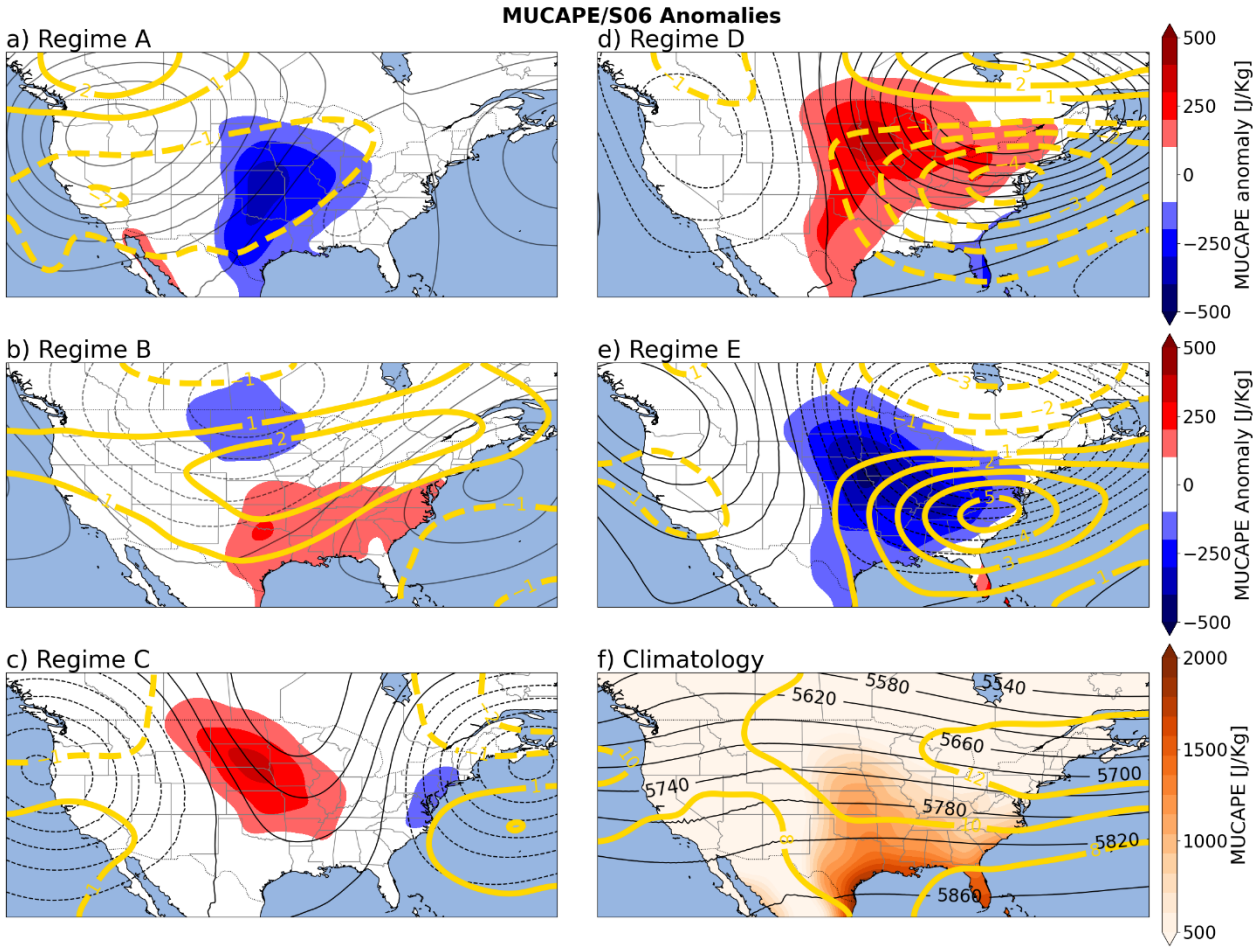

Figure 1: (a-e) Composite anomalies of 500H (black contours, +/- 10 m) for each warm-season WR and corresponding anomalies of daily maximum MUCAPE (units: J Kg^{-1}; color fill) and daily mean S06 (units: m s^{-1}; gold contours); (f) Warm-season climatology of 500H (black contours), daily maximum MUCAPE (color fill) and daily mean S06 (gold contours).

The WR-tornado activity link is illustrated by the composite anomalies of TD probability and CP for each WR. The climatological TD probability and CP are also shown (Fig. 2f) for reference. Here TD probability anomalies are evaluated following Eq. 2 with respect to P_c at each grid point. Convective-storm occurrences can be approximated using CP. Convective storms are a necessary but insufficient condition for tornadoes, so more CP does not necessarily lead to more tornadoes, but less CP likely means reduced tornado activity (Tippett et al., 2014). CP anomalies collocate well with the MUCAPE anomalies (Fig. 1) since non-zero CAPE is generally necessary for deep convection, but CP also includes information about convection initiation. WR-A has negative TD anomalies in the central-CONUS, where negative anomalies in CP and MUCAPE/S06 are also present. Positive TD anomalies in the Southeast and Midwest of WR-B are collocated with positive anomalies in CP, MUCAPE, and S06. Despite the negative S06 anomalies, positive TD anomalies occur in the central Great Plains in a region of positive MUCAPE and CP anomalies for WR-C. Weak, negative TD anomalies in association with WR-C are found in the Southeast where negative CP anomalies are present. Positive (negative) TD

anomalies in WR-D over the central-CONUS are collocated with positive (negative) CP and MUCAPE anomalies, and reduced S06 (Fig. 1d) also contributes to the negative TD anomalies in the Southeast. Finally, negative TD anomalies occur in the central-CONUS, collocated with negative anomalies of MUCAPE and CP associated with WR-E, while positive TD anomalies occur in the Southeast despite reduced MUCAPE. The latter can probably be attributed to the strong positive anomalies in S06 (Figs. 1e and 2e). However, given the low climatological TD probability in the Southeast (Fig. 2f), the absolute changes in TD days may not be high. Overall, the distribution of TD anomalies shows a good agreement with CP and MUCAPE anomalies of the same sign, and S06 seems to play a secondary role in most regions.

