INTERCOMPARISON OF SPATIOTEMPORAL VARIABILITY IN SEVERE WEATHER ENVIRONMENTAL PROXIES AND TORNADO ACTIVITY OVER THE UNITED STATES

A Thesis

Submitted to the Faculty

of

Purdue University

by

Shawn W. Simmons

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

December 2018

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF THESIS APPROVAL

Dr. Ernest M. Agee, Chair
Department of Earth, Atmospheric, and Planetary Sciences
Dr. Daniel R. Chavas
Department of Earth, Atmospheric, and Planetary Sciences
Dr. Michael Baldwin
Department of Earth, Atmospheric, and Planetary Sciences

Approved by:
Dr. Darryl E. Granger
Head of the School Graduate Program

ACKNOWLEDGMENTS

I thank my adviser, Dr. Dan Chavas, for all of the guidance and support he gave me throughout my time at Purdue University. I thank Dr. Kim Hoogewind for performing data calculations on the North American Regional Reanalysis output. Finally, I thank all of my friends and family for the love and support they have given to me during my time at Purdue University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

ABSTRACT

Simmons, Shawn W. M.S., Purdue University, December 2018. Intercomparison of spatiotemporal variability in severe weather environmental proxies and tornado activity over the United States. Major Professor: Daniel R. Chavas.

Tornadoes cause numerous deaths and significant property damage each year, yet how tornado activity varies across climate states, particularly under global warming, remains poorly understood. Importantly, severe weather events arise during transient periods of extreme thermodynamic environments whose variability may differ from that of the environmental mean state. This study analyzes the climatological relationships between commonly-used severe weather environmental proxies (the product of convective available potential energy and bulk vertical wind shear, energy-helicity index, and the significant tornado parameter) and tornado density on three dominant timescales of climate forcing: diurnal, seasonal, and interannual. We utilize reanalysis data to calculate the spatial distributions of the mean, median, and a range of extreme percentiles of these proxies across each timescale as well as for the full climatology. We then test the extent to which each measure captures the spatiotemporal variability of tornado density over the continental United States. Results indicate that the mean is a suitable statistic when used with the full climatology of the energy-helicity index and the significant tornado parameter without using convective inhibition in calculations, the diurnal cycle for convective available potential energy and the product of convective available potential energy and bulk vertical wind shear, and the interannual variations for all proxies except convective available potential energy. The mean is outperformed by extreme percentiles otherwise. This understanding of climatological relationships between tornadoes and the large scale environments can improve prediction of tornado frequency and provides a foundation

for understanding how changes in the statistics of large-scale environments may affect tornado activity in a future warmer climate state.

CHAPTER 1. INTRODUCTION

Severe weather is a hazard common to the continental United States. A thunderstorm is classified as severe if wind speeds exceed 58 miles per hour, hail size exceeds one inch in diameter, and/or a tornado forms [1]. These storms can cause significant damage and loss of life. Damage can arise from, but is not limited to, flooding, wind downing power lines and trees, lightning strikes starting fires, and tornadoes destroying structures. In 2015, tornadoes caused 320.42 million dollars of damage, killed 36, and injured 924 people [2]. Between the years 1991 to 2010, the annual average count of all (EF0+) tornadoes was 1253 [3].

Mesoscale convective storms come in multiple modes of organization. Classifications are typically broken down by severity, longevity, storm shape, intensity, as well as other factors [4]. Multiple studies have worked to classify convective storms by morphology [5,6]. The three general types of storm modes (single-cell, multi-cell, and supercell) have been broken down further. Smith et al. (2012) [5] used five major convective types: Quasi-Linear Convective System (QLCS), Linear Hybrid, Supercell, Marginal, and Disorganized, and these can be further broken down based on level of organization. Schoen and Ashley 2011 [7] found that deadly convective wind events can come from multiple storm types, but most fatal tornadoes were caused by supercells. The meteorological community has also recognized supercells as the storm type that most commonly produces tornadoes [8] and supercells also include most of the strong and violent tornadoes [9]. Forecasting convective modes proves to be difficult. Weissman and Klemp (1982) [4] showed that vertical wind shear and buoyancy can determine storm mode. However, the storm mode may change as the storm develops and evolves. The storm mode can be impacted and altered by the number of storms initiated, ability of the storms to organize, ability to continually initiate convection, and the type of boundary at which the storms initiate [10].

Mesoscale processes are significant to supercell and tornado development, both for preconditioning of the environment prior to an event as well as during the event itself [10]. For deep, moist convection to occur, a source of lifting must be present, whether it is convection from solar heating, orographic convection from mountains, frontal lifting along either a temperature (cold fronts) or moisture (drylines) gradient, or convergent flows of air near the surface [11]. These mesoscale processes are difficult to forecast. Though these boundaries can be easily identified in observations, convection rarely occurs along the entirety of the boundary, but rather, along segments of the boundary [11]. Detailed information on the full vertical structure of the local thermodynamic environment in the area of interest is needed in order to predict the spatiotemporal evolution of convective storms and their modes of organization. While nearby radiosondes provide some useful data toward this end, soundings are at discrete points in space that may not be representative of the full environment. The environment may also further evolve after a sounding was initially launched, as Beebe (1958) [12] showed that the vertical structure of the environment can distinctly change from a sounding that precedes the convective environment by multiple hours to a proximity sounding.

At the synoptic scale, key environmental ingredients necessary for severe weather have been identified. Such ingredients are easier to predict than the mesoscale forcings as they represent the thermodynamic and dynamic environment prior to any storm initiation. These are convective available potential energy [13] and lower tropospheric wind shear [11]. Two other key environmental ingredients that can give further insight into a severe storm are storm relative helicity and the lifting condensation level [14]. Convective available potential energy is the amount of energy that could be released if an air parcel is lifted to its equilibrium level [15]. Lower tropospheric wind shear is the magnitude of the vector difference between wind velocity at two atmospheric layers [11]. Storm relative helicity is a measure of horizontal vorticity in the storm relative reference frame [11]. The lifting condensation level is the level of the cloudbase [15]. Detailed descriptions of these parameters are provided below in Section 1.1.

Finally, certain *proxies* have been developed that combine the fundamental thermodynamic and dynamic parameters and are used specifically to define environments conducive to severe weather activity. These are the product of CAPE and lower-level tropospheric wind shear (CS from this point forward), Energy-Helicity Index (EHI), and the Significant Tornado Parameter (STP) [14, 16]. These are formulated from combinations of thermodynamic and wind shear parameters that favor supercell and tornado formation. The thermodynamic parameters are CAPE and the Lifting Condensation Level (LCL). The wind shear parameters are Storm-Relative Helicity (SRH) and vertical wind shear. CAPE and vertical wind shear are combined to calculate CS. CAPE and SRH are combined to calculate EHI. STP is a combination of all four of the parameters.

1.1 Parameters and Proxies

1.1.1 Thermodynamic Parameters

Convective available potential energy (CAPE)

CAPE is a measure of the potential for unstable buoyant acceleration of low-level air parcels displaced upwards within the atmosphere, typically associated with the combination of warm, moist boundary-layer air and steep lapse rates in the overlying free troposphere. Such unstable ascent is essential to deep moist convection that is associated with thunderstorms and lightning [13]. During convection, CAPE is converted to the kinetic energy of the ascending parcel, the majority of which is typically derived from the release of latent heat through the condensation of water vapor [17]. CAPE is defined as

$$CAPE = \int_{z_{LFC}}^{z_{EL}} B dz \approx -g \int_{z_{LFC}}^{z_{EL}} \frac{T_v'}{\overline{\overline{T}}_v} dz \tag{1.1}$$

where z_{LFC} is the level of free convection, z_{EL} is the equilibrium level, B is the buoyancy of the lifted parcel, g is the gravitational constant, T_v' is the virtual temperature perturbation of the air parcel, and $\overline{T}_v$ is the virtual temperature of the environment [11]. The value for the gravitational constant is provided in Table 1.1. By this equation, CAPE is the vertically integrated buoyancy from the level of free convection to the equilibrium level when the parcel is lifted pseudo-adiabatically. CAPE can be used to give an estimate for an upper bound of updraft strength by performing a calculation from potential energy into kinetic energy ($w_{max} = \sqrt{2 * CAPE}$) [11, 18, 19].

Table 1.1.
Constant values used in equations

Constant	**Value**
g	9.81 ms^{-2}
c_{pv}	1879 J kg^{-1} K^{-1}
R_v	461 J kg^{-1} K^{-1}
T_{trip}	273.16 K
E_{0v}	2.3740 x 10^6 J kg^{-1}
c_{vv}	1418 J kg^{-1} K^{-1}
c_{vl}	4119 J kg^{-1} K^{-1}

Lifting Condensation Level (LCL)

The LCL is an estimation of the height of the cloud base and occurs at the altitude where rising air reaches saturation and condensation begins to form clouds [15]. Espy (1836) [20] gave the first equation that was used to approximate LCL. Espy's equation is given by

$$z_{LCL} = (125m/K)(T - T_d) \tag{1.2}$$

where z_{LCL} is the LCL height in meters, T is the initial temperature of the rising air parcel, and T_d is the environmental dewpoint temperature. Further developments to an equation for LCL have been made and Romps (2017) [21] came to an exact equation of

$$z_{LCL} = \frac{c_{pm}}{g}(T - T_{LCL}) \tag{1.3}$$

where z_{LCL} is the LCL height in meters, c_{pm} is the specific heat at constant pressure for the air parcel, g is the gravitational constant, T is the air parcel's initial temperature, and T_{LCL} is the temperature at the LCL. T_{LCL} is given by the expression

$$T_{LCL} = c[W_{-1}(RH_l^{1/a} c e^c)]^{-1} T \tag{1.4}$$

where

$$a = \frac{c_{pm}}{R_m} + \frac{c_{vl} - c_{pv}}{R_v} \tag{1.5}$$

$$b = \frac{-E_{0v} - (c_{vv} - c_{vl})T_{trip}}{R_v T} \tag{1.6}$$

$$c = \frac{b}{a} \tag{1.7}$$

W_{-1} is the -1 branch of the Lambert W function, RH_l is the air parcel's initial relative humidity, c_{pm} is the air parcel's specific heat at constant pressure, R_m is the air parcel's specific gas constant, c_{vl} is the specific heat of liquid water, c_{pv} is the specific heat of water vapor at constant pressure, R_v is the specific gas constant for water vapor, E_{0v} is the difference in specific internal energy between water vapor and liquid water at the triple point, c_{vv} is the specific heat of water vapor at constant volume, and T_{trip} is the temperature at the triple point of water. Values for the constant parameters are provided in Table 1.1 Lower LCL values allow for less sub-cloud evaporation, which may reduce the potential for the mesocyclone to be undercut and weakened by cold outflow [14]. Rasmussen and Blanchard [14] found that the LCL height gave the most utility to distinguish between significant tornadoes and supercells that produced either weak tornadoes or non-tornadic supercells.

1.1.2 Dynamical Parameters

Vertical Shear

Bulk vertical wind shear (V_{shear}) is the absolute value of the difference in wind velocity vectors between the top and the bottom of the layer of interest [11].

$$V_{shear} = |\vec{u}_{top} - \vec{u}_{bot}| \tag{1.8}$$

where $\vec{u}_{top}$ is the wind velocity vector at the top of the atmospheric layer and $\vec{u}_{bot}$ is the wind velocity vector at the bottom of the atmospheric layer [11]. 0-6 kilometer V_{shear} is the most commonly used measure due to its ability to predict storm type and longevity [4]. Vertical wind shear is necessary for supercells as it prolongs the lifetime of a thunderstorm by tilting the updraft, however too strong of V_{shear} can can inhibit convection in areas of weak instability by increasing entrainment [11].

Storm Relative Helicity (SRH)

SRH is a measure of the streamwise horizontal vorticity available to feed into a storm's updraft and cause rotation [11]. Here, SRH is defined as the vertically-integrated dot product of the horizontal vorticity and the mean horizontal flow velocity in a reference frame moving with the storm, integrated from the bottom of an atmospheric layer to the top of the layer of interest:

$$SRH = \int_{z_0}^{z_t} (\overline{v} - c) \cdot \overline{\omega}_h dz \tag{1.9}$$

where z_0 is the height of the bottom of the atmospheric layer, z_t is the height of the top of the atmospheric layer, $\overline{v}$ is the environmental wind, c is the storm motion, and $\overline{\omega}_h$ is the horizontal vorticity [11].

1.1.3 Severe Weather Proxies

Product of CAPE and V_{shear} (CS)

CS is commonly used as a proxy for severe weather that combines the most essential thermodynamic (CAPE) and dynamic (shear) environmental ingredients for rotating convection. Weisman and Klemp (1982) [4] showed the dependence of storm structure on environmental wind shear and buoyancy, where under the same buoyancy level, low shear environments produced single-cell thunderstorms, moderate shear produced multi-cell thunderstorms, and high shear environments produced supercell structured thunderstorms. This has further been shown to be capable of distinguishing between significant and less severe thunderstorms [22, 23]. Here, equal weighting was given to both CAPE and V_{shear}.

$$CS = CAPE * V_{shear} \tag{1.10}$$

Observational studies have argued that V_{shear} may be more important in discriminating the severity of a storm [19, 24]. However, Seeley and Romps (2015) [25] suggests that weight given to V_{shear} is not the dominant source of uncertainty in predictions of the severity of future weather.

