Environmental Controls on the Climatological Scaling of Tornado Frequency with Intensity

CHIARA LEPORE[a] AND MICHAEL K. TIPPETT[b]

[a] *Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York;* [b] *Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York*

(Manuscript received 1 May 2020, in final form 26 August 2020)

ABSTRACT: The scaling of U.S. tornado frequency with enhanced Fujita (EF)-rated intensity is examined for the range EF1–EF3. Previous work has found that tornado frequency decreases exponentially with increasing EF rating and that many regions around the world show the same exponential rate of decrease despite having quite different overall tornado frequencies. This scaling is important because it relates the frequency of the most intense tornadoes to the overall tornado frequency. Here we find that U.S. tornado frequency decreases more sharply with increasing intensity during summer than during other times of the year. One implication of this finding is that, despite their rarity, when tornadoes do occur during the cool season, the relative likelihood of more intense tornadoes is higher than during summer. The environmental driver of this scaling variability is explored through new EF-dependent tornado environmental indices (TEI-EF) that are fitted to each EF class. We find that the sensitivity of TEI-EF to storm relative helicity (SRH) increases with increasing EF class. This increasing sensitivity to SRH means that TEI-EF predicts a slower decrease in frequency with increasing intensity for larger values of SRH (e.g., cool season) and a sharper decrease in tornado frequency in summer when wind shear plays a less dominant role. This explanation is also consistent with the fact that the fraction of supercell tornadoes is smaller during summer.

KEYWORDS: Storm environments; Tornadoes; Climatology; Regression analysis

1. Introduction

The societal risk from tornadoes is a combination of exposure and vulnerability with the underlying meteorological hazard (i.e., the frequency that tornadoes of a given intensity occur at specific locations; Ashley and Strader 2016). Tornado frequency can examined from different viewpoints, depending on the application. Short-range weather forecasts provide information about the immediate likelihood of tornado occurrence given the current conditions. For longer-time-scale applications, extended-range forecasts or predictable climate signals (e.g., ENSO and the MJO) can indicate when and where tornado activity is expected in the upcoming weeks to months (Allen et al. 2015; Lepore et al. 2018; Baggett et al. 2018; Gensini and Tippett 2019). In other cases such as risk management applications, the climatological perspective (i.e., where and when during the day and year do tornadoes occur on average; Brooks et al. 2003a; Krocak and Brooks 2018), or the climate change perspective (i.e., how is tornado frequency is expected to change in a warmer climate; Trapp et al. 2009; Diffenbaugh et al. 2013) are relevant.

Estimating the frequency of the strongest tornadoes is of particular interest because of their impact but is challenging because of their relative rarity in the observational record. One approach to gauging the climatological probability of strong tornadoes, even in locations where they have not been reported before, is to note that tornadoes follow a frequency–intensity scaling that allows the frequency of the strongest tornadoes to be predicted from the frequency of weaker tornadoes. In particular, the climatological frequency of tornadoes in the United States, Europe, South America, and other regions decreases with increasing intensity at an approximately exponential rate that varies little by region, except for steeper decreases in regions such as Florida, the Front Range of Colorado, and the United Kingdom where organized tornado storm modes (supercell and quasi-linear convective system; QLCS) are a smaller fraction of the tornado population (Brooks and Doswell 2001; Dotzek et al. 2003; Feuerstein et al. 2005; Smith et al. 2012).

Tornado intensity is measured in the United States using the damage-based EF-scale (Fujita 1981; Doswell et al. 2009), which goes from 0 (least intense) to 5 (most intense). In terms of EF rating, Brooks and Doswell (2001) identified two limiting frequency/intensity relations in the United States. For supercell tornado dominated regions, frequency $\sim \exp(-1.02 \times \text{EF rating})$, and in regions where supercell tornadoes are less dominant frequency $\sim \exp(-2.3 \times \text{EF rating})$. These log-linear scalings mean that tornado frequency decreases by 64% $[1 - \exp(-1.02)]$ from one EF class to the next higher one in supercell dominated regions, and by 90% $[1 - \exp(-2.3)]$ in regions where supercell tornadoes are less dominant. Brooks and Doswell (2001) used these log-linear slopes to extrapolate the frequency of violent (EF4 or EF5) tornadoes from the total number of tornadoes. Later, Feuerstein et al. (2005) found that tornado frequency and intensity satisfied an additional scaling constraint, namely that the overall ratio of the numbers of EF4 and EF3 tornadoes was approximately 0.177 (corresponding to a log-linear slope of −1.73). This ratio was smaller, ~0.1, indicating a sharper decrease, in regions where nonsupercell tornadogenesis was relatively frequent and larger, ~0.30, in supercell-dominated regions. The relation between storm

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

Corresponding author: Chiara Lepore, clepore@ldeo.columbia.edu

DOI: 10.1175/MWR-D-20-0138.1

mode and the frequency–intensity scaling was further clarified by Trapp et al. (2005) who used radar data to classify 3 years of tornadoes by storm type and found that tornado frequency decreased about 35% per EF class for supercell tornadoes and by about 22% for quasilinear convective storm (QLCS) tornadoes.

Improved understanding of the frequency–intensity relation in tornadic occurrences, for different regions and months of the CONUS, potentially has value in applications such as seasonal forecasting, climate change impacts, and hazard modeling. Previous studies have shown how storm mode climatologies can be informative in describing the mechanism behind these relations. However, these studies have so far focused on relatively small samples, primarily because of the extensive work that goes into extracting storm modes climatologies. In our work, instead of storm mode as an explanatory factor for these relations, we make use of large-scale environments, which are already computed in the majority of reanalysis, forecast, and climate models. We first examine the seasonal and regional variations in the U.S. tornado frequency–intensity scaling. We restrict our attention to the tornadoes rated EF1–EF3 in order to have adequate sample sizes, and because the exponential scaling is a good approximation over this range. To find additional evidence that differing scaling rates are related to different underlying convective modes and large-scale environments, we make use of a tornado environmental index (TEI), which is designed to relate the climatological number of monthly tornado reports (all intensities) to monthly averages of atmospheric variables (Tippett et al. 2012, 2014). The TEI captures some of the observed geographical variability and seasonality of tornado occurrences (Lepore et al. 2018). To address the scaling of frequency with intensity, we develop a EF-dependent TEI (TEI-EF), allowing us to explore the ability of an index to capture seasonal, spatial, interannual, and intensity variability of the observed process, and to relate the different scaling behaviors with environmental parameters. The TEI-EF shows that differing scaling rates are indeed associated with different environments. In particular, the sensitivity of TEI-EF to storm relative helicity (a parameter widely used as a proxy for supercell tornadogenesis), is overall in line with the supercell versus nonsupercell explanation previously proposed.

The tornado report and environmental data are described in section 2 along with the index methodology. The new EF-dependent TEI is introduced and evaluated in section 3. TEI-EF is used to interpret the observed U.S. tornado frequency/intensity distribution in section 4. A summary and discussion are given in section 5.

2. Data and methods

Tornado reports are taken from the NOAA Storm Prediction Center (SPC) Severe Weather Database. The starting latitude and longitude of the reported tornadoes are aggregated to a $1^\circ \times 1^\circ$ grid over the contiguous United States (CONUS) and to monthly resolution over the period 1979–2016. Report data are analyzed separately for tornadoes rated EF0 (23 458 cases), EF1 (12 799 cases), EF2 (4015 cases), and EF3 (1141 cases). Tornadoes rated EF4 and greater represent less than 1% of the tornadoes in the database and are not included in the analysis here.

Following Tippett et al. (2012), the TEI-EF is developed using environments from the North American Regional Reanalysis (Mesinger et al. 2006, NARR). The reanalysis data are averaged from their native 32-km resolution to the same $1^\circ \times 1^\circ$ grid used for the reports. We consider five NARR variables: convective precipitation (cPrcp), 0–3-km storm relative helicity (SRH3), mixed-layer convective available potential energy (MLCAPE), surface air temperature (T), and 0–6-km bulk shear (S06). A large body of literature has shown that these variables are useful predictors of conditions favorable to severe weather (e.g., Brooks 2013; Brooks and Doswell 2001; Brooks et al. 2003b; Trapp et al. 2005; Tippett et al. 2012, 2014; Lepore et al. 2018). Here we consider 0–3-km SRH primarily because it is directly available in NARR. However, the depth over which SRH is calculated can be an important factor in discriminating correctly between tornadic versus nontornadic supercells (Thompson et al. 2012; Coffer et al. 2019).