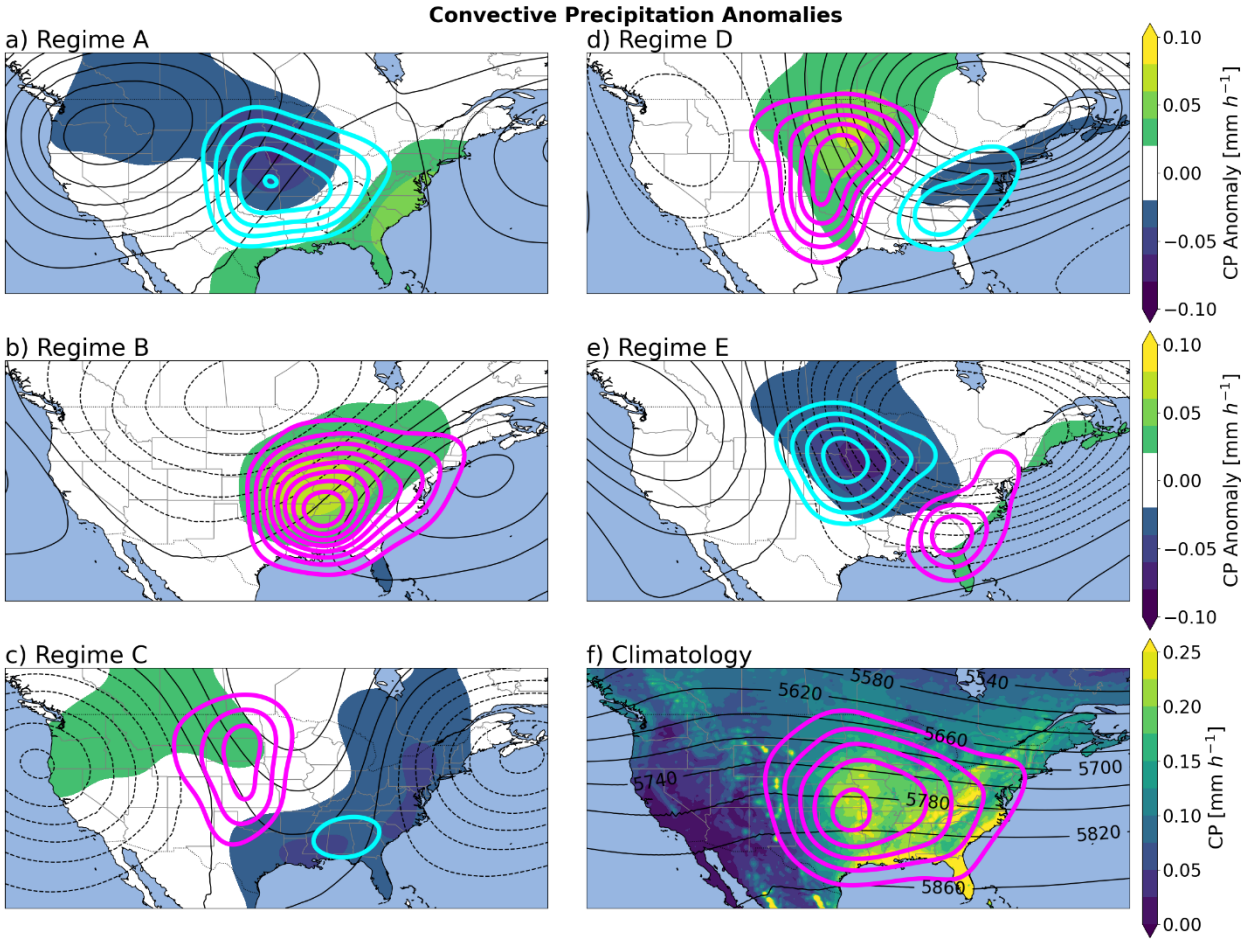

Figure 2: a-e): Significant composite anomalies of daily maximum CP rate (units: mm h^{-1}; shading) at each grid point (green/blue filled contours) for each WR A-E days with TD probability anomalies (contour intervals: +/- 10 %; magenta and cyan colors represent positive and negative values, respectively) and 500H for each WR (black contours: +/- 10 m); (f) Climatology of daily maximum CP (shading), 500H (black contours), and tornado day probability (magenta contours). Significance is tested at each grid point using a one-sample t-test with the null hypothesis that the sample mean of the anomalies is equal to zero, and anomalies with p-values≤0.05 are regarded as significant.