Energy Helicity Index (EHI)

EHI is an alternative severe weather proxy that similarly combines one thermodynamic parameter (CAPE) and one dynamic parameter (SRH). EHI is given by:

$$EHI = (CAPE * SRH)/160000 \tag{1.11}$$

, and was first developed by Hart and Korotky (1991) [26]. In forecasting, values over 1 indicate a potential for supercells and values over 2 show a high probability of supercell formation [14]. EHI can be used to differentiate between tornadic and non-tornadic supercells and as values of EHI become larger, the likelihood of tornadoes increases significantly [14, 27, 28].

Significant Tornado Parameter (STP)

STP is a non-dimensional quantity that combines all of the aforementioned environmental parameters: CAPE, vertical wind shear, SRH, and LCL. Here, the equation used for STP follows a fixed-layer approach [28],

$$STP = \left(\frac{CAPE}{1500}\right) * \left(\frac{V_{shear}}{20}\right) * \left(\frac{SRH}{150}\right) * \left(\frac{2000 - LCL}{1000}\right) \tag{1.12}$$

STP has been shown to have the ability to differentiate between supercell and non-supercell storms as well as between significant tornadic supercells and non-tornadic supercells [16]. STP approaches zero for small values of CAPE, vertical shear, and SRH along with LCL values higher than 2000 meters. In forecasting, STP values greater than 1 are associated with supercell storms that are able to produce a tornado, while non-tornadic and non-supercell storms are associated with values less than 1. STP has several alternative forms. In 2004, Thompson et al [29] released a modification to the STP equation. This new modification added the alternative of using the effective layer instead of a fixed layer, and it added a convective inhibition (CIN) term. Later, limits and caps were put on the STP equation. These bounds are listed in Thompson et al (2012) [30] and are such that the effective bulk vertical shear (EBWD) term would cap at 1.5 if EBWD is greater than 30 m/s and would be set to zero if EBWD is lower than 12.5 m/s, the LCL term is set to 1 if LCL is less than 1000 m and set to 0 if the LCL is greater than 2000 m, and the CIN term is set to 1 if CIN is greater than -50 J/kg and set to 0 if CIN lower than -200 J/kg.

1.1.4 Spatiotemporal Variability

Temporal Variability

Tornado activity is known to exhibit significant temporal variability, particularly on the diurnal and seasonal timescales. Studies have shown that increasing temperatures have led to changes in frequency of tornadoes in interannual trends. In the diurnal cycle, most tornadoes occur in the late afternoon to early evening, local mean solar time (LST) [31]. In the seasonal cycle, the spring (March, April, May) and summer (June, July, August) months have the most tornado occurrences, with the peak months being May and June [32]. Brooks et al. (2014) [33] and Elsner et al. (2015) [34] show that there has also been a shift in tornado occurrence in that there are less days with one tornado, but an increase in the number of days with multiple tornadoes.

Similarly, the environmental proxies for severe weather potential also exhibit significant temporal variability on diurnal, seasonal, and interannual timescales. CAPE is known to attain peak values during spring and summer [30] as well as during afternoon and early evening [35]. Vertical wind shear, both directional (SRH) and speed (V_{shear}), varies strongly seasonally, with greater values occurring in winter and the transition seasons of spring and fall [30].

Spatial Variability

The central United States is the peak area of favorable severe weather parameters in the continental United States [22]. This region is favorable for CAPE build-up as elevated land to the west (Rocky Mountains and Mexican Plateau) provides elevated dry air to act as a capping inversion, or "cap," to inhibit convection, and this cap allows for steep lapse rates and for CAPE to build throughout the day [11, 36, 37]. Moisture is supplied to the air below from the Gulf of Mexico, soil evaporation, and evapotranspiration from crops. Southerly winds from the Gulf in conjunction with easterly flow over the Rockies result in strong vertical shear [22].

Tornado activity exhibits spatial variability across the continental United States. Tornado trends vary spatially seasonally and interannually. The most pronounced spatial variability in tornado activity is its northward progression towards the northern Great Plains during the spring and summer months [32, 38]. During the cold season (Fall and Winter), tornadoes are more common to the southeastern United States and typically, are accompanied by high vertical shear, low CAPE environments [39].

The southeast shows a bimodal distribution in tornado activity with elevated tornado occurrences in the spring and in the fall/early winter while the Great Plains show a unimodal distribution with elevated tornado occurrences in spring [40].

1.1.5 Long-Term Trends and Future Projections

Several studies have shown trends in tornado activity. Studies have shown that an eastern shift in tornado frequency has taken place over the past 30 years, where tornado frequency has decreased in the southern Great Plains and an increase in the Tennessee/Alabama area [41, 42]. The results of Childs et al. [39] align with the pattern shown by Agee et al. [41] and Gensini and Brooks [42] for cold season tornado activity. Tippett et al. [43] has shown that there has been an increase in number of tornadoes per outbreak.

The effect a future and warmer climate will have on severe weather remains uncertain [19, 25]. There has been a growing consensus that there will be more frequent extreme values of CS under global warming [25, 44–46]. CAPE and V_{shear} have been examined in climate models. CAPE is expected to increase under climate change while V_{shear} is expected to decrease, but the increases in CAPE are expected to outweigh the decreases in V_{shear} to allow for more favorable combinations of severe environments [25]. Agard and Emanuel [47] argue that the diurnal cycle of CAPE increases in amplitude in a warmer climate, however, as temperatures increase, the time to peak CAPE also increases, so the diurnal cycle may become more of a limiting factor in the future. Hoogewind et al. [48] showed that the severe weather season may lengthen and possibly extend by a month. Likewise, several studies have found that the peak of the tornado season is also shifting to earlier in the year [40, 49]. How the environmental proxies and their modes of temporal variability will vary in response to these different components change under future climate change is uncertain.

Seeley and Romps [25] suggested that the future degree of severe thunderstorms could be closely related to humidification of the low-level atmosphere, as the models that predicted an aridification of the central United States showed a decrease to convective instability. Importantly, prediction of changes in severe weather events themselves in a future climate is particularly difficult since such events, especially tornadoes, are at a scale too small to be resolved by reanalysis nor climate models. Recent work has downscaled climate models to a resolution capable of resolving supercell-like rotating convection and demonstrated significant changes in the realization of storms themselves for a given environment under climate change [48]. Despite such shortcomings, it is still essential to understand the nature of the spatiotemporal variability of the large-scale environmental ingredients from which storms develop within the climate system. A better understanding of climate's control on these parameters and the parameters' influence on tornadoes is needed so that we can see how tornado occurrence may change in a warmer future climate.

This study is, in part, building on the work of Tippett et al. (2012) [32] to improve the framework for long term prediction of tornado frequency. This could be used to investigate possible effects of climate change on tornado activity such as when running reanalysis models under future climate conditions. Climate change may alter the diurnal and seasonal cycle as well as interannual variations. As these modes of temporal variability may change we want to use proxies that capture these modes correctly.

1.1.6 Caveats of Model Data

Reanalysis datasets are a model's best guess for the environmental conditions at a specific time. Several studies have examined the output of various reanalysis datasets [50–52]. Gensini et al (2014) [50] compared North American Regional Reanalysis (NARR) output to raw radiosonde data for 23 different severe weather variables. It was found that kinematic variables were best represented by NARR while thermodynamic variables are hindered by errors in low-level moisture [50]. Allen and Karoly (2014) [51] compared European Centre for Medium-Range Weather Forecasts

Interim Re-analysis (ERA-Interim) data to sounding observations for three variables (MLCAPE, 0-6 kilometer V_{shear}, and MLCIN). They found that ERA-Interim data can provide reasonable estimations of MLCAPE, 0-6 kilometer V_{shear} was well represented outside of the coastal region by the reanalysis dataset, and that ERA-Interim tended to do a poor job in representing MLCIN [51]. King and Kennedy (2018) [52] compared several reanalysis datasets (NARR, ERA-Interim, Modern-Era Retrospective Analysis for Research and Applications (MERRA 2), Japanese 55-year Reanalysis (JRA55), 20th Century Reanalysis (20CR), and the Climate Forecast System Reanalysis (CFSR)) to Rapid Update Cycle (RUC-2) proximity soundings. They found that NARR and JRA55 were the only reanalysis sets for which CAPE fell within the error-bars of RUC-2, all of the other reanalysis datasets were biased low for thermodynamic parameters and kinematic parameters that incorporate thermodynamic information, and all of the reanalysis datasets reasonably captured the kinematic environments [52].

Deficiencies in climate models also exist. Seeley and Romps' (2015) [25] results showed disagreement between even the models that performed highly in matching the radiosonde observations of severe thunderstorm environments. Allen et al (2014) [53] tested the performance of the Commonwealth Scientific and Industrial Research Organisation Mark, version 3.6 (CSIRO Mk3.6) and the Cubic-Conformal Atmospheric Model (CCAM) climate models over Australia by comparing them to ERA-Interim data. CSIRO MK3.6 significantly overestimated the frequency of severe thunderstorm environments while CCAM's distribution was closer to the ERA-Interim's distribution, but CCAM was influenced by negative biases in both CAPE and 0-6 kilometer V_{shear}. Limitations in knowledge and in model resolution pose significant challenges to predicting tornado activity [45] and downscaling models to simulate severe storms is computationally expensive [48]. .

1.2 Research Objectives

This work seeks to:

1. Compare the performance of common severe weather proxies in reproducing the climatological spatial variability in tornado activity;

2. Examine this performance across three dominant climate time-scales: diurnal variation, seasonal variation, and interannual variation;

3. Test the sensitivity of results to the chosen statistic, particularly mean vs. extremes

CHAPTER 2. DATA AND METHODS

2.1 Data

2.1.1 North American Regional Reanalysis

North American Regional Reanalysis (NARR) datasets are used for environmental parameters for the period from January 1979 to December 2015 [54]. NARR datasets are at a high resolution with a grid spacing of 0.3° and at a high temporal frequency providing 3 hourly data for the entirety of this historical period [55]. CAPE is taken from the NARR's internal variable for this parameter. V_{shear} and SRH are calculated from NARR's output of three dimensional wind fields. LCL is calculated using Romp's equation (see equation 1.3 in Section 2) from NARR output temperature and pressure data. CS is calculated with equal weight given to both CAPE and V_{shear}, while EHI is calculated using the equation mentioned in Section 2 (equation 1.11 using 0-3 kilometer SRH and STP is calculated using two methods. STP is calculated using equation 1.12 in Section 2 using 0-3 kilometer SRH and it is also calculated using 0-1 kilometer SRH and with a CIN term:

$$STP = \left(\frac{CAPE}{1500}\right) * \left(\frac{V_{shear}}{20}\right) * \left(\frac{SRH}{150}\right) * \left(\frac{2000 - LCL}{1000}\right) * \left(\frac{250 + CIN}{200}\right) \quad (2.1)$$

. Two limits were also placed on the STP equations. The LCL term was set to 0 if LCL was greater than 2000 m and the CIN term was set to 0 if CIN was less than -250 J/kg.

2.1.2 Tornado Database

Tornado observational data are sourced from the National Weather Service (NWS) Storm Prediction Center (SPC) "Actual Tornadoes" dataset [56]. This dataset is used for all calculations involving tornado count and density. Tornado touchdown points were used rather than paths or end points. We remove F/EF 0 tornadoes and hurricane induced tornadoes from the dataset. Hurricane induced tornadoes (HIT) are removed using methods similar to those described in Schultz and Cecil (2009) [57]. HITs are identified in the NWS data by incorporating NOAA's HURDAT2 data [58]. HURDAT2 is used to ascertain the date, time, and location of the center of all Atlantic tropical cyclones that occurred between 1979 and 2015. If a tornado occurred within 750 kilometers of the center of a tropical cyclone and within 3 hours of the cyclone's passing, it was considered "hurricane induced" and omitted from the dataset.

2.2 Methods

2.2.1 Tornado analysis

Datasets of tornado density are created from the NWS tornado database ("Actual tornadoes", [56]) for years 1979-2015. We use the NARR 0.3° spatial grid to calculate the tornado datasets. At each point in the grid, a fixed great-circle radius is used to count tornado and tornado day occurrence. A tornado day is defined as a day with at least 1 tornado occurring at any point in the day. We perform our analyses on strictly significant tornado events (EF2 or higher) as well. The radii (R) we use for counting tornadoes are $R = 50\ km$ and $R = 120\ km$. We utilize a 50 kilometer radius to approximate a 1° x 1° box, and a 120 kilometer radius to match the spatial smoother used by [38]. We divide by the total number of years to get our data in terms of tornadoes per year per 100 km^2. For both the seasonal and diurnal modes of variability, we normalize the time to one year by multiplying the tornado density by the number of time steps. From this, we calculate the spatial distribution of tornadoes over the continental United States. Our analysis is focused on the region east of the Rocky Mountains where the vast majority of tornadoes occur.