The original TEI was calibrated to match total monthly numbers of tornadoes (no EF dependence) and is a function of cPrcp and SRH3 of the form:

$$\begin{aligned}\mathrm{TEI} = \exp[&-\beta_0 + \beta_{\mathrm{cPrcp}} \log(\mathrm{cPrcp}) \\ &+ \beta_{\mathrm{SRH3}} \log(\mathrm{SRH3})] D \Delta x \Delta y \cos\phi, \end{aligned} \tag{1}$$

where log(cPrcp) and log(SRH3) are the natural logarithms of the monthly averages of the environments. The regression parameters β_0, β_{cPrcp}, and β_{SRH3} are the intercept and coefficients, respectively, of a Poisson regression. The regression parameters were calibrated on the CONUS-wide, *climatological* monthly values of the environments and tornado counts—the regression parameters do not vary spatially or by season (Tippett et al. 2012). The term $D\Delta x\Delta y\cos\phi$ is the offset, which accounts for the varying grid area and month length. Δx and Δy are the longitude and latitude spacings in degrees, both equal to 1° here, ϕ is the latitude in radians, and D is the number of days in a month. The value of TEI is the expected number of tornadoes.

The Poisson regression formulation means that the TEI coefficients represent the approximate percentage change in tornado numbers per percentage change in cPrcp or SRH3 (i.e., the sensitivity of TEI to changes in the environmental variables). For instance, if cPrcp increases by 1% and SRH3 is unchanged, then

$$\begin{aligned}\frac{\mathrm{TEI}(1.01\mathrm{cPrcp})}{\mathrm{TEI}(\mathrm{cPrcp})} &= \frac{\exp[\beta_{\mathrm{cPrcp}} \log(1.01\mathrm{cPrcp})]}{\exp[\beta_{\mathrm{cPrcp}} \log(\mathrm{cPrcp})]} \\ &= \exp[\beta_{\mathrm{cPrcp}} \log(1.01)] \approx \exp[0.01\beta_{\mathrm{cPrcp}}] \\ &\approx 1 + 0.01\beta_{\mathrm{cPrcp}}. \end{aligned} \tag{2}$$

The TEI coefficients in Tippett et al. (2012) are $\beta_{\mathrm{cPrcp}} = 1.34$ and $\beta_{\mathrm{SRH3}} = 1.82$, indicating that changes in SRH3 have a slightly larger effect on the index than do changes in cPrcp.

Here, we use the same approach but compute the regression parameters separately for each EF intensity class. This gives regression parameters that depend on EF rating. We use 10 128 samples to estimate the parameters for each intensity class. The sample size of 10 128 results from the number of months in

the climatology (12) times the number of grid points in the spatial domain (26 × 71) with ocean grid points, Mexico and Canada grid points (i.e., non-CONUS points), and a few cases where environments were equal to zero removed. All other CONUS grid points, including those with zero tornado reports, are included in the regression. Because we use the climatological monthly means of tornado counts and environments in the calibration, the interannual and regional variability can be used for independent validation.

The scaling of frequency with intensity at annual and monthly resolution is computed by fitting the logarithms of observed tornado counts and TEI-EF as a linear function of the intensity values EF1–EF3 (EF0 counts are not included in the fit) using weighted-least squares with weights that are proportional to the counts. Separate analyses are performed for the CONUS and for the 9 NOAA regions (Table 1). A by-product of the TEI-EF fitting procedure is that over the complete dataset, the TEI-EF scaling matches the observed scaling. The fitting procedure does not guarantee that the TEI-EF matches the observed scaling at the regional and monthly level, which thus allows for an independent validation of the results and further exploration of how the observed scaling is related to differences in convective environments.

3. An EF-dependent tornado environment index

a. Formulation and environmental sensitivity

To select which variables to include in the new indices, we first fit all two-parameter combinations of the five environments (cPrcp, SRH3, MLCAPE, T, S06), to the four EF levels separately (EF0–EF3), and computed their in-sample deviance-based *R*-squared values (Cameron and Windmeijer 1996). Based on this metric (not shown), we selected the three best performing models, which contain the variable pairs: (cPrcp, SRH3), (cPrcp, T), and (MLCAPE, SRH3). Adding a third variable to these three two-parameter models gives seven unique three-parameter models (see Fig. 1, letters A to G). We also include a three-parameter model that uses neither cPrcp nor SRH (Fig. 1, letter H). We then calculated their out-of-sample deviance-based *R*-squared values and their confidence intervals based on 10 repetitions of tenfold out-of-sample cross-validation (Tippett et al. 2012); all models passed also a deviance goodness of fit test. The results (Fig. 1) show that none of the three-parameter models perform significantly better than their two-parameter reduced form: models A–C fluctuate around (cPrcp, SRH) (gray line), models D–E around (MLCAPE, SRH3) (red line), and F–G around (cPrcp, T) (blue line); the H model (MLCAPE, T, S06) does not include any pair of the variables in the best two-parameter models, and performs worst. For EF0, models A and C perform slightly better than their two-parameter form (cPrcp, SRH3). When we repeated the subsequent analysis with models A and C, there were no significant improvements or different conclusions. In the end, the calibration procedure identified the same environments used in the original two-parameter TEI (cPrcp, SRH3) as able to describe the expected number of tornadoes across intensities, albeit with different coefficients. We refer to the TEI with EF-dependent coefficients as TEI-EF.

TABLE 1. States included in the nine NOAA regions (Lepore et al. 2018).

Region	States
South	TX, LA, MS, OK, KS, AR
Southeast	AL, FL, SC, NC, GA, VA
Central	TN, KY, MO, IL, IN, OH, WV
Upper Midwest	IA, MN, WI, MI
Plains	NE, SD, ND, WI, MT
Northeast	MD, DE, NJ, NY, PA, CT, RI, MA, NH, VT, ME
Southwest	UT, CO, AZ, NM
Northwest	OR, WA, ID
West	CA, NV

Although the TEI-EF coefficients are allowed to vary with EF level, the coefficients of cPrcp are essentially independent of intensity (Fig. 2, left panel, blue lines) and do not differ significantly from the TEI coefficient (red line). This result means that the percentage change in number of tornadoes per percentage change in cPrcp is essentially independent of EF level. In contrast, the SRH3 coefficients increase exponentially with intensity level. Consequently, a 1% increase in SRH3 produces a larger percentage change in TEI-EF at higher EF levels than at lower ones.

The differing sensitivity of each EF intensity class to cPrcp and SRH3 is also apparent in the 2D histogram of tornado report data binned by cPrcp and SRH3 (Fig. 3), where the isolines of TEI (dashed lines) and TEI-EF (solid lines) are overlaid. Figure 3 allows us to verify the modeling assumption of the Poisson regressions, which is that the log of the number of tornadoes is a linear function of the covariates (natural logarithms of cPrcp and SRH), and it highlights the ability of these two environments to discriminate tornadic climatological occurrences (Brooks et al. 2003b). Finally, the TEI-EF isolines match the report data better than the TEI isolines and are flatter for higher EF levels, which indicates a greater sensitivity to SRH3 than to cPrcp.

b. Annual cycle

Across EF levels, the peak of the monthly aggregated CONUS tornado reports occurs in May, as is also the case for TEI-EF (Fig. 4). Southeast activity peaks in April, where TEI-EF incorrectly puts peak activity in March for EF1–EF3 and fails to capture the amplitude. In most regions, the month of peak activity is the same across EF levels. However, in the Northeast, EF3 reports peak in May, and lower EF levels peak in July. In the Central region, EF3 reports peak in April while the peak is in May for lower EF levels. This difference in peak month between the weakest and strongest tornadoes in the Central region is captured to some degree by TEI-EF.