The link between WRs and geospatially aggregated tornado activity is summarized in Fig. 3 for different regions. There are a total of 4348 warm-season TDs from 1960-2022, therefore $P_c \approx$ 56.5%. TD probability is enhanced for WR-B and WR-D, with probability anomalies of **16.5%** and **8.2%**, respectively (corresponding to TD probabilities of **65.9%** and **61.3%**; Fig. 3). For reference, these are associated with large positive TD probability anomalies in the Southeast and central-CONUS, respectively (Fig. 2). TD probability is reduced for WR-A and WR-E, associated with negative tornado-report anomalies across the central-CONUS (Fig. 2). The TD probability anomaly associated with WR-C is close to zero (Fig. 3), which can be attributed to the cancellation between the opposite anomalies in the Southeast and central-CONUS (Fig. 2).

There are 415 warm-season TOs from 1960-2022, therefore P_c = 5.4%. In general, the TO probabilities have a stronger signal than TDs (yellow bars in Fig. 3). For TOs, WR-A has the strongest negative signal with a **-70.01%** anomaly while WR-B has the strongest positive signal with a **+69.81%** anomaly (Fig. 3). The TO anomalies are consistent with the analysis in Figs. 1-2, in which WR-A (WR-B) showed reduced (enhanced) MUCAPE, S06, and CP over the central and Southeast CONUS. The TO probability associated with WR C is around the climatological mean. WR-D has a positive, but not significant anomaly of TO probability (**+19.56%;** Fig. 3), which is consistent with enhanced MUCAPE and CP over the central-CONUS (Figs. 1d&2d). WR-E is associated with a negative TO probability anomaly, which can be linked to the reduced tornado activity over the central-CONUS (Fig. 2). Further analysis reveals that roughly **83%** of all TOs occur during a WR-B, C, or D.

We also checked TOs using >20 and >30 tornadoes thresholds (red and purple bars in Fig. 3). The analysis based on the >20 threshold yields similar results as that defined based on the >10 threshold. Although WR-A and WR-B demonstrate significant and consistent signals for days with >30 tornadoes, the other WRs exhibit contrasting signals for the >10 and >30 thresholds. This could be due to the small sample size of TOs when using the >30 threshold, and those anomalies are not significant.

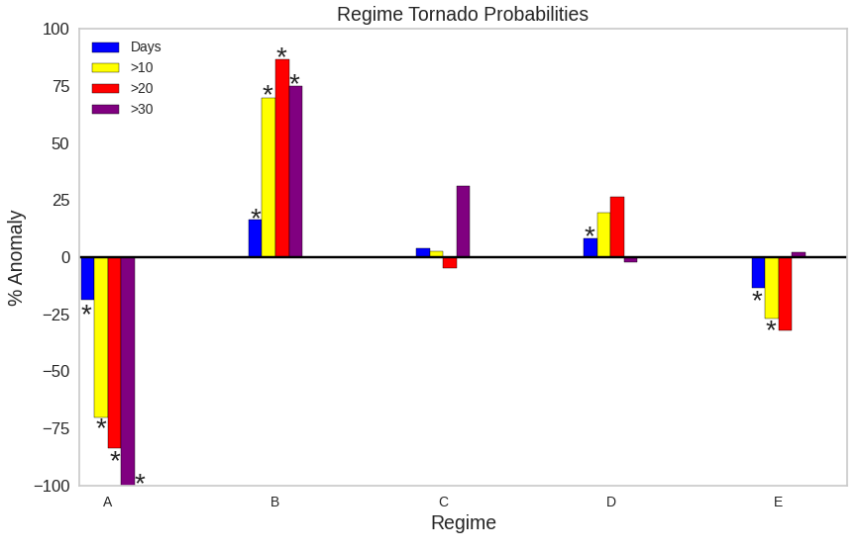

Figure 3: Tornado probability anomalies for days with > 0, > 10, > 20, & > 30 tornadoes for each WR in the CONUS (see Eq. 2 and the related discussion). Anomalies above the 95 % confidence level based on the Monte Carlo testing (with 10,000 resampling) are regarded as significantly different from zero and marked with an asterisk.

Next we compare persistent and non-persistent WRs to test the hypothesis that persistent regimes amplify the TD/TO probability anomalies. Persistent WRs are defined as those lasting for at least 5 days. The comparison of the TD probability anomalies between persistent and non-persistent WRs (Fig. 4a) does not fully support our hypothesis. Although persistent WR-B and WR-D are associated with a stronger positive anomaly in TD probability than their non-persistent counterparts, the negative TD probability anomalies are about the same for persistent and non-persistent WR-A, and persistent WR-E shows an even weaker decrease in TD probability than non-persistent WR-E. Persistent and non-persistent WR-C show TD probability anomalies of opposite signs, both with a small magnitude. Compared to TD probability, the anomalies of TO probability are generally stronger for both persistent WRs and non-persistent WRs. In particular, persistent WR-B is associated with an increase of TO probability by close 80% and accounts for about 30% of all TOs, while persistent WR-A is associated with a decrease of TO probability by about 70%. However, a consistent picture emerges: persistent WRs amplify the positive anomalies but do not further enhance the negative anomalies in comparison to the corresponding non-persistent WRs. The exception is persistent WR-C, which is associated with nearly zero TO anomaly, in contrast to a positive but not significant anomaly for non-persistent WR-C (Fig. 4b).

The asymmetric impacts of WR persistence on positive and negative tornado activity anomalies are also illustrated in Fig. S5. One possible interpretation is that tornado activity indices are positively defined metrics so they cannot be reduced much further when already close to zero. Additionally, the results should also be interpreted with caution given the limited sample sizes for certain groups (Table S1).