Tornado density and tornado day density are calculated for the three temporal cycles of interest (diurnal, seasonal, and interannual), and for the full climatology. To match the times of tornadoes (CST) to the NARR data (UTC), we synchronize the data to be in terms of Coordinated Universal Time (UTC). The tornado dataset records explicit times for the tornado events, while NARR uses 8 3-hourly time steps in a day starting at 00 UTC. Thus, we bin all tornado data into 3-hourly segments in Coordinated Universal Time (UTC) to align with the NARR temporal frequency. For the diurnal cycle, tornadoes are binned by the closest preceding NARR timestep, for example, all tornadoes that occur within 00-03 UTC are assigned to the 00Z NARR timestep. For the seasonal cycle, we group the tornadoes by month. For interannual variations, we group the tornadoes by year. Our primary analysis and the examples used hereafter are done for tornadoes rated EF/F1+ with $R = 50\ km$. We then examine sensitivities to these results based on using significant tornadoes (EF/F2+) and tornado days as well as testing the sensitivity to increasing the radius to 120 km.

An example of the time series for tornado density across all three climatological time scales at Lafayette Indiana is given in Figure 2.1a-c. These show the times of peak tornado occurrence. The diurnal cycle (Figure 2.1a) displays a maxima at 18Z with relatively high values at 21Z and 00Z. The seasonal cycle (Figure 2.1b) displays an absolute maxima in June and two local maxima in April and November. The interannual variations (Figure 2.1c) displays an absolute maxima in 2013.

2.2.2 Environmental parameter analysis

We next calculate our environmental proxies at each NARR grid point and calculate distributions of these proxies for the full climatology and across climatological modes of temporal variability. We first calculate the probability distributions for the full climatology of each severe weather environmental proxy at each point in the NARR grid. For a given distribution, we extract both standard central tendency statistics(mean, median) and several extreme percentiles (75^{th}, 90^{th}, 95^{th}, 99^{th}, and 99.9^{th}). This method is then identically applied to analyze variability across the diurnal cycle, the seasonal cycle, and in interannual variations in the same manner as for tornadoes as described above.

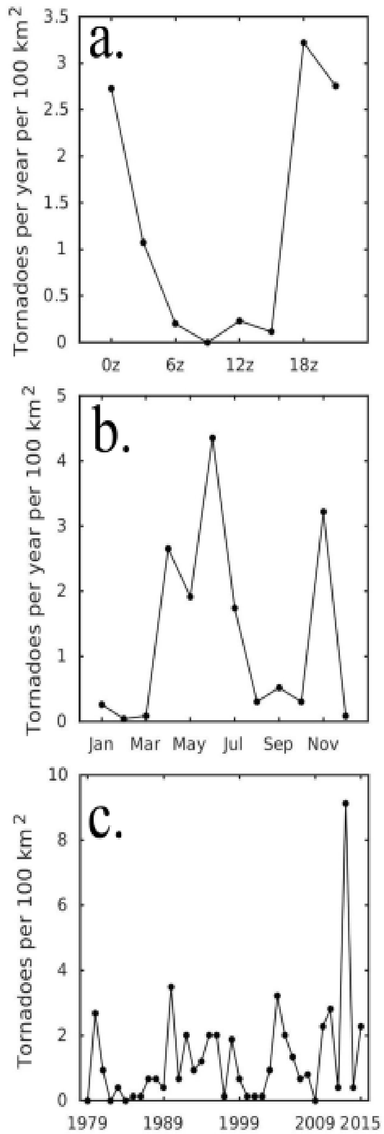

Figure 2.1. Time series of tornado density calculated when $R = 50\ kms$ for Lafayette, Indiana: a. diurnal tornado counts per year, b. seasonal tornado counts per year, c. interannual tornadoes. The full climatology tornado density is 1.29 tornadoes per year.

An example of this analysis is displayed at Lafayette, IN in Figure 2.2 for each of the three climatological time scales for each of the proxies. At every grid point, we calculate our desired metrics from the binned data. This allows us to capture the spatial distribution of each metric of the different environmental proxies. The full climatological PDF is given by the black curves in Figures 2.2a-i. From these PDFs, we calculate our set of statistics in order to generate timeseries of each statistic across each of the three climatological timescales at each gridpoint in our domain.

Figure 2.3 is an example for Lafayette, Indiana for time series of the 99^{th} percentile of the three environmental proxies along each of the climatological time scales. For the diurnal cycle, CS (Figure 2.3a) and EHI (Figure 2.3d) both have their highest values at 21Z and 00Z. STP (Figure 2.3g) has its highest value at 03Z. For the seasonal cycle, all three proxies have their highest values in June and July. For interannual variability, CS (Figure 2.3c) has its absolute maxima in 1980. However, EHI (Figure 2.3f) and STP (Figure 2.3i) have two local maxima in 1980 and 2011. A slight upward trend can also be seen in the interannual values of the 99^{th} percentile of each of the three proxies.

2.2.3 Covariability Analyses

Finally, we combine the aforementioned datasets to test how well each statistic for each proxy captures the spatial distribution of tornado activity. We begin with the full climatology and end with our three individual modes of climatological temporal variability.

Full Climatology

We perform spatial correlations between the NARR full climatology proxy datasets (CS, EHI, STP) and the tornado density datasets. CAPE is used as a baseline comparison against the other four proxy calculations. These correlations are performed for all of the metrics specified for the NARR data. The values of a metric of a proxy

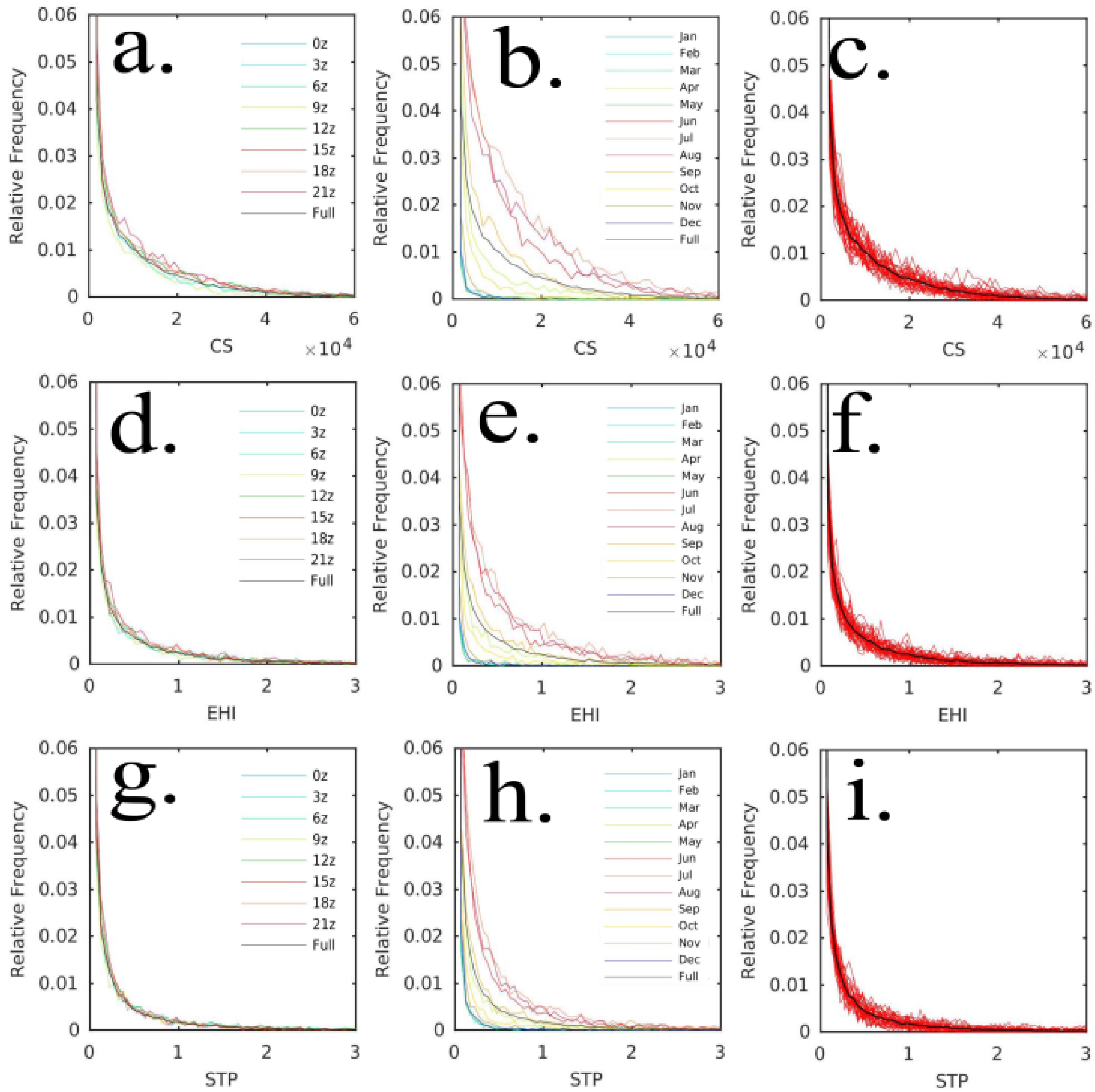

Figure 2.2. Probability distributions for the three environmental proxies for Lafayette Indiana, across the three dominant time scales: a. diurnal CS, b. seasonal CS, c. interannual CS, d. diurnal EHI, e. seasonal EHI, f. interannual EHI, g. diurnal STP without using CIN, h. seasonal STP without using CIN, i. interannual STP without using CIN. The black lines represent the full climatology. Bin width for all three quantities is 0.05.

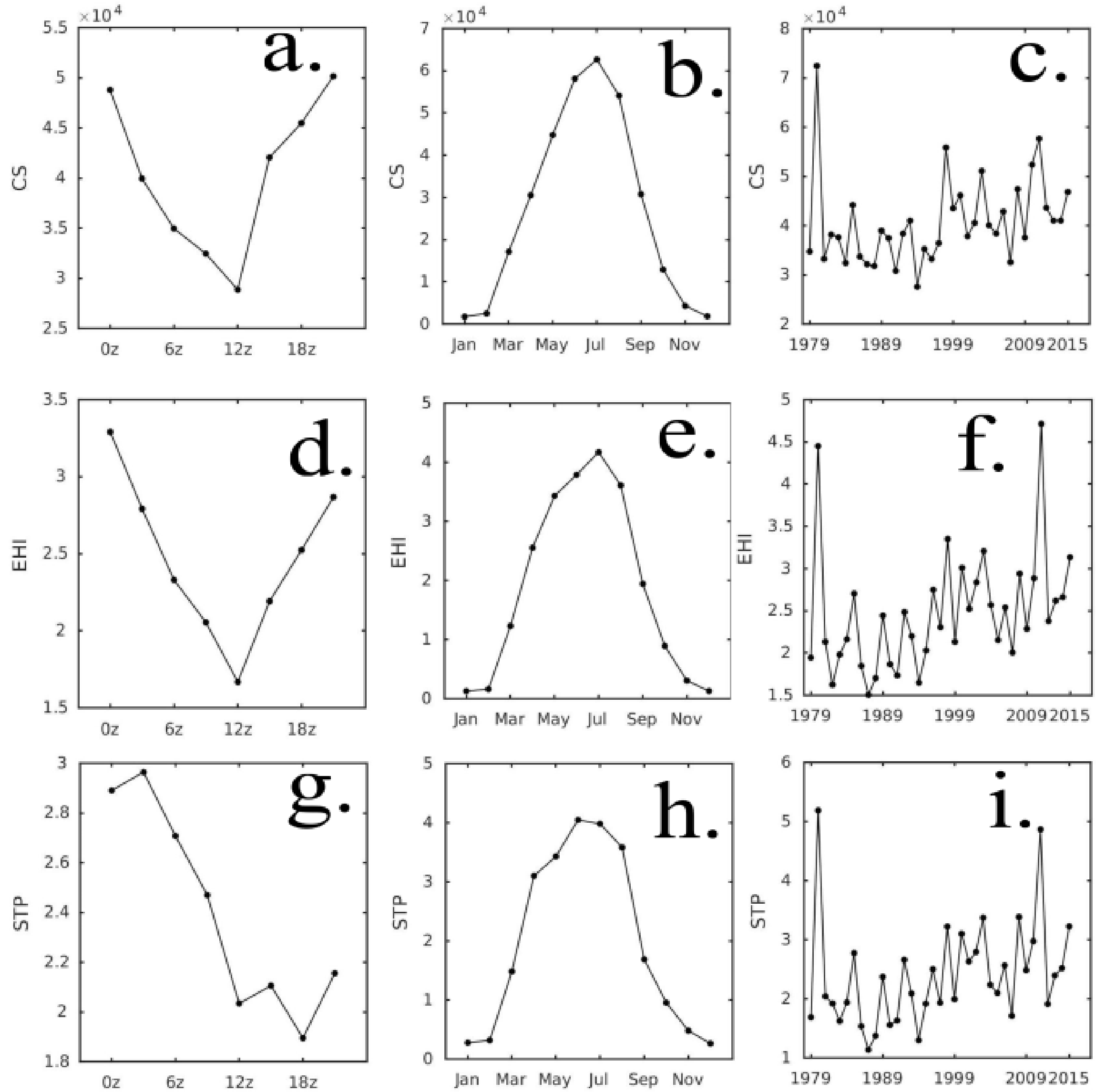

Figure 2.3. As in Figure 2.1, but for the 99th percentiles of the three proxies for Lafayette, Indiana across the three dominant time scales: a. diurnal CS, b. seasonal CS, c. interannual CS, d. diurnal EHI, e. seasonal EHI, f. interannual EHI, g. diurnal STP without using CIN, h. seasonal STP without using CIN, i. interannual STP without using CIN.

are plotted against the values for tornado density at the same location, and the correlation coefficient "R" is calculated between them. To reduce a weighting of the correlation coefficient toward zero, we require that more than five tornadoes need to have occurred in the 37 years of our data to be used in the correlation calculation.