At the CONUS level, the index underestimates the peak in May for all EF levels, and slightly overestimates the decline in activity from July to September. In the regional results, except for the western regions (Southwest, Northwest, and West, where the original TEI model showed similar behavior), the index reproduces the timing and ascending part of the seasonal

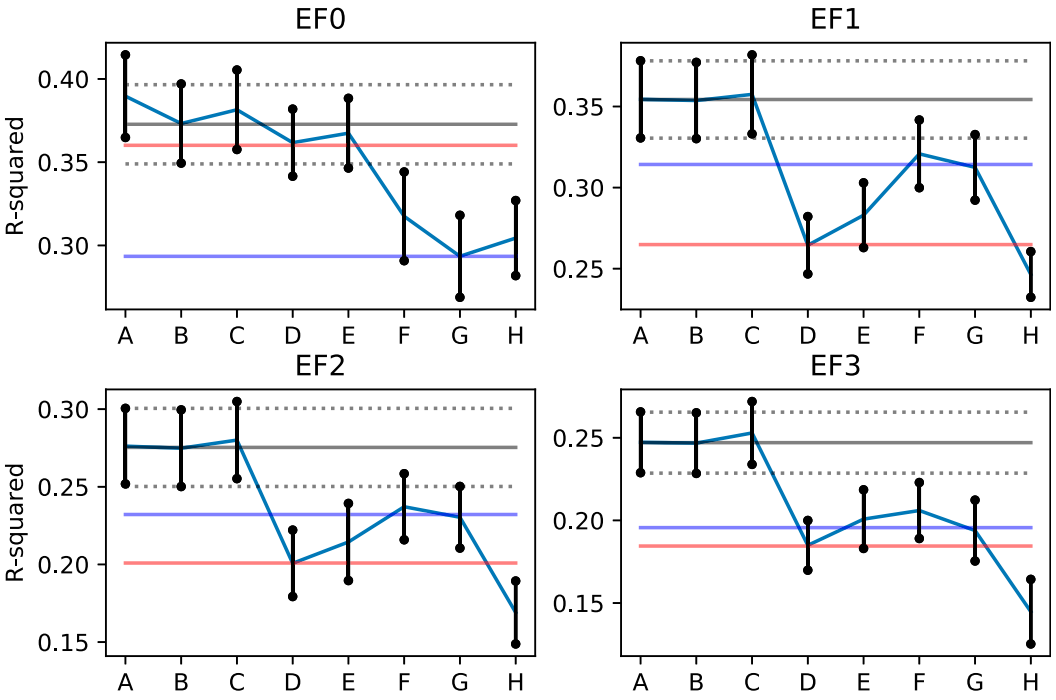

FIG. 1. The deviance based *R*-squared values for eight three-parameter regression models (A to H) and three two-parameter model (cPrcp–SRH3 in gray, MLCAPE–SRH3 in red, and cPrcp–*T* in blue solid lines) for each EF intensities classes (four panels). Error bars (vertical black lines for three-parameter models and dashed lines for cPrcp–SRH3 two-parameter model) show plus and minus one standard deviation, computed from 10 repetitions of tenfold out-of-sample cross-validation. The three-parameter models include A: cPrcp, SRH3, and MLCAPE; B: cPrcp, SRH3, and S06; C: cPrcp, SRH3, and *T*; D: MLCAPE, SRH3, and S06; E: MLCAPE, SRH3, and *T*; F: cPrcp, MLCAPE, and *T*; G: cPrcp, S06, and *T*; and H: MLCAPE, S06, and *T*.

cycle to a fair degree, even when peaks are located in different months for different regions. Generally, the index performs better for those regions where the seasonal cycle is similar across all intensities (Upper Midwest, Plains, and South), which may be because the same environments (cPrcp and SRH) are used in the index for all four EF classes.

For EF1 and greater intensities, the index somewhat overestimates the August–September values. For the same EF ranges, a secondary peak in November is present at the CONUS level and in the Central, South, Southeast, and Southwest regions. The index is unable to reproduce exactly the secondary cool season peak; however, it does record an

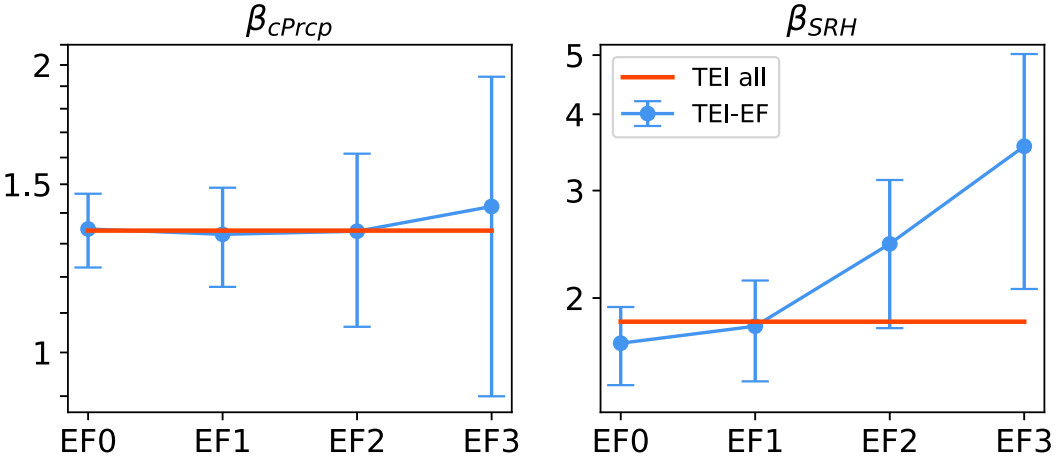

FIG. 2. TEI-EF coefficients (blue lines; logarithmic scale) (left) β_{cPrpc} and (right) β_{SRH3} and the TEI coefficients (red lines). Error bars indicate the 95% confidence intervals of the TEI-EF coefficients.

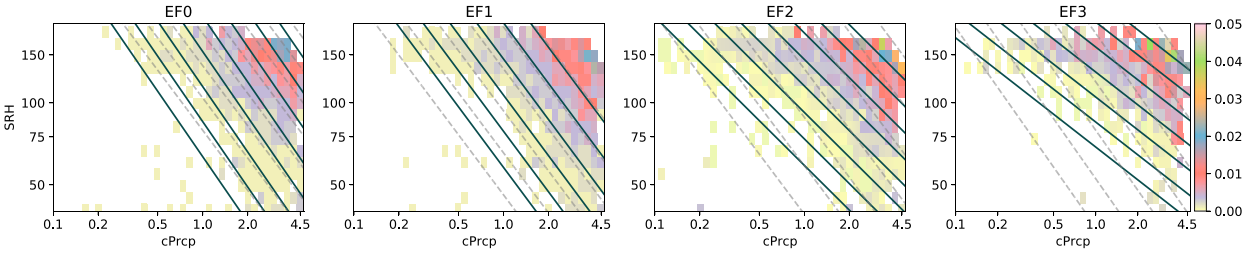

FIG. 3. 2D histograms (each panel sums to 1) for observed (in color) EF0–EF3 tornado counts as a function of cPrcp and SRH3 (bins follow a logarithmic scale). The black lines are isolines of 2D histograms of the TEI-EF index, and gray dashed lines are isolines for the TEI index, which are identical across EF classes.

upturn in October tornado occurrence for all EF classes in the South, and for EF2 and EF3 in the Southeast and Central. The Pearson correlation values between the report and TEI-EF annual cycles are high.

The spatial distributions of tornado counts per EF classes (Fig. 5) clearly identify differences between observed EF0 and higher classes at the spatial level, and features that the TEI-EF model can and cannot reproduce. The observed location of the peak of EF0 tornadoes is over Texas, Oklahoma, and Kansas, whereas the peaks of EF1 and greater intensities are more south and into Arkansas and the coastal areas of Louisiana, Mississippi and Alabama. The TEI-EF model for EF1 and greater does not reproduce the peak over Alabama and Louisiana, in line with the lower correlation values for the annual cycle in Fig. 4 for both southern regions, but does correctly reproduce the peak in northern Texas and Oklahoma. Overall, the four TEI-EF spatial distributions are more similar to each other than are the corresponding reports because, once again, TEI-EF is based on the same two underlying large-scale environments, albeit with different coefficients.