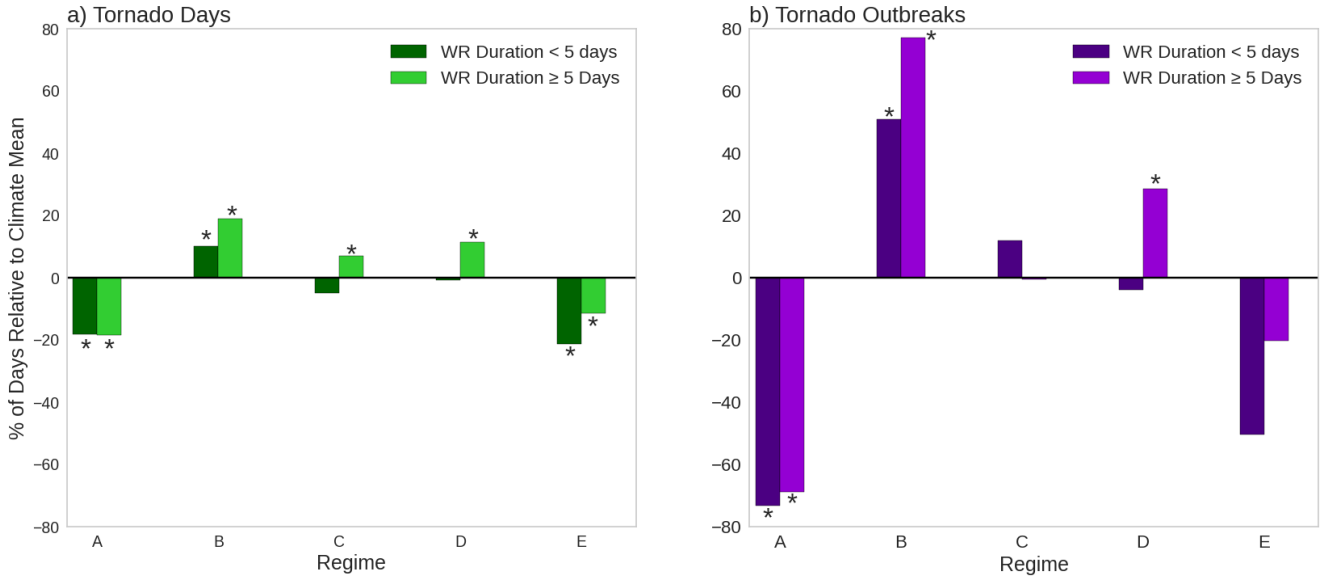

Figure 4: (a) TD and (b) TO (days with > 10 EF-1+ tornadoes) probability anomalies for each WR for persistent and non-persistent days. Anomalies above the 95 % confidence based on the Monte Carlo testing are marked with an

asterisk. Anomalies above the 95 % confidence level based on the Monte Carlo testing (with 10,000 resampling) are regarded as significant and marked with an asterisk.

4 Variability of WRs and Tornado Activity

In this section, we further quantify the link between WRs and tornado activity. WR frequencies demonstrate strong interannual and decadal variability (Fig. S6a-e). In particular, WR-A exhibits a frequency increase during the 1980s and reached a peak around 1990, coinciding with the steepest decrease in TDs (Brooks et al., 2014; Graber et al., 2024); and WR-D shows a negative trend of occurrence in the recent two decades. The frequencies of persistent WRs also show changes across different multidecadal time periods (Fig. S6f).

To examine whether WRs can help explain the interannual and decadal variability of tornado activity over the period 1960-2022, an empirical model was developed following Eq. 3. Figure 5a shows the empirically modelled TDs along with the observed TDs. Despite the decadal variability of WR and persistent WR frequencies (Fig. S6), the empirical model fails to capture the observed decreasing trend or the decadal shift in the 1980s. After detrending the observations using the least-squares fit, the model reasonably represents the interannual variability of TDs (Fig. 5b), with a rank correlation of **0.35** (p-value ~0.04). An empirical framework for EF-3+ TDs was also tested, yielding a rank correlation of **0.37** (Fig. S7). It is interesting to note that the modelled TDs are nearly out of phase with observations in the 1960s, when tornado reports are less reliable (Trapp, 2013). After excluding these years, we reconstructed the empirical model using updated TD probabilities during 1970-2022, and the correlation increases to **0.46** (Fig. S8a)**.** The empirical framework was also tested for EF-3+ TDs during 1970-2022, and the correlation is **0.49** (Fig. S8b).

We also examined TOs. The TO time series from the empirical model has a significant rank correlation (above the 95% confidence level) with the observations but it underestimates the observed variance. Since the observed TOs do not have a strong trend, detrending the data does not affect the results appreciably. Similar to the TD model results (Fig. 5a, b), the TO model is nearly out of phase with the observations in 1960s. After excluding the data in 1960s, the correlation between the empirical model and observation increases to **0.43** from 1970-2022 (Fig. S8c).