We employ bootstrapping to calculate the spatial correlation coefficient and the correlation coefficient's 95 percent confidence interval. To accomplish this, we utilize 1000 bootstrapped samples of our spatial distribution of points and calculate correlation coefficients for each sample. From this set of 1000 correlation coefficients, we take the median, 2.5^{th} percentile, and 97.5^{th} percentile. This process is done for each metric of all proxies across all of the different methods used for tornado counting. We consider the proxies to be statistically similar if their respective 95 percent confidence intervals overlap.

Temporal Variations

Next, we analyze the extent to which our severe weather environmental proxies capture the spatial distribution of temporal variability in the tornado density on three dominant time scales of climate forcing: diurnal, seasonal, and interannual. Each point has its own sets of values for tornado density (example across all three time scales for Lafayette, Indiana in Figure 2.1a-c) and for a metric of a proxy (example of the 99th percentile of the three proxies at Lafayette, Indiana in Figure 2.3a-i) across the time modes of interest. With each set, the datasets will be tested to determine how each proxy captures variability. The values across the time scales are plotted against each other, and the correlation coefficient is calculated between each set of data. Performing this at all points, results in a spatial distribution of correlation coefficients over our selected area of the United States. This spatial distribution can then be plotted on a map to show how well each proxy maps the variations in tornado density.

The temporal analysis described above is applied to all grid points in our domain and local correlations are calculated to yield a map of the spatial distribution of co-variability for a given climatological mode of temporal variability. As a simple means of measuring the performance across the domain and comparing across metrics, the correlation coefficients are averaged across the entire domain. Again, we utilize 1000 bootstrapped samples of our spatial distribution of points and calculate correlation coefficients for each sample. From this set of 1000 correlation coefficients, we take the median, the 2.5^{th}, and the 97.5^{th} percentiles of the average of the local correlation coefficients to obtain a median and a 95 percent confidence interval to represent the domain. We apply the same definition for statistically similar as mentioned prior. This process is done for all metrics of each of the proxies and all the methods used for counting tornadoes.

CHAPTER 3. RESULTS

3.0.1 Analysis: Full Climatology

The spatial distribution of tornado density for $R = 50\ km$ for all EF1/F1+ tornadoes is shown in Figure 3.1. This reveals five maxima in tornado occurrence located in northern Colorado, central Oklahoma, along a northern part of the border between Texas and Louisiana, central Arkansas, and southern Mississippi.

The spatial distributions of the 99th percentile of each of the three proxies are shown in Figure 3.2. For example, the 99th percentile of CS (Figure 3.2a) has maxima over land located in central Texas and in the northern plains in northeastern Kansas, southeastern Nebraska, southwestern Iowa, and northwestern Missouri. Interestingly, CS has maxima over the ocean. These over-ocean maxima are in the central Gulf of Mexico, off the eastern coast of Florida, and off the western coast of Mexico. EHI (Figure 3.2b) has one peak that occurs in the southern great plains in Northern Texas, eastern/central Oklahoma, and eastern Kansas. STP (Figure 3.2c) peaks in the southern great plains in south-central Kansas, central Oklahoma, and Northern Texas. Interestingly, like CS, EHI and STP both have relatively high values over the ocean in the Gulf of Mexico.

Figures 3.3 and 3.4 display the spatial relationship between tornado density and the 99th percentile of each of our proxies for the full climatology. Figures 3.3 and 3.4 display the spatial distribution of both the 99^{th} percentile of a given proxy (Figures 3.3ace) and the statistical relationship between these spatial distributions, including their linear correlation (Figures 3.3bdf and 3.4bdf); results are displayed for CAPE (Figures 3.3ab and 3.4ab), CS (Figures 3.3cd), EHI (Figures 3.3ef), our STP formulation without CIN (Figures 3.4cd), and our STP formulation with CIN (Figures 3.4ef.

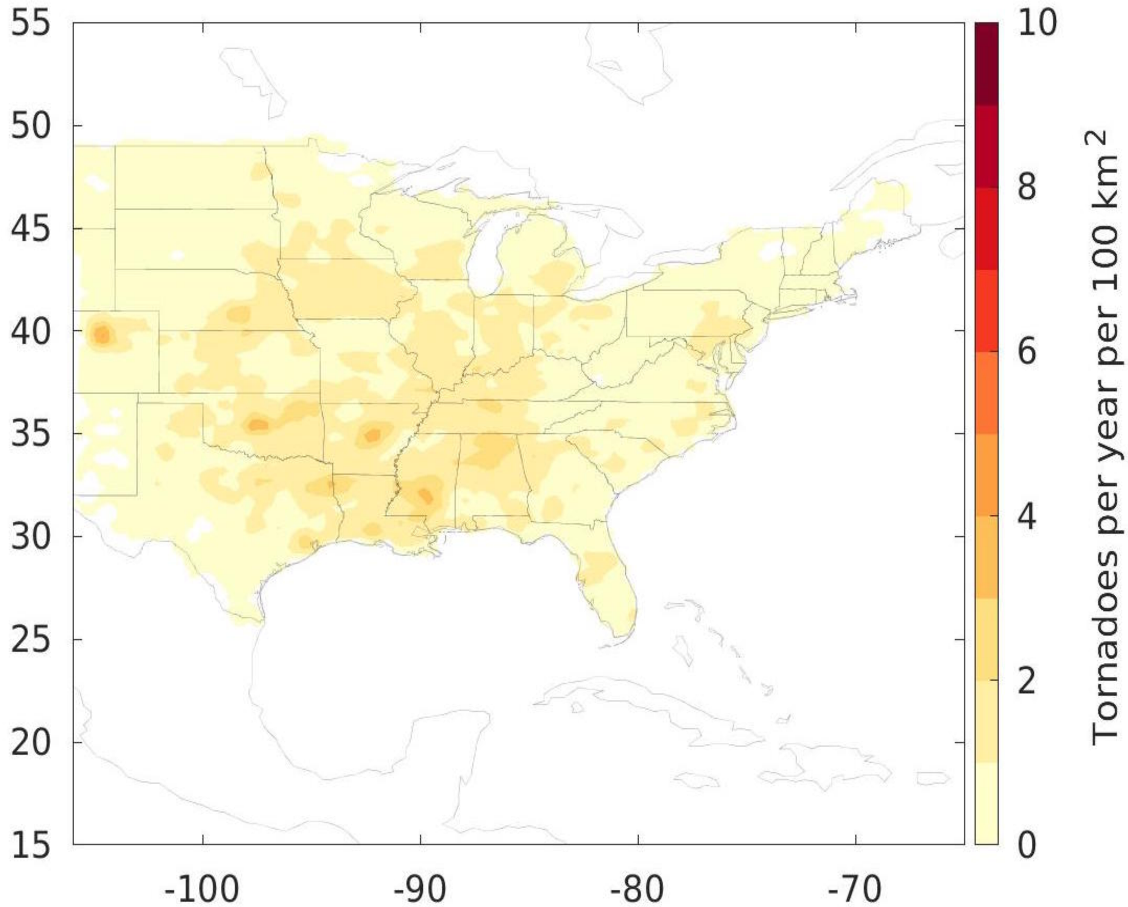

Figure 3.1. Tornado density calculated for EF1+ tornadoes at R = 50 km over the continental United States, east of the Rocky Mountains. White areas are those without applicable tornado data.

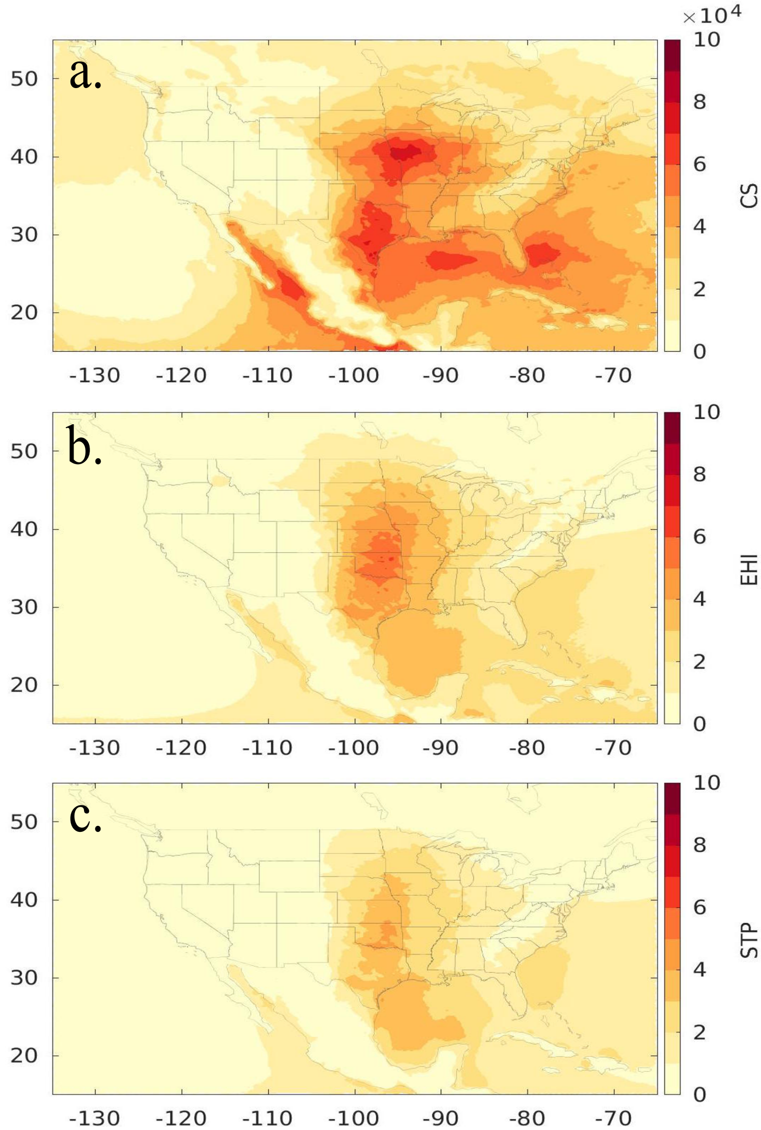

Figure 3.2. The 99th percentiles of a. CS, b. EHI, and c. STP without using CIN over a large domain that includes the continental United States.

For our analysis we exclude grid points with tornado density values less than 0.15 tornadoes per year per 100 km^2. These regions are blocked out (white).

Correlation coefficients between tornado activity metrics and our statistics for each environmental proxy are displayed in Figure 3.5; results are displayed for EF/F1+ tornado density (Figure 3.5a), EF/F2+ tornado density (Figure 3.5b), and tornado day density (Figure 3.5c). Beginning with EF/F1+ (Figure 3.5a), for CAPE, the 95^{th}, the 99^{th}, and the 99.9^{th} percentiles are the best performers. These three percentiles outperform the mean. For CS, the 90^{th} and the 95^{th} percentiles are the best performers. These two percentiles outperform the mean. For EHI, the mean and percentiles over the 75^{th} (90^{th}, 95^{th}, 99^{th}, and 99.9^{th}) are all statistically similar to each other. For STP calculated without CIN, the highest extreme percentiles (99^{th} and 99.9^{th}) perform best, while the mean is slightly less well correlated but is still statistically similar to these two extreme percentiles. For STP with CIN included in calculations, the highest extreme percentiles (99^{th} and 99.9^{th}) perform best. These two percentiles outperform the mean. The median consistently exhibits the lowest correlation coefficient values, which is due to the fact that these proxies often take values near zero, so that the median is also often near or at zero. When restricting tornadoes to EF/F2+, the correlation coefficients decrease across all statistics except at the 75^{th} percentile of STP calculated without the CIN term where the number does not change. All other statistics decrease by 7-77%. The qualitative results remain largely unchanged. For the analysis using tornado days, the correlation coefficients increase by 10-40% across all statistics except the median. We disregard the median for calculating a percent change due to its values' proximity to zero and that some values transition from negative to positive. For CS, the mean is now outperformed by the 90^{th} and 95^{th} percentiles. The qualitative results remain largely unchanged for both EHI and STP.