Finally, at the CONUS and regional levels, Tables 2–5 show the correlations of the 38 (one for each year considered) TEI-EF values and the observed tornado counts at the monthly (first 12 columns) and annual (last column) scale. Because each month is evaluated separately and the monthly annual cycle is removed, the correlations depend only on the interannual variability (year-to-year differences), which play no role in the fitting procedure. Therefore, the interannual correlations are a separate assessment of the performance of the model. Month-by-month correlations are overall positive and significant (in bold) for the majority of the CONUS-wide analysis and for the most active regions and seasons (Tables 2–5). These results are in line with the correlations of the all-tornado TEI (Table 1 of Lepore et al. 2018). Across EF levels and regions, either May or June show the best performing skill, which then diminishes rapidly in the remaining summer months. The worst performing month is August, with only 5 significant correlations out of

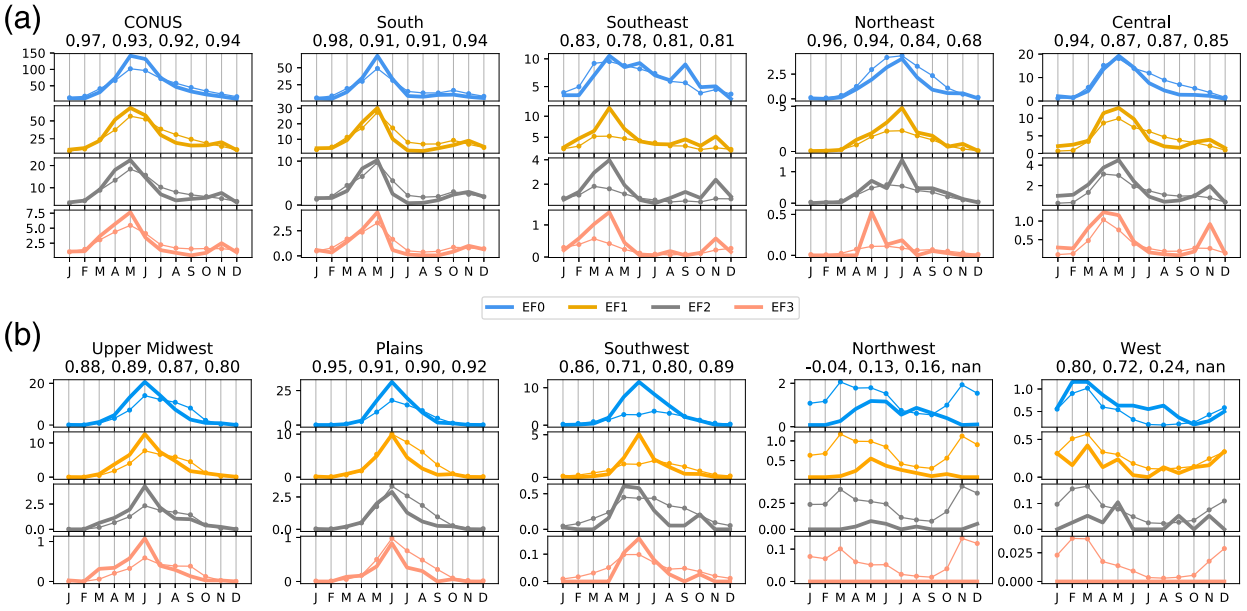

FIG. 4. Annual cycle of the average monthly tornado reports (thick curves) and TEI-EF (thin curves) at the CONUS and regional level. Pearson correlations between observations and TEI-EF results for the four intensities (EF0–EF3) are indicated in the title.

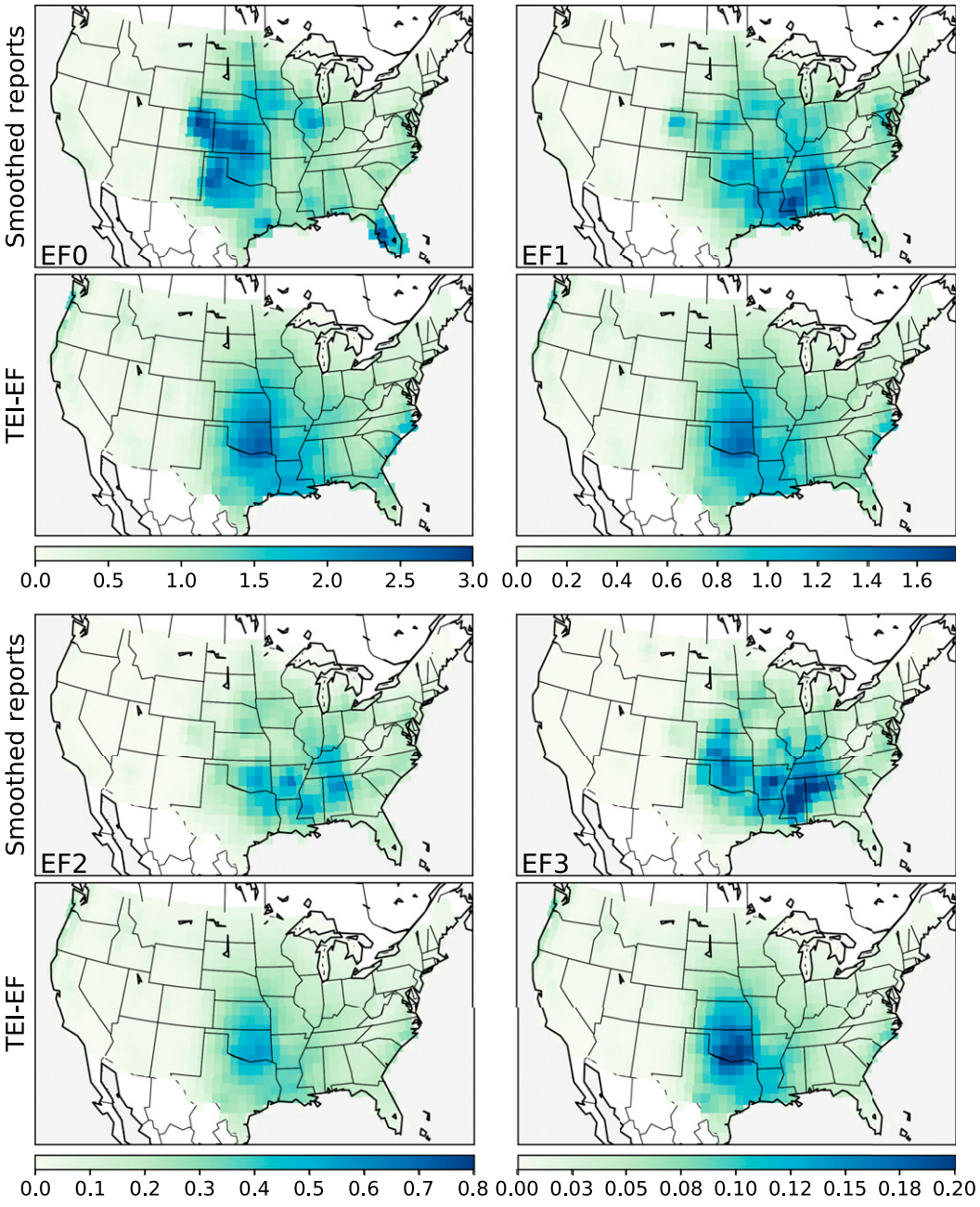

FIG. 5. Observed spatial distribution of the yearly average tornado counts for the four intensity classes: (first row) EF0 and EF1 and (third row) EF2 and EF3 compared with (second row),(fourth row) TEI-EF model. Color bars across EF classes are different.

the available 24 values, followed by September with 9 significant values out of 21. Hart and Cohen (2016) argue that environmental parameters are less effective at predicting significant tornadoes June–October because they are poorer representations of significant tornado potential and that there is a decrease in overall convective predictability related to a transition from strongly forced regimes to weakly forced regimes. The annual correlations are high for the majority of the EF classes and regions. The same correlation analysis was performed using three-parameter models (not shown), but their use did not substantially change the results.

The use of the same environments, cPrcp and SRH, for all EF classes has some limits in describing the variability of the intensity process, especially at the spatially distributed scale. However, both the seasonal (Fig. 4), and interannual variability (Tables 2–5) are reasonably well reproduced at the CONUS and regional scales.