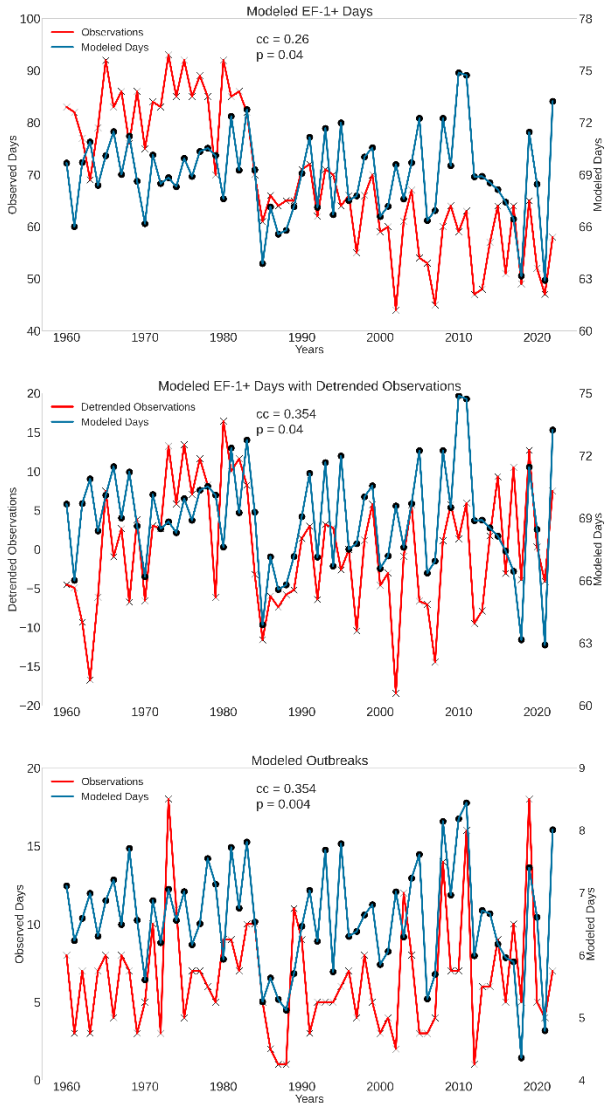

Figure 5: Empirically modeled TDs (blue with circles) per year overlaid with (a) observed TDs (red with crosses) and (b) detrended observed TDs (red with crosses) with spearman rank correlation coefficient (cc) and p-value; (c) empirically modeled (blue with circles) and observed (red with crosses) TOs per year with the spearman rank correlation coefficient and p-value.

It is worth noting that although the empirical model captures the interannual variability of TDs reasonably well, it misses the negative trend or decadal variability of TDs. The empirical model is constructed under the assumption that probability anomalies of tornado activity associated with the WRs do not change during the period of analysis. This assumption, however, may not be strictly valid. For example, MUCAPE increases over time for all five WRs, although S06 undergoes smaller changes (Fig. 6a-f, S9-S10). Additionally, convective inhibition (CIN) increases in the Southeast for WR-B (Fig. 6h) and in the central-CONUS for WR-C (Fig. 6i) from P1 to P3. Further analysis reveals lower TD probability for various regimes in P3 than in P1 (not shown). This may explain why the empirical model fails to capture the trend of TDs. A better understanding of dynamic and thermodynamic anomalies associated with WRs and the

role of internal climate variability and anthropogenic forcing in modulating WRs will help us better understand tornado activity on the decadal and longer time scales.

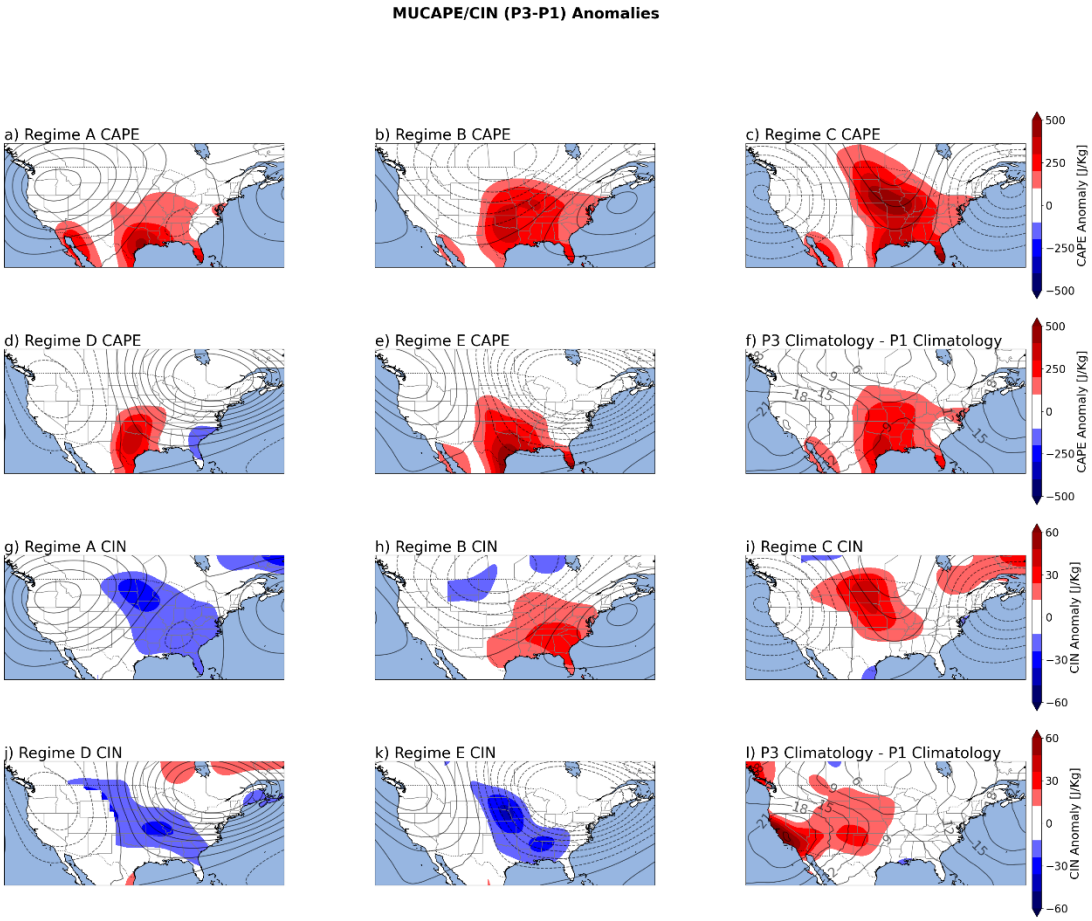

Figure 6: Change in MUCAPE (shading) anomalies (a-e) and CIN (shading) anomalies (g-k) from period 1 (1960-1979) to period 3 (2000-2022) for each WR and (f, l) all WR days.