Sensitivity of these results to a larger tornado radius is shown in Figure 3.6, which is analogous to Figure 7 but with $R = 120\ km$. Most correlation coefficients are found to increase between 2-107% across the different proxies and statistics. The statistics

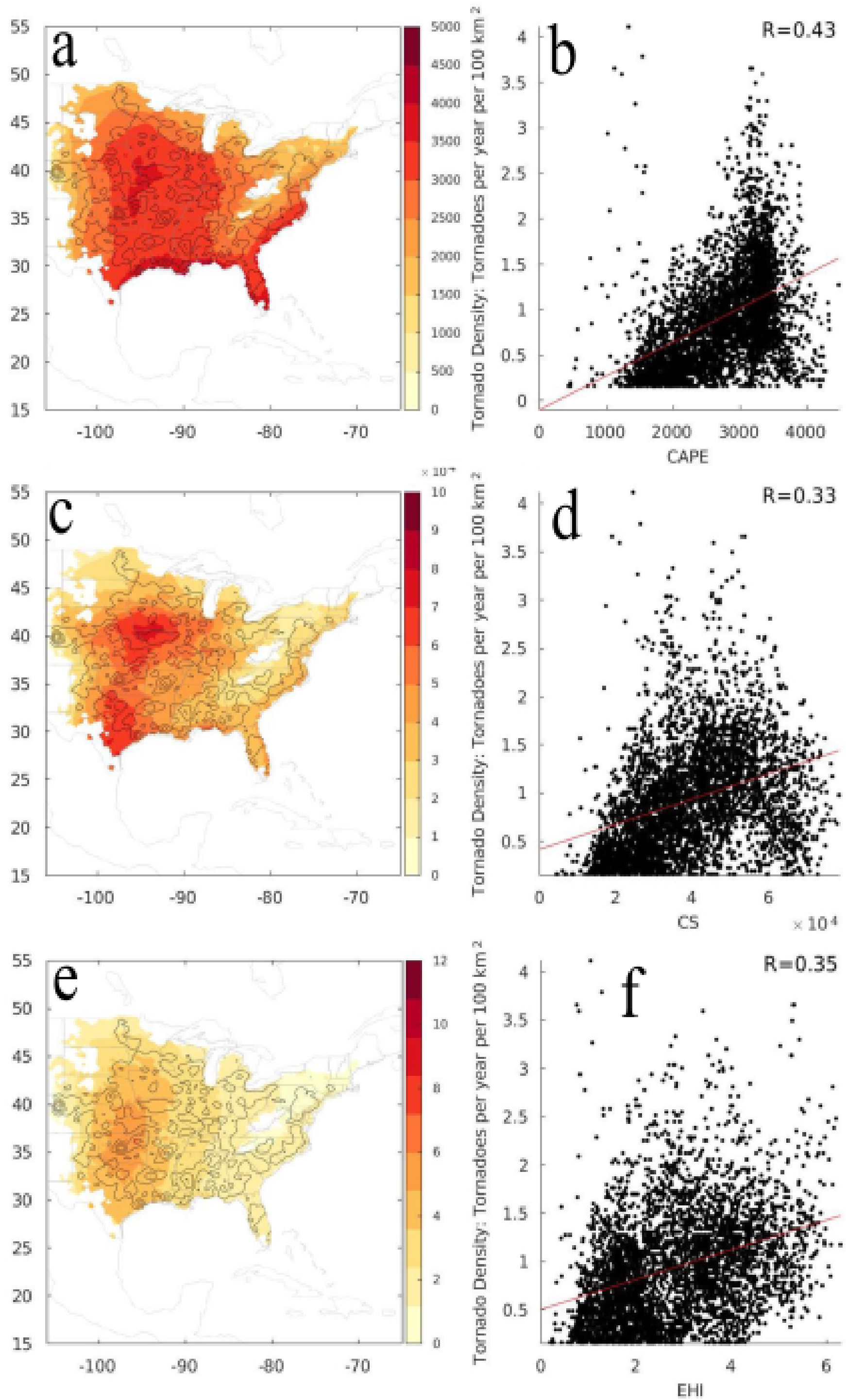

Figure 3.3. Parts a, c, and e: Color-filled contour of the 99th percentile of the full climatology of a. CAPE, c. CS, and e. EHI with a contour of EF1+ tornado density calculated at R = 50 km overlayed. Locations with less than .15 tornadoes per year per 100 km^2 and points over the ocean have been removed. Parts b, d, and f: Scatter plot of 99th percentile of b. CAPE, d. CS, and f. EHI vs. tornado density with a line of best fit and the correlation coefficient in the top right corner.

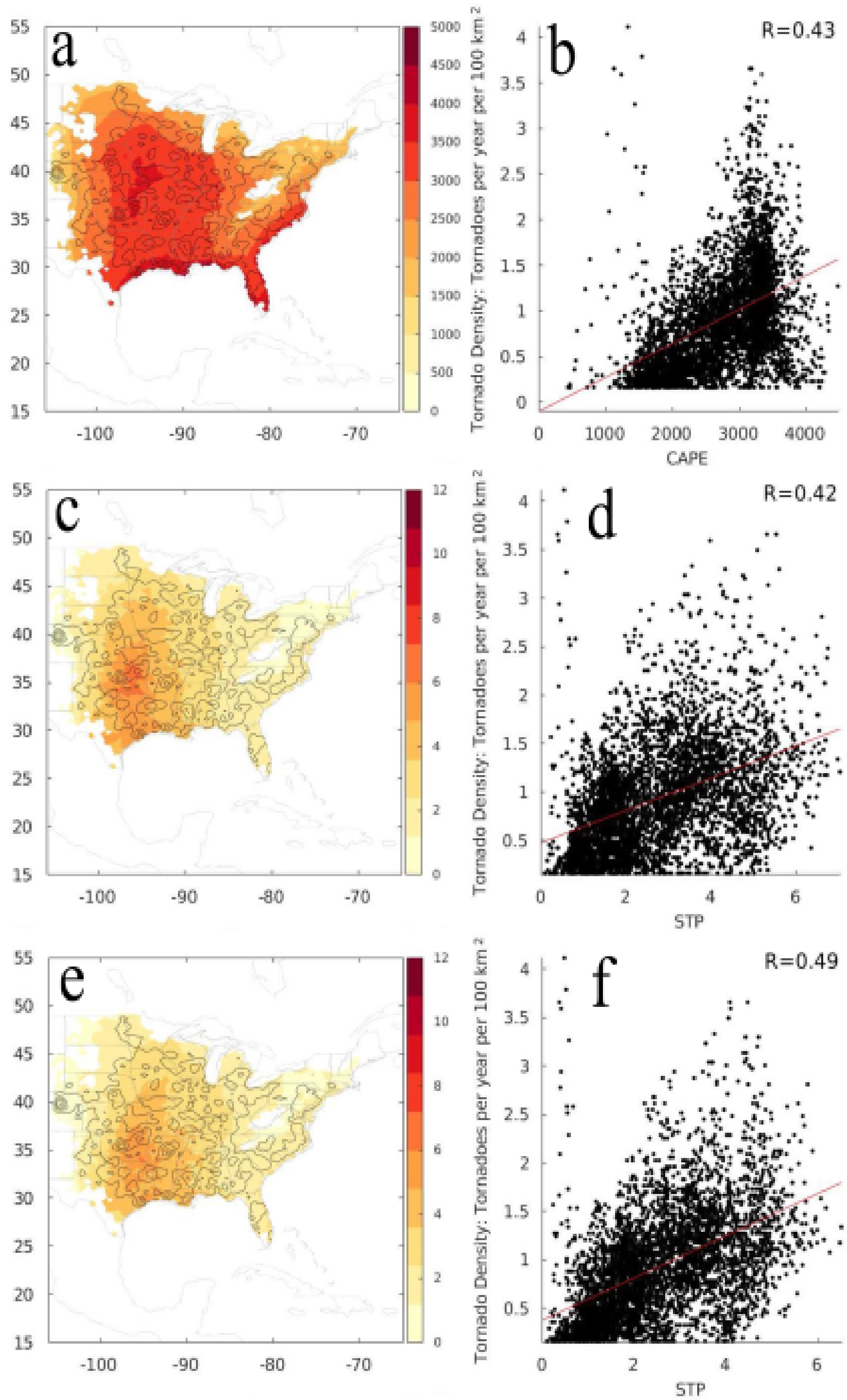

Figure 3.4. Parts a, c, and e: Color-filled contour of the 99th percentile of the full climatology of a. CAPE, c. STP equation without CIN, and e. STP equation with CIN included with a contour of EF1+ tornado density calculated at R = 50 km overlayed. Locations with less than .15 tornadoes per year per 100 km^2 and points over the ocean have been removed. Parts b, d, and f: Scatter plot of 99th percentile of b. CAPE, d. CS, and f. EHI vs. tornado density with a line of best fit and the correlation coefficient in the top right corner.

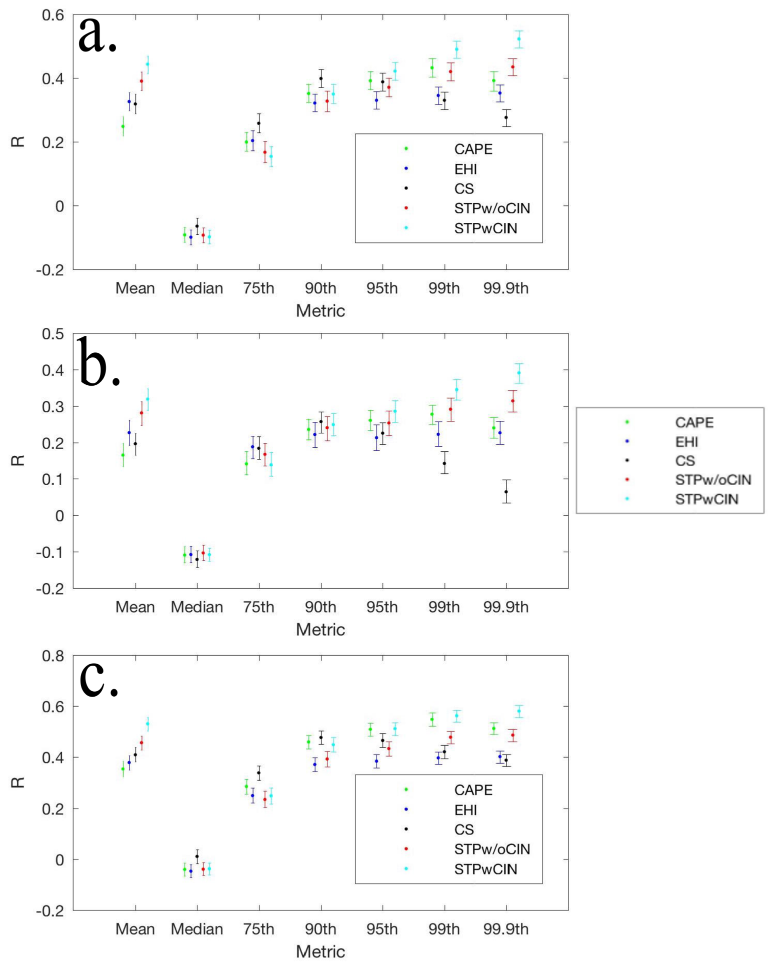

Figure 3.5. Scatter plots between full climatological proxy metrics and correlation coefficients using different methods for counting tornadoes for $R = 50km$ for: a. all tornadoes rated EF/F1+, b. all significant tornadoes (EF/F2+), and c. all tornado days. Green is CAPE. Blue is EHI. Black is CS. Red is STP without using CIN. Turquoise is STP with CIN included. Error bars represent the 95 percent confidence intervals derived from 1000 bootstrapped re-samplings of the data.

that do no increase when changing radius are the 75^{th} percentile of STP without the CIN term for EF/F1+ tornadoes which decreases by 3% , the 75^{th} percentile of STP without the CIN term for tornado days which stays constant, and the 75th percentile of STP with the CIN term for tornadoes rated EF/F1+. This result is qualitatively consistent across the different tornado activity definitions (EF/F1+, EF/F2+, and tornado days). Overall, the qualitative results remain largely unchanged, with a slight exception for EHI. For EHI, the highest two percentiles (99^{th} and 99.9^{th}) have increased relative to the other statistics of EHI when tornadoes rated EF/F 1+ and tornado days are used. Under these circumstances, the two highest percentiles become the best performers and the other three statistics (mean, 90^{th}, and 95^{th}) are slightly less well correlated.

3.0.2 Analysis: Temporal Variations

We next analyze variability across climatological timescales. We begin with an example demonstration for analyzing the relationship between tornado activity and environmental proxies over diurnal, seasonal, and interannual timescales for Lafayette, IN. Figure 9 displays the relationship between the 99th percentile of each environmental proxy and tornado density along the three temporal modes of variation at Lafayette. For the diurnal timescale, (Figures 3.7a, 3.7d, and 3.7g) CS has the highest temporal correlation coefficient and STP has the lowest temporal correlation coefficient. For the seasonal timescale (Figures 3.7b, 3.7e, and 3.7h), STP calculated without the CIN term has the highest correlation coefficient, while CS has the lowest correlation coefficient. For the interannual timescale, (Figures 3.7c, 3.7f, and 3.7i), EHI has the highest correlation coefficient and STP has the lowest correlation coefficient. The outlier year for Lafayette, Indiana that none of the proxies capture well is 2013.