TABLE 2. Interannual monthly (first 12 columns) and annual (last column) correlations between TEI-EF and observed tornadoes for EF = 0. Missing values are for combinations with less than 10 values recorded. Boldface values are significant at the 0.05 level.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
CONUS	**0.64**	**0.52**	**0.51**	**0.39**	**0.53**	**0.59**	**0.63**	0.19	0.1	0.29	**0.5**	**0.53**	**0.83**
South	**0.61**	**0.48**	**0.51**	**0.51**	**0.4**	**0.58**	**0.57**	0.17	0.24	**0.47**	**0.58**	**0.59**	**0.74**
Southeast	**0.54**	0.31	0.31	**0.44**	**0.44**	**0.39**	0.17	0.25	**0.4**	**0.43**	**0.42**	**0.53**	**0.49**
Central	—	**0.6**	**0.6**	**0.47**	**0.68**	**0.52**	**0.32**	0.23	**0.54**	**0.38**	**0.42**	**0.57**	**0.73**
Upper Midwest	—	—	**0.53**	**0.35**	**0.63**	**0.4**	**0.55**	0.3	**0.37**	**0.39**	—	—	**0.80**
Plains	—	—	—	**0.46**	**0.58**	**0.65**	**0.74**	**0.32**	**0.36**	0.24	—	—	**0.86**
Northeast	—	—	—	**0.36**	**0.32**	**0.39**	**0.37**	**0.35**	0.28	—	—	—	**0.69**
Southwest	—	—	0.27	0.24	0.29	0.22	0.13	0.3	−0.01	**0.46**	—	—	**0.69**
Northwest	—	—	—	−0.05	**0.39**	**0.34**	0.1	0.0	0.21	—	—	—	0.10
West	**0.44**	0.31	**0.56**	0.17	0.21	**0.63**	**0.32**	**0.45**	—	—	—	**0.5**	**0.36**

4. Tornado frequency intensity distribution

a. Annual scaling

By construction, the TEI-EF regression model matches the CONUS occurrence frequency of each EF level. Therefore, there is a perfect match between the observed CONUS frequency–intensity scaling and that of TEI-EF (Fig. 6, top left panel). However, there is no such guarantee at the regional level (Fig. 6, second to last panels). Nonetheless, the index works particularly well for the South, Central, and Upper Midwest regions, matching the lower frequency of the EF0 class as well. TEI-EF reproduces well the rate of frequency reduction with increasing intensity (slope) that is observed in the Southeast, Northeast, and Plains, despite failing to match the frequencies themselves. For the western regions, neither the frequencies nor their slopes match observations. Overall, the exponential decay model for frequency/intensity scaling describes well the observed behavior over the EF1–EF3 range. For the Southeast and Northeast regions, the local EF2–EF3 slopes (solid lines) are different from the CONUS slope (dashed lines).

There are two possible explanations of the fact that EF0 frequencies do not follow the exponential scaling of the other intensity classes. One is that underreporting is the reason for lower EF0 frequencies (Dotzek et al. 2003). EF0 tornadoes are believed to be somewhat under reported due to their relatively short lifetimes, small damage potential (i.e., their damage can be masked by straight-line winds), and overnight occurrence (Trapp et al. 2005). If EF0 tornadoes are underreported, we could extrapolate the EF1–EF3 back to EF0 and determine the "true" number of EF0 tornadoes. Alternatively, the reduction in intensity in going from EF1 to EF0 may be less than that between other classes. In that case, by translating the observed EF0 frequency to the right, we can identify "EF0.5" (red solid circle in Fig. 6 first panel) rather than EF0 as a more representative measure of the intensity of that class of tornadoes.

Figure 7 compares our findings with previous ones, in terms of the occurrence ratios instead of slopes [with ratio = $N_{\mathrm{EF}_i}/N_{\mathrm{EF}_{i-1}} = \exp(\text{slope})$]. For the CONUS, 6 NOAA regions (the three western regions where TEI-EF performs poorly are dropped from here on), and Florida and Colorado separately, we plot the observed and TEI-EF based occurrence ratios. On the same figure we add the TEI-EF CONUS-wide ratio (black line), the two limiting behaviors discussed in Brooks and Doswell (2001, blue lines B&D), and the corresponding values from Feuerstein et al. (2005, magenta lines, see caption for details). First of all, TEI-EF ratios reproduce the observed ones. Although the TEI-EF ratios match the observed values at the CONUS-wide level by construction, their ability to reproduce the regional variability, especially for those regions such as the Northeast and Florida that deviate the most from the CONUS-wide ratio value (0.3), is driven only by the regional variability of the environments included in the TEI-EF indices. The lower value of the northeast ratio is in line with that of the NYNEX region in Feuerstein et al. (2005). The two sets of ratios (EF3/EF2 and EF2/EF1) decrease as a function of EF intensity, but less than the magenta lines from Feuerstein

TABLE 3. As in Table 2, but for TEI-EF 1.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
CONUS	**0.73**	**0.59**	**0.64**	**0.55**	**0.66**	**0.59**	**0.33**	0.31	0.04	**0.52**	**0.54**	**0.58**	**0.76**
South	**0.52**	**0.48**	**0.57**	**0.61**	**0.71**	**0.32**	**0.48**	0.03	0.14	**0.58**	**0.61**	**0.7**	**0.64**
Southeast	**0.7**	**0.46**	**0.44**	**0.67**	**0.59**	**0.42**	**0.36**	0.11	**0.41**	**0.43**	**0.41**	**0.59**	**0.53**
Central	**0.77**	**0.57**	**0.52**	**0.52**	**0.7**	**0.74**	**0.43**	0.26	**0.61**	**0.59**	**0.48**	**0.58**	**0.68**
Upper Midwest	—	—	**0.7**	0.3	**0.35**	**0.64**	**0.66**	**0.41**	**0.46**	**0.34**	—	—	**0.79**
Plains	—	—	—	**0.41**	**0.51**	**0.44**	**0.41**	0.14	—	—	—	—	**0.73**
Northeast	—	—	—	0.26	0.29	**0.43**	**0.48**	**0.51**	0.22	—	**0.61**	—	**0.67**
Southwest	—	—	—	—	0.12	0.23	0.12	0.13	−0.02	—	—	—	**0.54**
Northwest	—	—	—	—	0.19	−0.03	—	—	—	—	—	—	0.11
West	—	—	—	—	—	—	—	—	—	—	—	—	0.28

TABLE 4. As in Table 2, but for TEI-EF 2.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
CONUS	**0.7**	**0.48**	**0.39**	**0.33**	**0.41**	**0.55**	**0.41**	0.13	0.14	0.24	**0.49**	**0.55**	**0.63**
South	**0.5**	0.25	0.19	**0.47**	**0.59**	0.17	0.26	—	0.23	**0.52**	**0.62**	**0.7**	**0.55**
Southeast	**0.33**	**0.45**	**0.37**	**0.44**	**0.46**	0.22	—	—	0.29	0.19	**0.36**	**0.45**	**0.42**
Central	—	—	**0.63**	**0.37**	**0.7**	**0.7**	**0.39**	0.18	**0.45**	**0.39**	**0.35**	—	**0.58**
Upper Midwest	—	—	—	0.23	0.19	**0.47**	**0.44**	0.01	**0.41**	—	—	—	**0.64**
Plains	—	—	—	0.24	**0.46**	**0.38**	**0.61**	−0.05	—	—	—	—	**0.58**
Northeast	—	—	—	—	0.3	—	0.22	0.24	—	—	—	—	**0.43**
Southwest	—	—	—	—	0.08	−0.04	—	—	—	—	—	—	0.28
Northwest	—	—	—	—	—	—	—	—	—	—	—	—	0.08
West	—	—	—	—	—	—	—	—	—	—	—	—	0.16

et al. (2005). For Colorado, the TEI EF2/EF1 ratio (red empty triangle) is higher than the observed one, due to TEI-EF1 being lower than observed (see TEI-EF1 in Fig. 5). Finally, groups A and B display different ranges of ratios with higher values in group A and lower values in group B, consistent with previous work.

b. *Seasonal scaling*

The same frequency/intensity scaling analysis can be performed at the monthly scale. We find that the frequency of both tornado reports and TEI-EF fall off more rapidly with increasing intensity (more negative slope) during warm months than during cool months (less negative slopes; Fig. 8). At the CONUS scale (Fig. 8, first panel), the TEI-EF slopes (black) capture well the extent of the observed seasonal variability (red), with slightly poorer results for the months of June–August. At the regional level, the observed regional slopes closely follows the CONUS ones, except in the Southeast, Plains, and Southwest regions. TEI-EF slopes capture well the seasonal and regional variability of the slopes, which is large for the Southeast region, but quite small for the Plains region. The behavior of TEI-EF is driven by the environments, since the regression coefficients are constant spatially and throughout the year—all the variation comes from the environments.

c. *Explaining the yearly and seasonal scaling of EF1+ tornado counts*

Brooks and Doswell (2001) and Feuerstein et al. (2005) speculated that differences in the observed slopes/ratios were due to differences in the dominant convective modes (supercell versus nonsupercell), with smaller slopes, corresponding to higher ratios (∼−1 for slope, corresponding to 30% for ratio) in supercell dominated areas, and more negative slopes, or smaller ratios (∼−2 for slope, 10% for ratio) in regions where nonsupercell tornadoes have a greater role. Trapp et al. (2005) noted an intermediate ratio of ∼22% for QLCS tornadoes. A radar-based CONUS-wide climatology of convective modes developed by Smith et al. (2012) for tornadoes and significant severe thunderstorms provides evidence supporting this hypothesis. In particular, they found that the relative frequency of right-moving supercells is relatively high over most of the CONUS, except for Colorado and Florida (their Fig. 10). The two states that have lower ratio values have the highest percentage of disorganized (nonsupercell) events. These results are an indication that the climatological frequency of convective modes plays an important role in modulating the regional frequency–intensity scaling of tornadoes.