5. Summary

The weather regime concept was used to investigate the link between synoptic-scale circulation patterns during the warm season and the variability of corresponding tornado activity over the U.S. on the interannual time scale. Five WRs were identified over North America using the K-means clustering analysis of daily 500H anomalies from the ERA-5 reanalysis. WR-A is a three-cell wave-pattern and is associated with negative anomalies of tornado activity in the central-CONUS, which is consistent with negative anomalies of MUCAPE and S06 over that region. WR-B is a three-cell wave-pattern that contributes to increased tornado activity in the Southeast as evidenced by positive anomalies in MUCAPE and S06 there. WR-C is a three-cell wave-pattern with negative 500H anomalies over both coasts. It is associated with positive MUCAPE anomalies over the central-CONUS and negative S06 anomalies. It exhibits climatologically

average tornado activity, but it does make a positive, spatially small, contribution to tornado activity in the Great Plains. WR-D and WR-E are both dipole patterns with positive and negative 500H anomalies over the east coast, respectively. WR-D contributes to anomalously positive tornado activity in the Great Plains while WR-E contributes to anomalously negative tornado activity in the Great Plains. WR-E also contributes to positive anomalies of tornado activity in the Southeast. A year that includes a high number of WR-B days is likely to have an above average number of TDs and TOs. In contrast, a year with a high number of WR-A days would likely have a below average number of TDs and TOs

We tested the hypothesis that WR persistence amplifies the tornado activity anomalies, regardless of positive or negative anomalies. However, the impacts of WR persistence on positive and negative tornado activity anomalies are found to be asymmetric: persistent WRs amplify the positive anomalies but may not further enhance the negative anomalies. This can probably be attributed to the positive-definite nature of tornado activity indices. While persistent WRs with favorable environmental conditions (such as WRs B and D) may further increase tornado activity, TD or TO probability cannot be reduced much further by the persistence of a tornado-unfavorable WR (such as WR-A) when they are already close to zero.

Using WR frequency and persistence, an empirical model was developed to quantify the relationship between tornado activity and warm-season WRs. The performance of the empirical model shows promising skill in estimating the interannual variability of tornado days or TO days, and the model performance was better after excluding the data in the 1960s. The empirical model, however, misses the trend or the multi-decadal variability of TDs. This model deficiency could be attributed to the non-stationary relationship between WRs and tornado activity on the multi-decadal time scale, which is illustrated by the increase in CAPE for all WRs in the more recent decades. The roles of internal variability and anthropogenic forcing, however, are outside the scope of the present study and merit further investigation. Furthermore, although not explored in this study, WRs and tornado activity may both be modulated by large-scale, low-frequency climate modes. WRs could potentially act as an intermediary between large-scale climate modes and tornado activity, while the low-frequency modes may be important sources of predictability for the interannual variability of tornado activity. Overall, weather regimes offer a promising path for developing skillful seasonal tornado prediction models. Such efforts are ongoing and will be reported in due course.

Code Availability

Weather Regime identification code is available at:

https://github.com/Matt0604/Kmeans

Data Availability

The ERA5 data are available at the NCAR research data archive (RDA) (ds 633.0) and Copernicus Climate Data Store (CDS).

doi: 10.5065/BH6N-5N20

The tornado report data used in this study are available through the NOAA Storm Prediction Center severe weather database.

https://www.spc.noaa.gov/wcm/#data

Author Contributions

Conceptualization: ZW, RJT, MG

Methodology: MG, ZW, RJT

Project Administration: RJT, ZW

Supervision: ZW, RJT

Writing – Original Draft: MG

Writing – Review and Edits: MG, ZW, RJT

Competing Interests:

Authors declare they have no competing interest.

Acknowledgements

We acknowledge the NCAR Computation and Information Systems Laboratory (CISL) for providing computing resources through Derecho. All ERA5 data are available used in this study are available at the research data archive. Tornado report data is available through the NOAA Storm Prediction Center severe weather database.

References

Allen, J. T., Tippett, M. K., and Sobel, A. H.: Influence of the El Nino/Southern Oscillation on tornado and hail frequency in the United States, Nat. Geosci., 8, 278–283, https://doi.org/10.1038/ngeo2385, 2015.

Ashley, W. S. and Strader, S. M.: Recipe for Disaster: How the Dynamic Ingredients of Risk and Exposure Are Changing the Tornado Disaster Landscape, Bull. Am. Meteorol. Soc., 97, 767–786, https://doi.org/10.1175/BAMS-D-15-00150.1, 2016.

Brooks, H. E., Lee, J. W., and Craven, J. P.: The Spatial Distribution of Severe Thunderstorm and Tornado Environments from Global Reanalysis Data, Atmospheric Res., 67–68, 73–94, https://doi.org/10.1016/S0169-8095(03)00045-0, 2003.