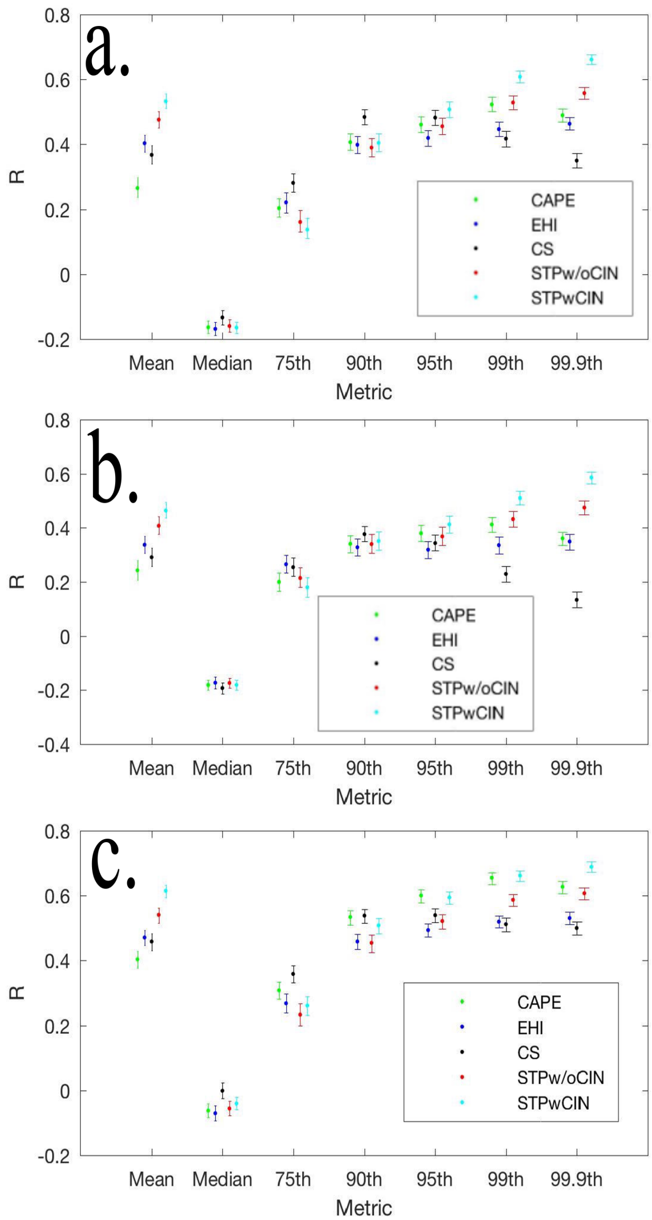

Figure 3.6. As in Figure 3.5, but for R = 120 km radii. Green is CAPE. Blue is EHI. Black is CS. Red is STP without using CIN. Turquoise is STP with CIN included.

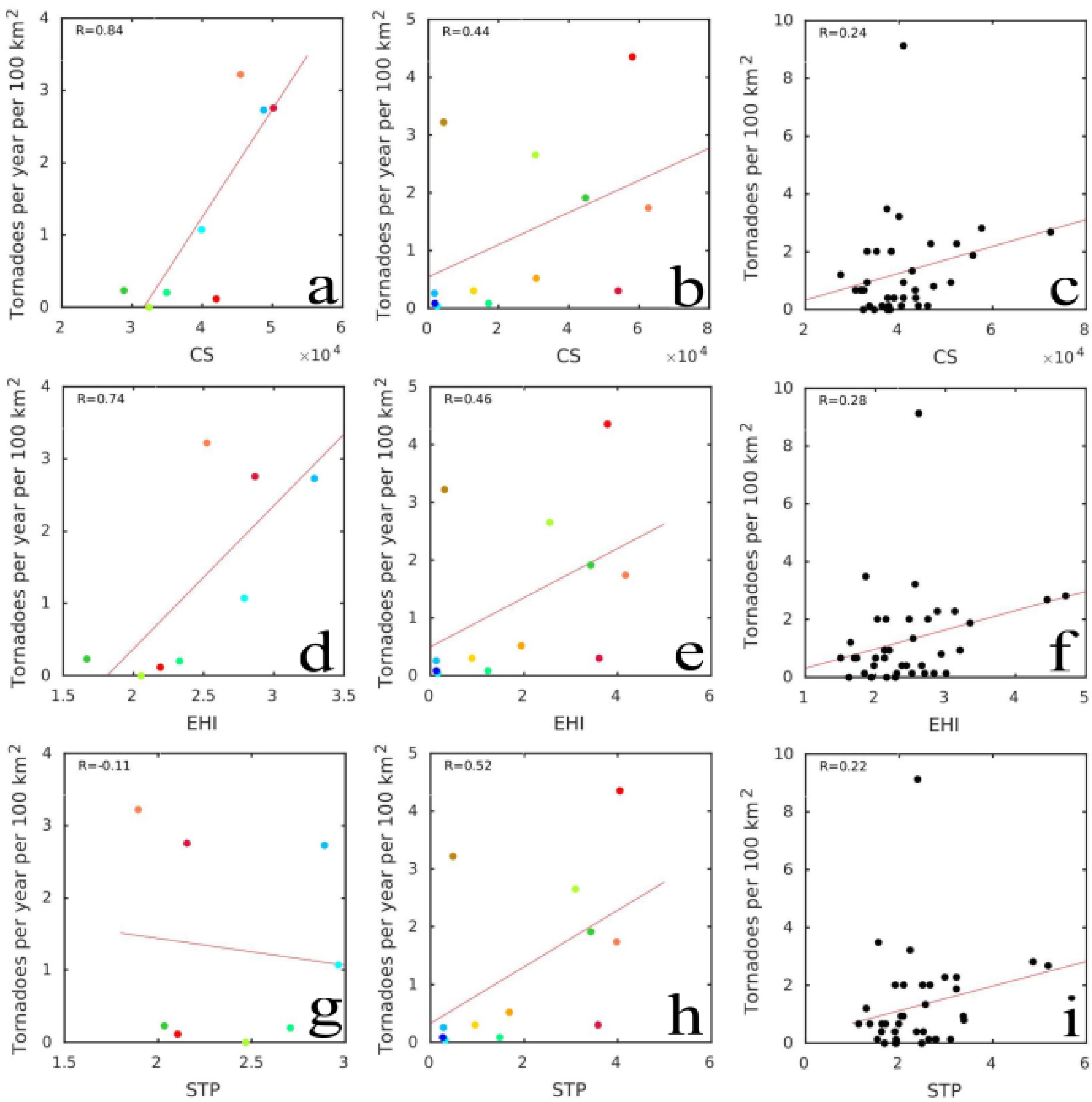

Figure 3.7. Scatter plots between 99th percentile proxies and tornado densities calculated at R = 50 km: a. Diurnal CS, b. Seasonal CS, c. Interannual CS, d. Diurnal EHI, e. Seasonal EHI, f. Interannual EHI, g. Diurnal STP without including CIN, h. Seasonal STP without including CIN, and i. Interannual STP without including CIN. The colors of the points in the diurnal and seasonal plots correspond to the same hours and months as in figure 2.2.

Seasonal Cycle

Maps of the temporal correlation coefficients between the seasonal variations of the 99th percentile of the environmental proxies and tornado density are displayed in Figure 3.8; results are displayed for CAPE (3.8a and 3.9a), CS (3.8b), and EHI (3.8c), the STP formulation without CIN (3.9b), and the STP formulation with CIN (3.9c). The areas in white are the same locations as before at which less than .15 tornadoes per year per $100km^2$ have occurred. The 99^{th} percentiles of these proxies work well seasonally in the Great Plains, the Midwest, and the Northeast. However, all proxies exhibit relatively smaller or negative correlations in the southeast. CAPE shows extremely poor correlations throughout the entirety of the southeast. CS shows negative correlations in eastern Florida, northwestern Florida, southeastern Georgia, central South Carolina, eastern North Carolina, southern Alabama, southern Louisiana, along the Arkansas/Mississippi border, north-central Arkansas, and south-central Missouri. EHI shows negative correlations in southeastern Georgia, northwestern Florida, along the Arkansas/Mississippi border, north-central Arkansas, and south-central Missouri. Both forms of STP show negative correlations in south-western Florida and north-western Florida.

Correlation coefficients between seasonal tornado activity metrics and our desired statistics for each environmental proxy are displayed in Figure 3.10. Beginning with results for EF/F1+ (Figure 3.10a), among all of the proxies, the two highest percentiles (99^{th} and 99.9^{th}) are the best correlated and the means have lower correlations. The 99^{th} and 99.9^{th} percentiles' correlations are higher than the mean by 36% and 40% respectively for CAPE, 45% and 53% respectively for CS, 21% and 30% respectively for EHI, 8% and 10% respectively for STP without CIN, and 8% and 5% respectively for the STP calculation with CIN. Other than the median, where STP is comparable to EHI, STP outperforms CS and EHI at every metric. EHI outperforms CS at every metric. These results are largely insensitive to our different tornado activity metrics. Using EF/F2+ tornadoes (Figure 3.10b) causes the correlation co-

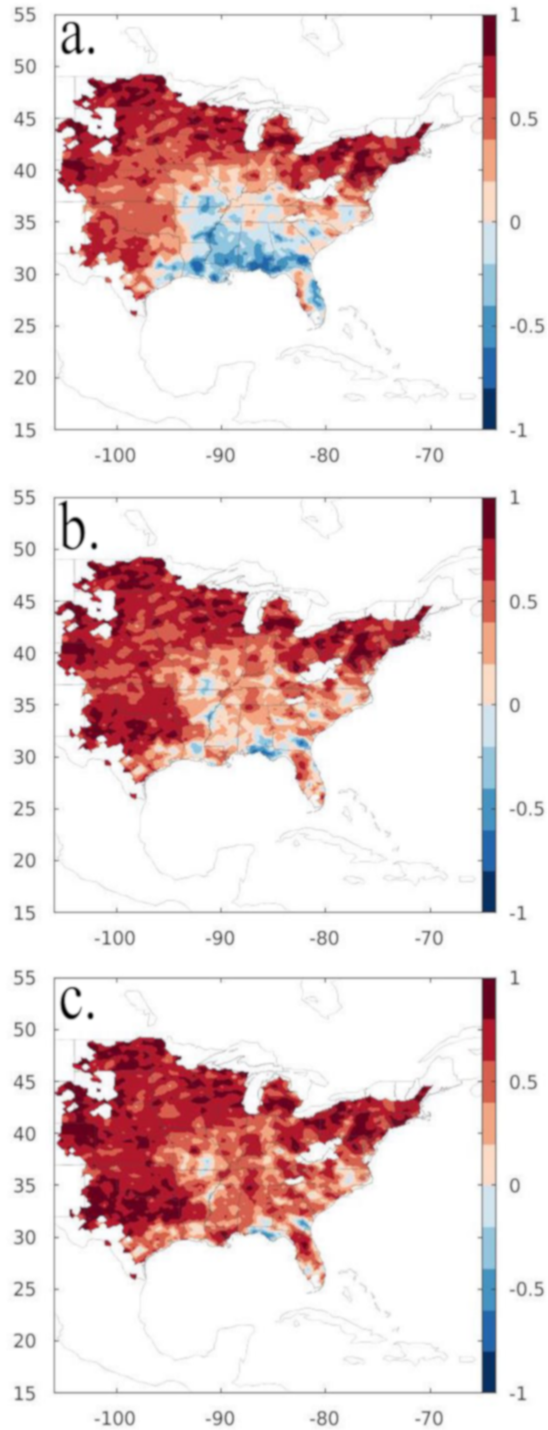

Figure 3.8. Maps of correlation coefficients between seasonal 99th percentile a. CAPE, b. CS, and c. EHI and tornado (EF/F 1+) density when R = 50 km.