In a companion work to Smith et al. (2012) and Thompson et al. (2012) examined the environments of tornadoes and significant severe thunderstorms for 8 years of data (2003–11), which they classified into three groups: supercell, QLCS, and "disorganized" storms. They found that MLCAPE was a poor discriminator between storm modes except in winter when it was effective in discriminating between QLSC and supercell. Also they found no meaningful dependence of EF-scale on MLCAPE in either mode category. Thompson et al. (2012) also found that measures of bulk wind shear and storm relative helicity were not very effective in separating supercell and

TABLE 5. As in Table 2, but for TEI-EF 3.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
CONUS	**0.4**	**0.34**	**0.37**	0.31	**0.4**	**0.59**	**0.42**	0.17	−0.08	0.11	0.31	**0.39**	**0.53**
South	—	—	0.15	0.26	**0.54**	**0.43**	—	—	—	—	**0.38**	—	**0.51**
Southeast	—	—	0.14	0.16	—	—	—	—	—	—	**0.43**	—	**0.32**
Central	—	—	**0.5**	**0.33**	**0.4**	—	—	—	—	—	0.18	—	**0.40**
Upper Midwest	—	—	—	—	−0.02	0.2	—	—	—	—	—	—	**0.36**
Plains	—	—	—	—	—	0.18	—	—	—	—	—	—	**0.39**
Northeast	—	—	—	—	—	—	—	—	—	—	—	—	0.22
Southwest	—	—	—	—	—	—	—	—	—	—	—	—	—
Northwest	—	—	—	—	—	—	—	—	—	—	—	—	—
West	—	—	—	—	—	—	—	—	—	—	—	—	—

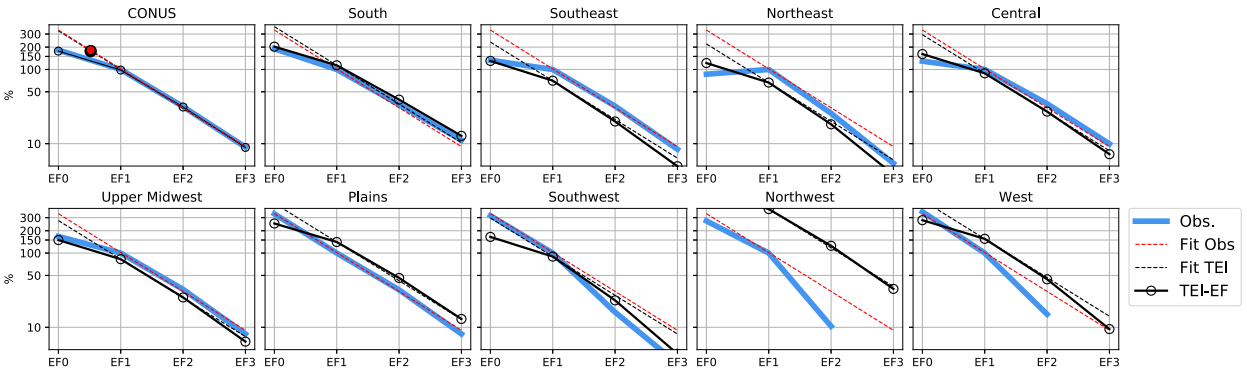

FIG. 6. Annual totals of tornado reports (blue solid lines) and TEI-EF values (black circle and line) at the (top left) CONUS and (remaining panels) regional level as a function of EF intensity over the 1979–2016 period. Reports have been normalized to 100% EF1 tornadoes for each panel. The red dashed line is a weighted log-linear fit of the EF1–EF3 CONUS values, and the black dashed line is fit to the CONUS TEI-EF values; both slopes are used to draw the dashed lines in the regional plots (anchored to each regional observed EF1 or TEI-EF1). The red and black solid circles, which closely overlap, identify the EF class for which the observed EF0 counts align on the (log) linear fit, and is equal to 0.517.

QLCS cases, though values were somewhat lower for QLCS cases. On the other hand, they were effective in separating organized (supercell and QLCS) and disorganized storms, with lower values for the disorganized types. The fact that bulk wind shear and storm-relative helicity ranges for QLCS tornadoes were intermediate those for supercell and disorganized storms would suggest that their frequency–intensity scaling would also be intermediate, which is consistent with Trapp et al. (2005) who found an occurrence ratio of 22% QLCS tornadoes.

If climatological frequency of convective modes is a key factor, then the question arises of how the TEI-EF represents fairly well both the regional and seasonal variation in tornado frequency–intensity scaling using its two environmental quantities, cPrcp and SRH3. One clue is that, SRH3 is widely accepted in the operational forecasting community as a supercell and tornado forecast parameter (Colquhoun and Riley 1996; Rasmussen and Blanchard 1998; Thompson et al. 2002; Davies-Jones 2015, and references therein). There is evidence that a minimum amount of SRH3 needs to be present for most of the lifetime of a convective event to maintain supercell storms, and that environments characterized by large SRH3 are the ones that generate long-lived storms (Droegemeier et al. 1993; Colquhoun and Riley 1996). Moreover, as mentioned above, SRH has shown to be effective in discriminating between organized (supercell and QLCS) versus disorganized storms. Because of all this, it is plausible that the monthly averages of SRH3 that appear in TEI-EF may provide some information about the relative climatological frequencies of supercell and nonsupercell (disorganized) dominated processes.

To determine how much of the observed scaling variability is directly linked to the regional and temporal variability of SRH3, we first compute the frequency/intensity scaling for a TEI-EF model in which we artificially fix the cPrcp coefficients to their EF0 value. The resulting slopes (cPrcp-fixed, green, Fig. 8) are nearly identical to those of the full TEI-EF model (black, Fig. 8). This result is somewhat expected because β_{cPrcp} is nearly constant across EF classes (Fig. 2). Therefore, we conclude that nearly all the scaling variability in TEI-EF results from its varying sensitivity to SRH3. To make this statement precise, consider the TEI-EF occurrence ratio between two EF classes, and take the corresponding cPrcp coefficients to be equal. In that case, we have

$$\text{Occurrence ratio} = \exp[\beta_0' - \beta_0 + (\beta_{\mathrm{SRH3}} - \beta_{\mathrm{SRH3}}') \log(\mathrm{SRH3})], \quad (3)$$

where the prime indicates the coefficients for the higher intensity. First, we see that the occurrence ratio depends on the difference $\beta_0 - \beta_0'$, which does not vary by month or location, and that there is no dependence on cPrcp because its coefficient is independent of EF level. Second, the only variation by location and month in the occurrence ratio is through

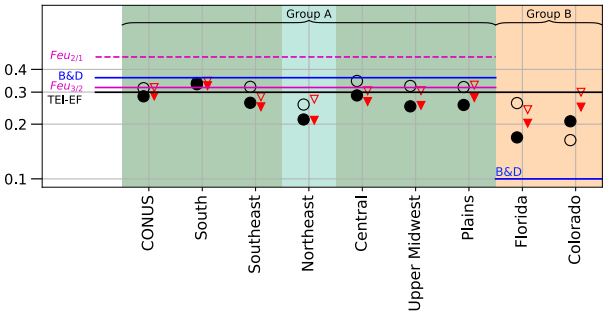

FIG. 7. Observed (black circles) and TEI-EF based (red triangles) occurrence ratios of annual totals $N_{\mathrm{EF}_i}/N_{\mathrm{EF}_{i-1}}$ (EF2/EF1: empty markers: EF3/EF2: solid markers) for CONUS, six NOAA regions, Florida, and Colorado, divided into groups A (supercell dominated) and B (not supercell dominated) as in Feuerstein et al. (2005). The black solid line is our TEI-EF CONUS ratio (0.3), the two blue lines (B&D) report the two limiting behaviors (0.36 and 0.1) found in Brooks and Doswell (2001) for supercell and nonsupercell regions, respectively, and the magenta lines show the EF2/EF1 (dashed; 0.468) and EF3/EF2 (solid; 0.318) ratios reported in Fig. 6 of Feuerstein et al. (2005).