Brooks, H. E., Carbin, G. W., and Marsh, P. T.: Increased Variability of Tornado Occurrence in the United States, Science, 346, 349–352, https://doi.org/10.1126/science.1257460, 2014.

Charney, J. G. and DeVore, J. G.: Multiple Flow Equilibria in the Atmosphere and Blocking, J. Atmospheric Sci., 36, 1205–1216, https://doi.org/10.1175/1520-0469(1979)036<1205:MFEITA>2.0.CO;2, 1979.

Cook, A. R. and Schaefer, J. T.: The Relation of El Nino-Southern Oscillation (ENSO) to Winter Tornado Outbreaks, Mon. Weather Rev., 136, 3121–3137, https://doi.org/10.1175/2007MWR2171.1, 2008.

Corti, S., Molteni, F., and Palmer, T. N.: Signature of Recent Climate Change in Frequencies of Natural Atmospheric Circulation Regimes, Nature, 398, 799–802, https://doi.org/10.1038/19745, 1999.

Cwik, P., McPherson, R. A., Richman, M. B., and Mercer, A. E.: Climatology of 500-hPa Geopotential Height Anomalies Associated with May Tornado Outbreaks in the United States., Int. J. Climatol., 43, 893–913, https://doi.org/10.1002/joc.7841, 2022.

Del Genio, A. D., Yao, M.-S., and Jonas, J.: Will Moist Convection be Stronger in a Warmer Climate?, Geophys. Res. Lett., 34, https://doi.org/10.1029/2007GL030525, 2007.

Diffenbaugh, N. S., Scherer, M., and Trapp, R. J.: Robust Increases in Severe Thunderstorm Environments in Response to Greenhouse Forcing, Proc. Natl. Acad. Sci. U. S. A., 110, 16361–16366, https://doi.org/10.1073/pnas.1307758110, 2013.

Falkena, S. K. J., Wiljes, J. de, Weisheimer, A., and Shepherd, T. G.: Revisiting the identification of wintertime atmospheric circulation regimes in the Euro-Atlantic sector, Q. J. R. Meteorol. Soc., 146, 2801–2814, https://doi.org/10.1002/qj.3818, 2020.

Gensini, V. A. and Brooks, H. E.: Spatial Trends in United States Tornado Frequency, Nature, 1, https://doi.org/10.1038/s41612-018-0048-2, 2018.

Graber, M., Trapp, R. J., and Wang, Z.: The Regionality and Seasonality of Tornado Trends in the United States, Npj Clim. Atmospheric Sci., 7, https://doi.org/10.1038/s41612-024-00698-y, 2024.

Grams, C. M., Beerli, R., Pfenninger, S., Staffell, I., and Wernli, H.: Balancing Europe's Wind-power Output through spatial development informed by weather regimes, Nat. Clim. Change, 7, 557–562, https://doi.org/10.1038/nclimate3338, 2017.

Grams, C. M., Ferranti, L., and Magnusson, L.: How to make use of weather regimes in Extended-range Predictions for Europe, ECMWF Newsl., 2020.

Hannachi, A., Straus, D. M., Franzke, C. L. E., Corti, S., and Woollings, T.: Low-Frequency Nonlinearity and Regime Behavior in the Northern Hemisphere Extratropical Atmosphere, Rev. Geophys., 55, 199–234, https://doi.org/10.1002/2015RG000509, 2017.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. R. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Kodinariya, T. M. and Makwana, P. R.: Review on Determining Number of Clusters in K-Means Clustering., Int. J. Adv. Res. Comput. Sci. Manag. Stud., 1, 90–95, 2013.

Lee, S. H. and Messori, G.: The Dynamical Footprint of Year-Round North American Weather Regimes, Geophys. Res. Lett., 51, 2024.

Lee, S. H., Tippett, M. K., and Polvani, L. M.: A New Year-Round Weather Regime Classification for North America, J. Clim., 36, 7091–7108, https://doi.org/10.1175/JCLI-D-23-0214.1, 2023.

Mercer, A. E. and Bates, A.: Meteorological Differences Characterizing Tornado Outbreak Forecasts of Varying Quality, Atmosphere, 1, 16, https://doi.org/10.3390/atmos10010016, 2019.

Mercer, A. E., Shafer, C. M., Doswell III, C. A., Leslie, L. M., and Richman, M. B.: Synoptic Composites of Tornadic and Nontornadic Outbreaks, Mon. Weather Rev., 140, 2590–2608, https://doi.org/10.1175/MWR-D-12-00029.1, 2012.

Michelangeli, P.-A., Vautard, R., and Legras, B.: Weather Regimes: Recurrence and Quasi Stationarity, J. Atmospheric Sci., 52, 1237–1256, https://doi.org/10.1175/1520-0469(1995)052%3C1237:WRRAQS%3E2.0.CO;2, 1995.

Miller, D., Wang, Z., Trapp, R. J., and Harnos, D. S.: Hybrid Prediction of Weekly Tornado Activity Out to Week 3: Utilizing Weather Regimes, Geophys. Res. Lett., 47, https://doi.org/10.1029/2020GL087253, 2020.

Miller, D., Gensini, V. A., and Barrett, B. S.: Madden-Julian Oscillation Influences United States Springtime Tornado and Hail Frequency, Npj Clim. Atmospheric Sci., 5, https://doi.org/10.1038/s41612-022-00263-5, 2022.