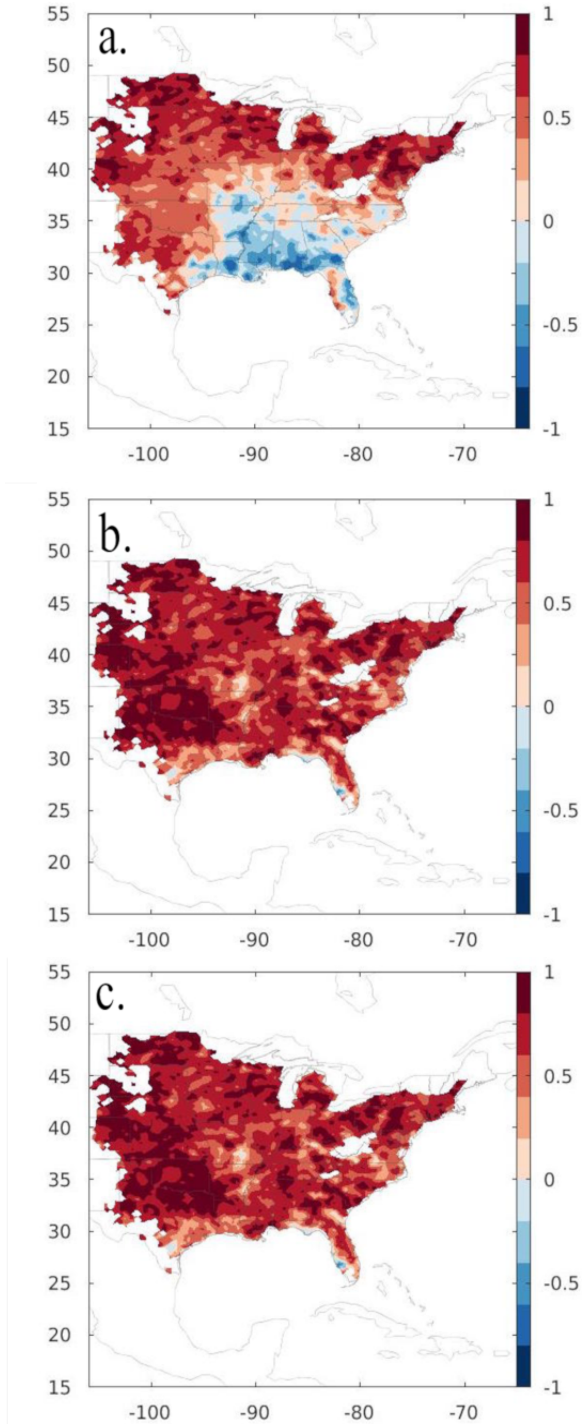

Figure 3.9. Maps of correlation coefficients between seasonal 99th percentile a. CAPE, b. STP without including CIN, and c. STP using CIN and tornado (EF/F 1+) density when R = 50 km.

efficients to decrease by 18-45% based on the proxy and statistic. Using tornado days (Figure 3.10c) causes the correlation coefficients to increase by 3-8% depending on the proxy and statistic. The qualitative results remain the same. Sensitivity of these results to a larger tornado radius is shown in Figure 3.11. These results are also largely insensitive to the use of a larger tornado radius, except as the radius increases from 50 kilometers to 120 kilometers, all mean correlation coefficients increase by 16-52% based on proxy and statistic.

Diurnal Cycle

Figure 3.12 displays maps of the correlation coefficients between the diurnal variations of 99th percentile of the three proxies and tornado density. CS (Figure 3.12b) performs well across much of this area of the United States and has the smallest area of negative correlations and these are in southern Florida, along the southern coast from western Florida to Louisiana, and in central Texas. For EHI (Figure 3.12c, in the southeast, negative correlations are found in southern Florida, western Florida, eastern Georgia, and southern Louisianna. In the Great Plains, negative correlation coefficients are found in southern Texas, central Texas, western Oklahoma, western Kansas, western Nebraska, northwestern South Dakota, central North Dakota, and southeast North Dakota. For STP without CIN (Figure 3.13b), the Great Plains exhibit almost exclusively negative correlations; the Southeast is variable where most of Florida, the Piedmont Plateau, southern Alabama, southern Mississippi, central Tennessee, northern Alabama, and at the border between Arkansas, Missouri, and Tennessee exhibit negative correlations; in the Midwest, the northern half of Indiana, most of Ohio, northern Kentucky, central Michigan, and southern Michigan exhibit negative correlations. For STP with CIN (Figure 3.13c), the central and southern Great Plains exhibit extremely low correlation coefficients west of -98^o latitude and high correlation values east of -98^o latitude, the northern Great Plains exhibit low correlation values, the Southeast and Mid- Atlantic are variable and follow extremely

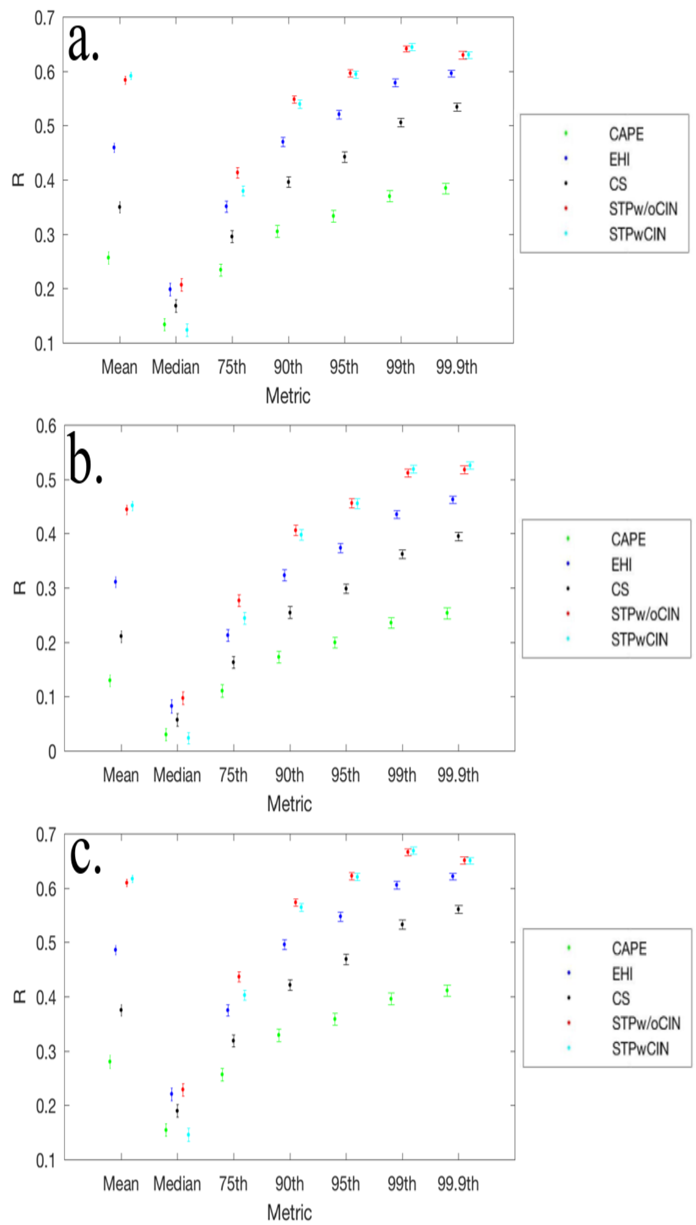

Figure 3.10. Scatter plots between monthly proxy metrics and correlation coefficients using different methods for counting tornadoes for R = 50 km for: a. all tornadoes rated EF/F 1 and greater, b. all significant tornadoes (EF/F2+), and c. all tornado days. Green is CAPE, EHI is blue, CS is black, STP without including CIN is red, and STP using CIN is turquoise. Error bars represent the 95 percent confidence intervals derived from bootstrapped re-sampling of the data.

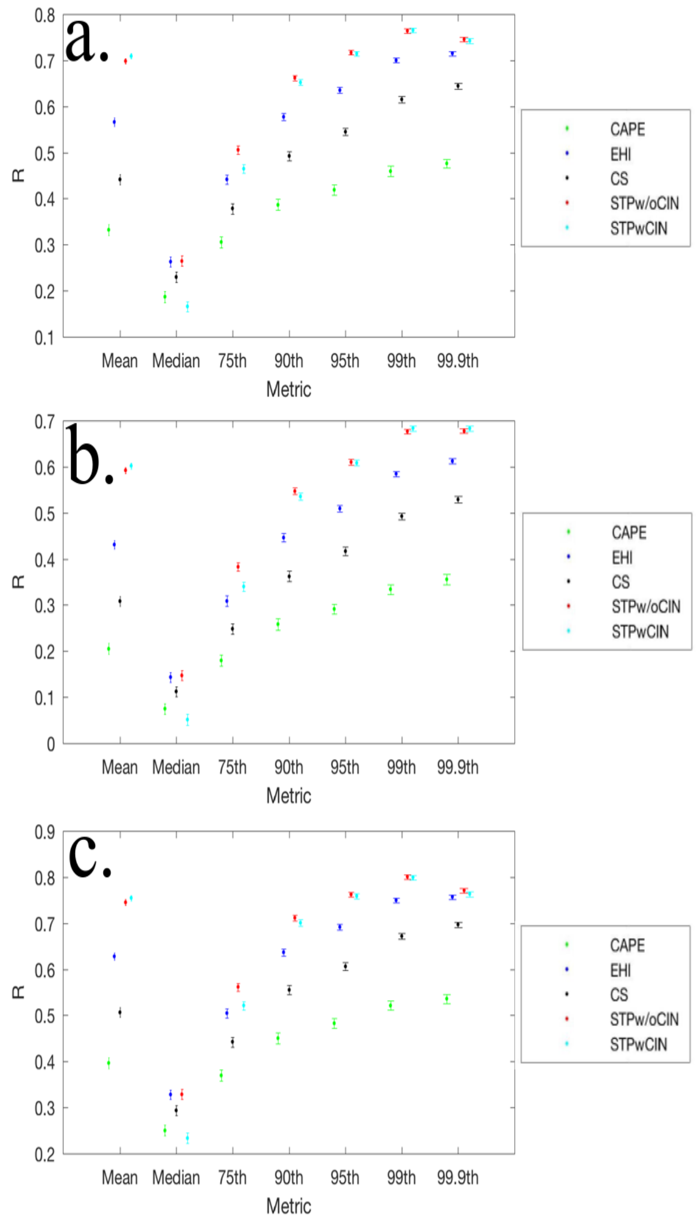

Figure 3.11. As in Figure 3.10, but for R = 120 km.

similar spatial patterns to previously those discussed for the other formula for STP, and the Midwest shows variability with low correlation values in Ohio, Michigan, Indiana, and Kentucky, but relatively high values throughout the rest of the Midwest..

Correlation coefficients between diurnal tornado activity metrics and our statistics for each environmental proxy are displayed in Figure 3.14. Beginning with EF/F1+ (Figure 3.14a), for CAPE, the mean and 75^{th} percentile are the highest correlated metrics and are statistically similar. The other extremes have lower mean correlation coefficients by 3-12%. For CS, the mean and 75^{th} percentile are the highest correlated metrics and statistically similar, while the extreme percentiles have slightly lower correlations by 3-6%. For EHI, the median and the 75^{th} percentile have the highest correlation and they outperform the mean by 11% and 14% respectively. For STP calculations without CIN, the median has the highest correlation coefficient, while the mean and other statistics have much lower correlation coefficients. For STP calculations with CIN, the median has the highest correlation coefficient and the other statistical metrics have much lower mean correlation coefficients. STP percent differences are not calculated due to the proxy's coefficients' proximities to zero. CS is the highest performing proxy for the diurnal cycle except at the median where CS and EHI are statistically similar. CAPE outperforms EHI at all metrics except the median. EHI outperforms both forms of STP at every metric. STP calculations with CIN outperform STP calculated without CIN at every metric.

When using EF/F2+ to test (Figure 3.14b), the qualitative results remain largely unchanged but correlation coefficients decrease by 12-21% from when EF/F1+ is used. When using tornado days (Figure 3.14c), the qualitative results are largely unchanged, but there is a small increase in correlation values of 3-4%. Sensitivity of these results to a larger tornado radius is shown in Figure 3.15. Results are largely insensitive to the use of a larger tornado radius, except as the radius increases from 50 kilometers to 120 kilometers, all correlation coefficients increase by 14-25% based on proxy and statistic.

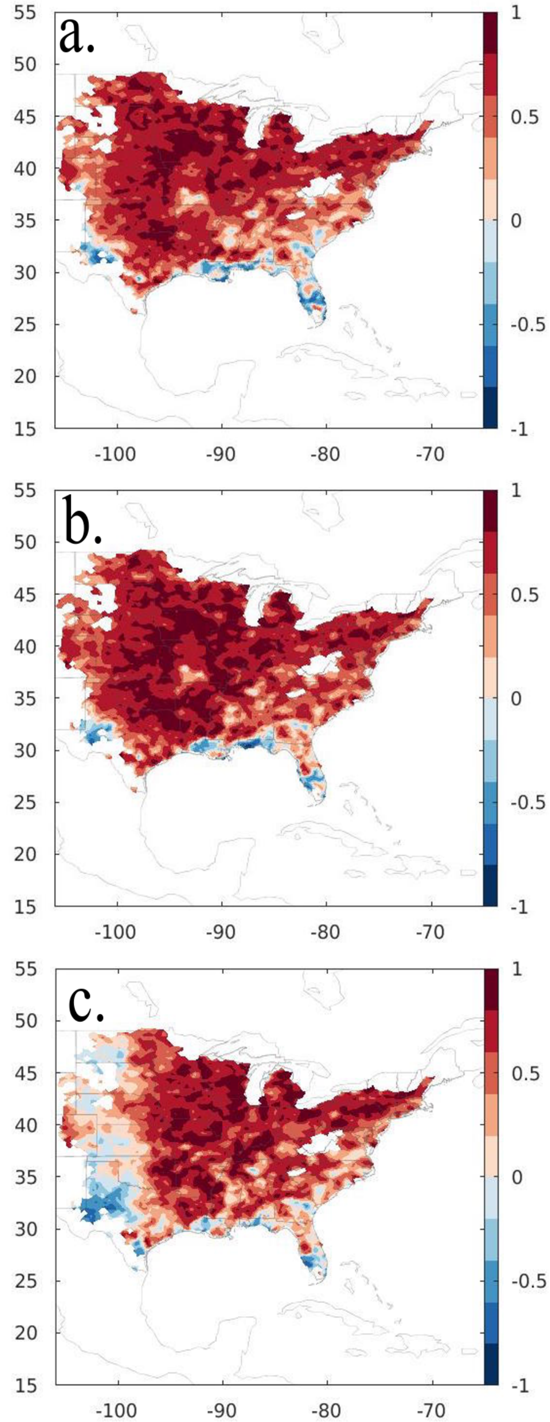

Figure 3.12. As in Figure 3.8, but for the diurnal cycle.