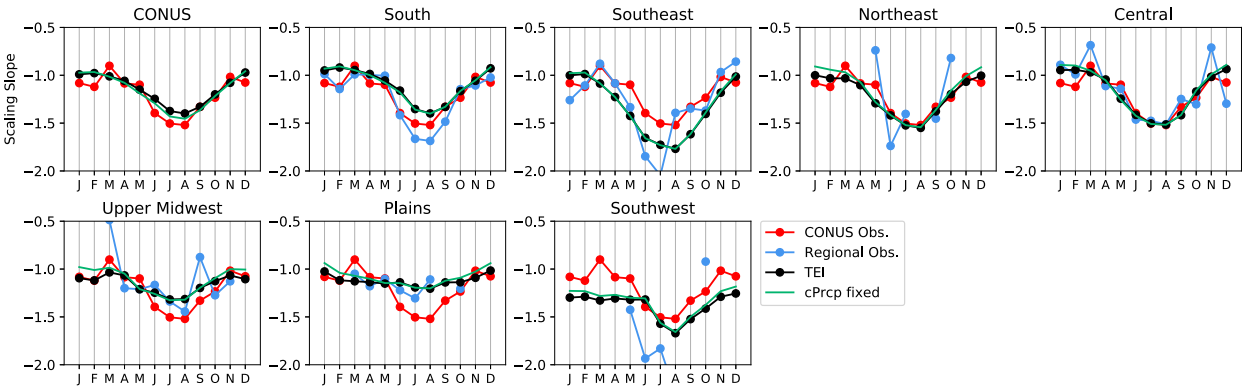

FIG. 8. Seasonal variability of the slope of weighted log-linear fits for monthly EF1–EF3 counts for (top left) CONUS-wide totals and (remaining panels) seven NOAA regions. CONUS observed seasonal cycle is in red, and reported in the regional panels for reference; regional observed values are in blue; and TEI-EF based ones in black. cPrcp fixed in green refers to TEI-EF indices obtained by fixing cPrcp coefficient to be equal for all EF intensities.

SRH3 and the term $(\beta_{\mathrm{SRH3}} - \beta'_{\mathrm{SRH3}})\log(\mathrm{SRH3})$. Moreover, since the sensitivity to SRH3 increases with intensity (Fig. 2), $(\beta_{\mathrm{SRH3}} - \beta'_{\mathrm{SRH3}})$ is negative, which means that the occurrence ratio is larger (smaller decrease in frequency) where and when SRH3 tends to be larger (i.e., during cool months and where supercells are more frequent). Therefore, the regional and seasonal differences in the occurrence slopes/ratios of TEI-EF in Figs. 8 and 7 are essentially driven by SRH3.

To further illustrate the role of SRH3 in setting the frequency–intensity scaling of TEI-EF, we constructed two sets of idealized annual cycles of cPrcp and SRH3 by spatially averaging over the CONUS and over Florida (Fig. 9 left panel).When TEI-EF is applied to the idealized annual cycles, the resulting slopes (Fig. 9, right panel) are essentially identical to the monthly varying and annual CONUS slopes obtained from TEI-EF (in black), and also match well the observed values (in red). In particular, for the CONUS, both the monthly (circle and line) and the annual values (solid lines) are not significantly different, and so for the annual slopes in Florida (dashed lines). The January–March (JFM) and May–August (MJJA) results for Florida (blue and black empty circles and squares), although not as good in reproducing the observed values (red empty circles and squares), capture the shift in slope values, from flatter slopes in JFM to steeper ones in MJJA. The only difference between CONUS and Florida TEI-EF* results is once again the different seasonal cycle of the environments, presented in the left panel. Although cPrcp seasonality is strong, the seasonal changes in tornado counts due to this environment will be almost identical among all EF classes, because the coefficients are essentially constant with EF. Instead, the different sensitivity of each EF class to SRH3 (as indicated by the parameters in Fig. 2), causes the model for higher intensity tornadoes (TEI-EF3, with higher sensitivity to SRH3) to increase more relatively to lower intensities (TEI-EF1) for any unit increase in SRH3; this leads to a less steep slope in cold months, when SRH3 is highest, and a steeper slopes in warm months, when SRH3 is lowest.

5. Summary and discussion

We have extended a tornado environment index (TEI; Tippett et al. 2012, 2014), which was designed to predict the total number of tornadoes from monthly environments, to the

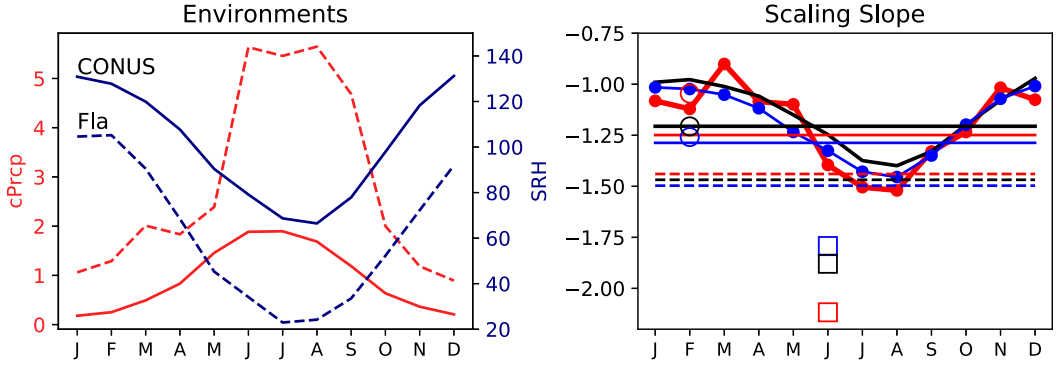

FIG. 9. (left) cPrcp and SRH3 climatological CONUS-wide (solid lines) and Florida only (dashed lines) monthly values. (right) The observed (red), TEI-EF (black), and TEI-EF* (blue; see text for details) slope values for tornado counts taken from CONUS monthly (circle and line), and annual (solid lines) data, Florida annual (dashed lines), JFM (empty circles), and MJJA (empty squares) data.

EF-dependent TEI (TEI-EF), which predicts separately the number of tornadoes rated EF0, EF1, EF2, and EF3. After a model selection procedure, the same environmental parameters used in TEI, convective precipitation (cPrcp) and 0–3-km storm relative helicity (SRH3), were found to give a good performing and parsimonious model. The TEI-EF reproduces the seasonal and annual variability of tornado numbers by EF rating, with correlation values (as shown in Tables 2–5 and Fig. 4) that are not far from those obtained by the original TEI (Lepore et al. 2018).

The scaling of tornado frequency with intensity shows a log-linear relationship that varies regionally and seasonally. The regional scaling variability (Brooks and Doswell 2001; Feuerstein et al. 2005; Trapp et al. 2005) has been explained as the result of differing frequency of convective modes, with regions that are mostly dominated by supercell tornadic occurrences displaying higher occurrence ratios (lower slopes), versus regions that are dominated by disorganized (nonsupercell) severe storms with lower occurrence ratios (higher slopes) (Brooks and Doswell 2001; Feuerstein et al. 2005; Trapp et al. 2005). The seasonal variability of the slope, and its regional features, is instead an interesting new finding. In particular, we observe steeper slopes during warm seasons, and flatter slopes for colder season once again in line with the observed seasonal distribution of supercell versus disorganized convective modes (Fig. 12 in Smith et al. 2012). This seasonally dependent scaling means that the fraction of high-intensity tornadoes is greater during cool months than warm months.

TEI-EF well reproduces the scaling at both the spatial and seasonal scale, and it allows us to explore which large-scale environment is responsible for its variability. In fact, based on the widely accepted use of SRH3, one of the two parameters used in our index, as a proxy for supercell tornadogenesis, we are able to give a more support to the hypothesis that convective modes are the driver of the scaling variability. In particular, our findings confirm that, at both regional and seasonal levels, scaling variability depends primarily on SRH3 variability, with occurrence ratios that are higher (corresponding to lower slopes) for regions and seasons where SRH3 is higher, and the opposite for regions and seasons where the SRH3 is lower.

A possible future application of TEI-EF and similar indices would be to explore the extent to which in the context of seasonal forecasting (Lepore et al. 2018) numbers of tornadoes by intensity can be predicted. Our analysis has found that the model has some clear weaknesses, such as the relatively poor fit for the late summer and fall months. This results in missing the secondary peak seen in November for EF1+ classes (South and Southeast regions Fig. 4) and the lower correlation values in August–October (as seen in Tables 2–5). Similar issues were noted for TEI (Tippett et al. 2014; Lepore et al. 2018) and are a likely consequence of the simplicity of the model, which uses the same parameters in all locations and seasons. A related issue is the fact that months that display similar averages of the environments considered (i.e., March vs November or May vs September; see for example Fig. 9 left panel), can have quite different distributions at daily resolutions, and this difference is lost in the TEI and TEI-EF. Moreover, events in the cool months (particularly in the Southeast region) can be dominated by QLCS tornadoes (rather than supercell) whose dependence on SRH has been shown to be different from that of supercells (Trapp et al. 2005; Thompson et al. 2012). Directions for future model improvement would be to incorporate sub-monthly environmental statistics or explicit seasonality.