Moore, T. W.: Annual and Seasonal Tornado Trends in the Contiguous United States and its Regions, Int. J. Climatol., 38, 1582–1594, https://doi.org/10.1002/joc.5285, 2018.

Moore, T. W. and DeBoer, T. A.: A Review and Analysis of Possible Changes to the Climatology of Tornadoes in the United States, Prog. Phys. Geogr. Earth Environ., 43, https://doi.org/10.1177/0309133319829398, 2019.

NCEI: U.S. Billion-Dollar Weather and Climate Disasters, 2024.

Niloufar, N., Devineni, N., Were, V., and Khanbilvardi, R.: Explaining the Trends and Variability in the United States Tornado Records using Climate Teleconnections and Shifts in Observational Practices, Sci. Rep., 11, https://doi.org/10.1038/s41598-021-81143-5, 2021.

Rasmussen, E. N. and Blanchard, D. O.: A Baseline Climatology of Sounding-Serived Supercell and Tornado Forecast Parameters., Weather Forecast., 13, 1148–1164, https://doi.org/10.1175/1520-0434(1998)013<1148:ABCOSD>2.0.CO;2, 1998.

Robertson, A. W. and Ghil, M.: Large-Scale Weather Regimes and Local Climate over the Western United States, J. Clim., 12, 1796–1813, https://doi.org/10.1175/1520-0442(1999)012%3C1796:LSWRAL%3E2.0.CO;2, 1999.

Robertson, A. W., Vigaud, N., Yuan, J., and Tippett, M. K.: Toward Identifying Subseasonal Forecasts of Opportunity Using North American Weather Regimes, Mon. Weather Rev., 148, 1861–1875, https://doi.org/10.1175/MWR-D-19-0285.1, 2020.

Storm Prediction Center: U.S. Killer Tornado Statistics, 2024.

Strader, S. M., Ashley, W. S., Pingel, T. J., and Krmenec, A. J.: Projected 21st Century Changes in Tornado Exposure, Risk, and Disaster Potential, Clim. Change, 141, 301–313, https://doi.org/10.1007/s10584-017-1905-4, 2017.

Strader, S. M., Gensini, V. A., Ashley, W. S., and Wagner, A. N.: Changes in Tornado risk and Societal Vulnerability Leading to Greater Tornado Impact Potential, Npj Nat. Hazards, 1, 2024.

Thompson, D. B. and Roundy, P. E.: The Relationship between the Madden-Julian Oscillation and US Violent Tornado Outbreaks in the Spring, Mon. Weather Rev., 141, 2087–2095, https://doi.org/10.1175/MWR-D-12-00173.1, 1998.

Thompson, R. L., Smith, B. T., Grams, J. S., Dean, A. R., and Broyles, C.: Convective Modes for Significant Severe Thunderstorms in the Contiguous United States. Part II: Supercell and QLCS Tornado Environments, Weather Forecast., 27, 1136–1154, https://doi.org/10.1175/WAF-D-11-00116.1, 2012.

Tippett, M. K.: Changing Volatility of U.S. Annual Tornado Reports, Geophys. Res. Lett., 41, 6956–6961, https://doi.org/10.1002/2014GL061347, 2014.

Tippett, M. K., Sobel, A. H., Camargo, S. J., and Allen, J. T.: An Empirical Relation between U.S. Tornado Activity and Monthly Environmental Parameters, J. Clim., 27, 2983–2999, https://doi.org/10.1175/JCLI-D-13-00345.1, 2014.

Tippett, M. K., Lepore, C., and L'Heureux, M. L.: Predictability of a Tornado Environment Index from El Nino Southern Oscillation (ENSO) and the Arctic Oscillation, Weather Clim. Dyn., 3, 1063–1075, https://doi.org/10.5194/wcd-3-1063-2022, 2022.

546 Tippett, M. K., Malloy, K., and Lee, S. H.: Modulation of U.S. Tornado Activity by year-round North
547 American Weather Regimes, Mon. Weather Rev., https://doi.org/10.1175/MWR-D-24-0016.1, 2024.

548 Trapp, R. J.: Mesoscale-Convective Processes in the Atmosphere, Cambridge University Press, 2013.

549 Trapp, R. J.: On the Significance of Multiple Consecutive Days of Tornado Activity, Mon. Weather Rev.,
550 142, 1452–1459, https://doi.org/10.1175/MWR-D-13-00347.1, 2014.

551 Trapp, R. J., Diffenbaugh, N. S., Brooks, H. E., Baldwin, M. E., Robinson, E. D., and Pal, J. S.: Changes in
552 Severe Thunderstorm Environment Frequency during the 21st Century caused by Anthropogenically
553 Enhanced Global Radiative Forcing, Proc. Natl. Acad. Sci. U. S. A., 104, 19719–19723,
554 https://doi.org/10.1073/pnas.0705494104, 2007.

555 Vigaud, N., Tippett, M. K., and Robertson, A. W.: Probabilistic Skill of Subseasonal Precipitation Forecasts
556 for the East Africa-West Asia Sector during September-May, Weather Forecast., 33, 1513–1532,
557 https://doi.org/10.1175/WAF-D-18-0074.1, 2018.

558 Wakimoto, R. M. and Wilson, J. W.: Non-supercell Tornadoes, Mon. Weather Rev., 117, 1113–1140,
559 https://doi.org/10.1175/1520-0493(1989)117<1113:NST>2.0.CO;2, 1989.

560