Acknowledgments. This work was supported by the Willis Research Network. The authors thank two anonymous reviewers and Jeff Trapp for their careful reading, insightful comments and helpful suggestions. Severe Weather Database tornado and hail reports are available from the Storm Events Database at http://www.spc.noaa.gov/wcm/#data. NARR data are downloaded from the Research Data Archive (RDA) at the National Center for Atmospheric Research (NCAR), Computational and Information Systems Laboratory at http://rda.ucar.edu/datasets/ds608.0/. All analysis are performed in python using: NumPy (van der Walt et al. 2011), xarray (Hoyer and Hamman 2017), pandas (McKinney 2010; pandas development team 2020), statsmodels (Seabold and Perktold 2010), SciPy (Virtanen et al. 2020), and matplotlib (Hunter 2007).

REFERENCES

Allen, J. T., M. K. Tippett, 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.

Ashley, W. S., and S. M. Strader, 2016: Recipe for disaster: How the dynamic ingredients of risk and exposure are changing the tornado disaster landscape. *Bull. Amer. Meteor. Soc.*, **97**, 767–786, https://doi.org/10.1175/BAMS-D-15-00150.1.

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.

Brooks, H. E., 2013: Severe thunderstorms and climate change. *Atmos. Res.*, **123**, 129–138, https://doi.org/10.1016/j.atmosres.2012.04.002.

——, and C. A. Doswell III, 2001: Some aspects of the international climatology of tornadoes by damage classification. *Atmos. Res.*, **56**, 191–201, https://doi.org/10.1016/S0169-8095(00)00098-3.

——, ——, 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.

Cameron, A. C., and F. A. Windmeijer, 1996: R-squared measures for count data regression models with applications to health-care utilization. *J. Bus. Econ. Stat.*, **14**, 209–220, https://doi.org/10.1080/07350015.1996.10524648.

Coffer, B. E., M. D. Parker, R. L. Thompson, B. T. Smith, and R. E. Jewell, 2019: Using near-ground storm relative helicity in supercell tornado forecasting. *Wea. Forecasting*, **34**, 1417–1435, https://doi.org/10.1175/WAF-D-19-0115.1.

Colquhoun, J. R., and P. A. Riley, 1996: Relationships between tornado intensity and various wind and thermodynamic variables. *Wea. Forecasting*, **11**, 360–371, https://doi.org/10.1175/1520-0434(1996)011<0360:RBTIAV>2.0.CO;2.

Davies-Jones, R., 2015: A review of supercell and tornado dynamics. *Atmos. Res.*, **158–159**, 274–291, https://doi.org/10.1016/j.atmosres.2014.04.007.

Diffenbaugh, N. S., M. Scherer, and R. J. Trapp, 2013: Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proc. Natl. Acad. Sci. USA*, **110**, 16 361–16 366, https://doi.org/10.1073/pnas.1307758110.

Doswell, C. A., III, H. E. Brooks, and N. Dotzek, 2009: On the implementation of the enhanced Fujita scale in the USA. *Atmos. Res.*, **93**, 554–563, https://doi.org/10.1016/j.atmosres.2008.11.003.

Dotzek, N., J. Grieser, and H. E. Brooks, 2003: Statistical modeling of tornado intensity distributions. *Atmos. Res.*, **67–68**, 163–187, https://doi.org/10.1016/S0169-8095(03)00050-4.

Droegemeier, K. K., S. M. Lazarus, and R. Davies-Jones, 1993: The influence of helicity on numerically simulated convective storms. *Mon. Wea. Rev.*, **121**, 2005–2029, https://doi.org/10.1175/1520-0493(1993)121<2005:TIOHON>2.0.CO;2.

Feuerstein, B., N. Dotzek, and J. Grieser, 2005: Assessing a tornado climatology from global tornado intensity distributions. *J. Climate*, **18**, 585–596, https://doi.org/10.1175/JCLI-3285.1.

Fujita, T. T., 1981: Tornadoes and downbursts in the context of generalized planetary scales. *J. Atmos. Sci.*, **38**, 1511–1534, https://doi.org/10.1175/1520-0469(1981)038<1511:TADITC>2.0.CO;2.

Gensini, V. A., and M. K. 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.

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.

Hoyer, S., and J. Hamman, 2017: xarray: N-D labeled arrays and datasets in Python. *J. Open Res. Software*, **5**, 10, https://doi.org/10.5334/jors.148.

Hunter, J. D., 2007: Matplotlib: A 2d graphics environment. *Comput. Sci. Eng.*, **9**, 90–95, https://doi.org/10.1109/MCSE.2007.55.

Krocak, M. J., and H. E. Brooks, 2018: Climatological estimates of hourly tornado probability for the United States. *Wea. Forecasting*, **33**, 59–69, https://doi.org/10.1175/WAF-D-17-0123.1.

Lepore, C., M. K. Tippett, and J. T. Allen, 2018: CFSv2 monthly forecasts of tornado and hail activity. *Wea. Forecasting*, **33**, 1283–1297, https://doi.org/10.1175/WAF-D-18-0054.1.

McKinney, W., 2010: Data structures for statistical computing in Python. *Proc. Ninth Python in Science Conf.*, (SciPy 2010), Austin, TX, SciPy.org, https://doi.org/10.25080/Majora-92bf1922-00a.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, https://doi.org/10.1175/BAMS-87-3-343.

pandas development team, 2020: pandas-dev/pandas: Pandas. Zenodo, accessed 23 October 2020, https://doi.org/10.5281/zenodo.3509134.

Rasmussen, E. N., and D. O. Blanchard, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting*, **13**, 1148–1164, https://doi.org/10.1175/1520-0434(1998)013<1148:ABCOSD>2.0.CO;2.

Seabold, S., and J. Perktold, 2010: Statsmodels: Econometric and statistical modeling with Python. *Ninth Python in Science Conf. (SciPy 2010)*, Austin, TX, SciPy.org, https://doi.org/10.25080/Majora-92bf1922-011 .

Smith, B. T., R. L. Thompson, J. S. Grams, C. Broyles, and H. E. Brooks, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Wea. Forecasting*, **27**, 1114–1135, https://doi.org/10.1175/WAF-D-11-00115.1.

Thompson, R. L., R. Edwards, and J. A. Hart, 2002: Evaluation and interpretation of the supercell composite and significant tornado parameters at the Storm Prediction Center. *21st Conf. on Severe Local Storms/19th Conf. in Weather Analysis and Forecasting/15th Conf. on Numerical Weather Prediction*, San Antonio, TX, Amer. Meteor. Soc., J3.2, https://ams.confex.com/ams/SLS_WAF_NWP/techprogram/paper_46942.htm.

——, B. T. Smith, J. S. Grams, A. R. Dean, and C. Broyles, 2012: Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Wea. Forecasting*, **27**, 1136–1154, https://doi.org/10.1175/WAF-D-11-00116.1.

Tippett, M. K., 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.

——, ——, ——, and J. T. Allen, 2014: An empirical relation between U.S. tornado activity and monthly environmental parameters. *J. Climate*, **27**, 2983–2999, https://doi.org/10.1175/JCLI-D-13-00345.1.

Trapp, R. J., S. A. Tessendorf, E. S. Godfrey, and H. E. Brooks, 2005: Tornadoes from squall lines and bow echoes. Part I: Climatological distribution. *Wea. Forecasting*, **20**, 23–34, https://doi.org/10.1175/WAF-835.1.

——, N. S. Diffenbaugh, and A. Gluhovsky, 2009: Transient response of severe thunderstorm forcing to elevated greenhouse gas concentrations. *Geophys. Res. Lett.*, **36**, L01703, https://doi.org/10.1029/2008GL036203.

van der Walt, S., S. C. Colbert, and G. Varoquaux, 2011: The NumPy array: A structure for efficient numerical computation. *Comput. Sci. Eng.*, **13**, 22–30, https://doi.org/10.1109/MCSE.2011.37.

Virtanen, P., and Coauthors, 2020: SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods*, **17**, 261–272, https://doi.org/10.1038/s41592-019-0686-2.