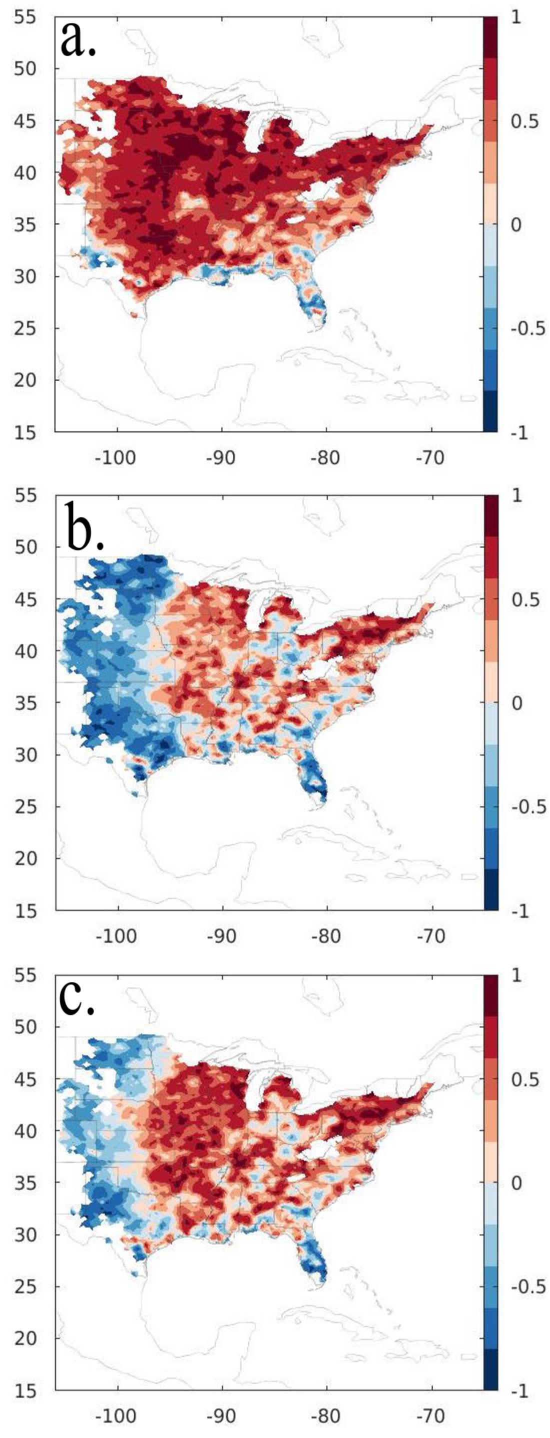

Figure 3.13. As in Figure 3.9, but for the diurnal cycle.

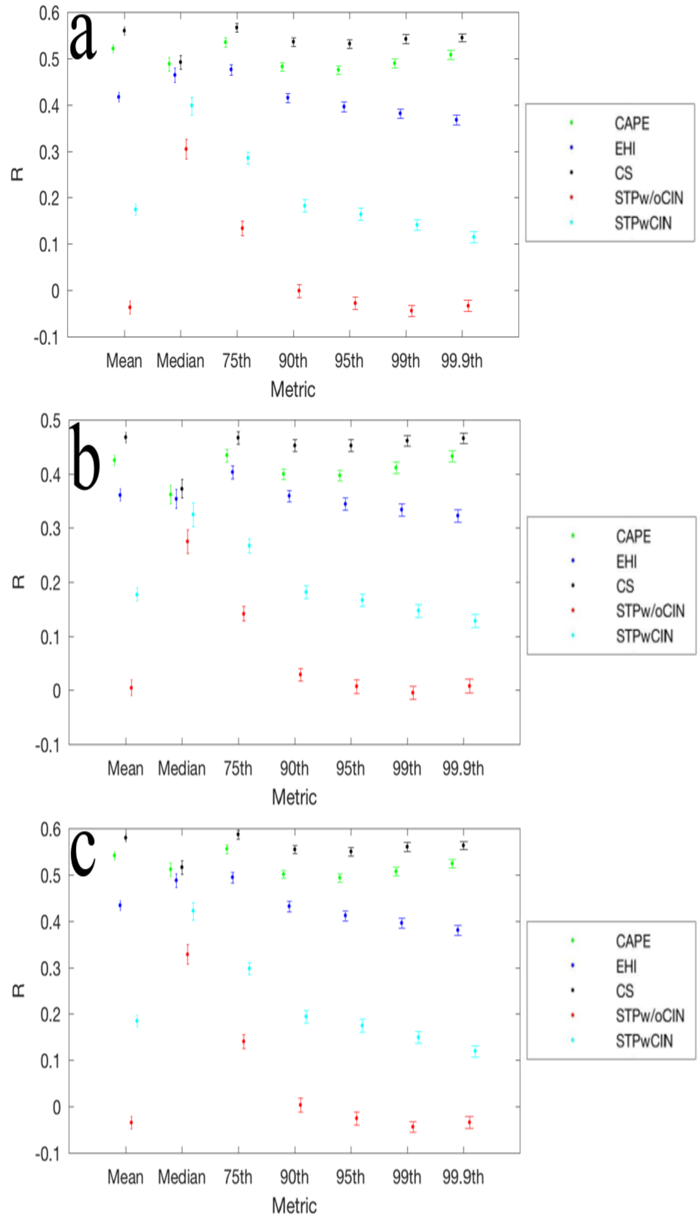

Figure 3.14. Scatter plots between diurnal proxy metrics and correlation coefficients using different methods for counting tornadoes for R = 50 km for: a. all tornadoes rated EF/F 1 and greater, b. all significant tornadoes (EF/F2+), and c. all tornado days. CAPE is green, EHI is blue, CS is black, STP without using CIN is red, and STP using CIN is turquoise. Error bars represent the 95 percent confidence intervals derived from bootstrapped re-sampling of the data.

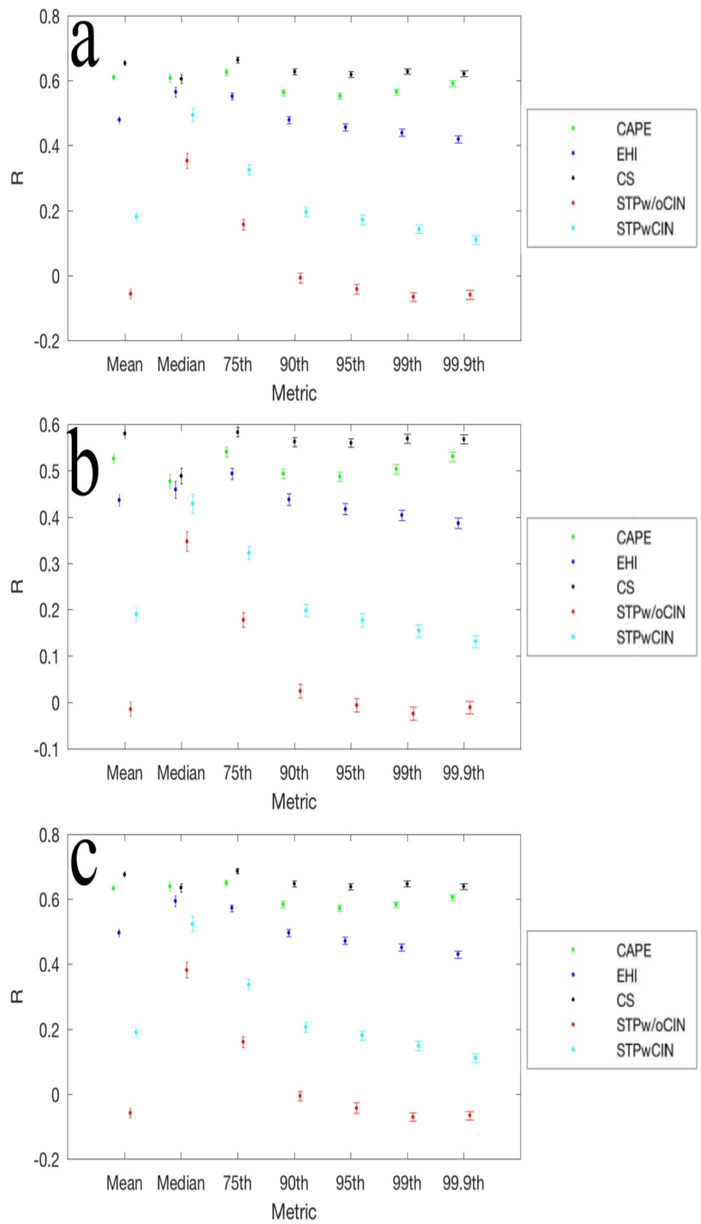

Figure 3.15. As in Figure 3.14, but for R = 120 km.

Interannual Variations

Maps of the correlation coefficients between the interannual variations of 99th percentile of the three proxies and tornado density are displayed in Figures 3.16 and 3.17. The correlation coefficients are very low across this area of the United States for all of the proxies and there are sparse locations of highly correlated areas. The proxies share very similar spacial distributions of correlation coefficients.

Correlation coefficients between interannual tornado activity metrics and our statistics for each environmental proxy are displayed in Figure 3.18. Beginning with EF/F1+ (Figure 3.18a), for CAPE, the mean is outperformed by the 99^{th} and the 99.9^{th} percentiles. For CS and EHI, the mean is statistically similar to the highest correlated extreme percentiles. For both forms of STP, the mean is the highest correlated metric, while the extreme percentiles' correlations are lower by 8-40% depending on statistic. When EF/F2+ tornadoes are used (Figure 3.18b), for CS and STP, the qualitative results are largely unchanged. For EHI, the 99.9^{th} percentile has the highest correlation. The correlation coefficients decreased by 21-54% depending on which proxy and statistic. When tornado days are used (Figure 3.18c), the results are largely unchanged. Sensitivity of these results to a larger tornado radius is shown in Figure 3.19. Results are largely insensitive to the use of a larger tornado radius, except as the radius increases from 50 kilometers to 120 kilometers, all correlation coefficients also increase.

3.1 Discussion

Correlation values increase when the smoothing radius is increased from 50 kilometers to 120 kilometers. This could be attributed to environmental conditions being at a much larger scale than a tornado, and a larger counting radius smooths the tornado to cover a larger fraction of the environment. Interestingly, STP is not the highest correlated predictor at all time scales. Of the three proxies, STP is the only proxy specifically designed for tornadoes so one would assume that it should out-

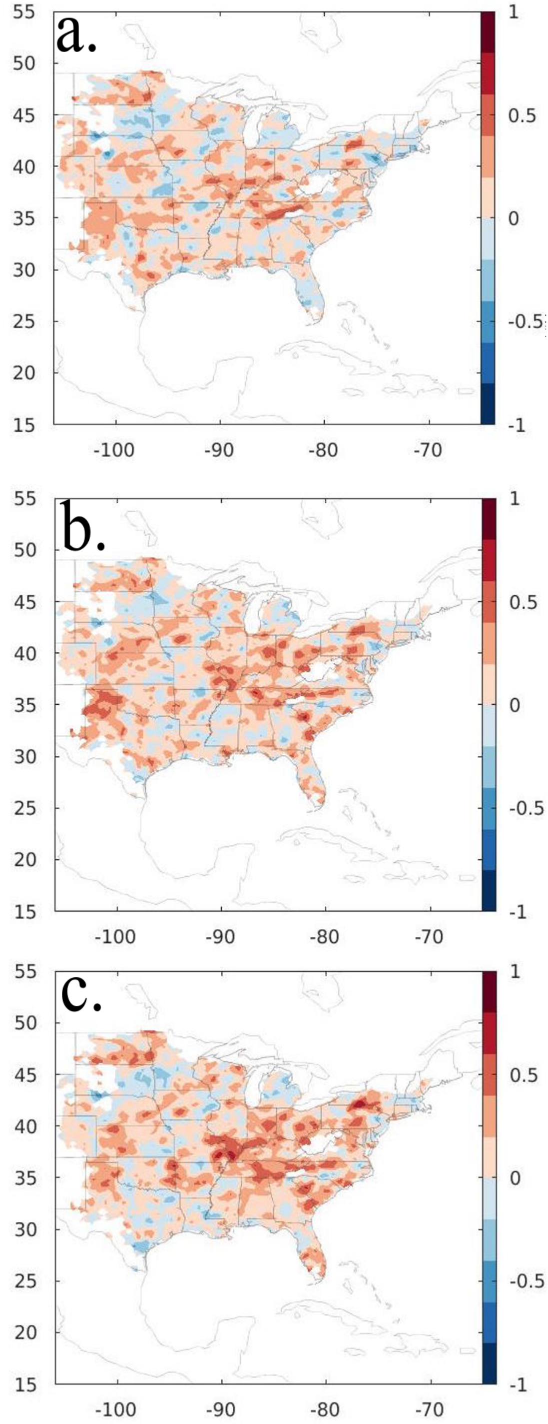

Figure 3.16. As in Figure 3.8, but for the interannual variations.

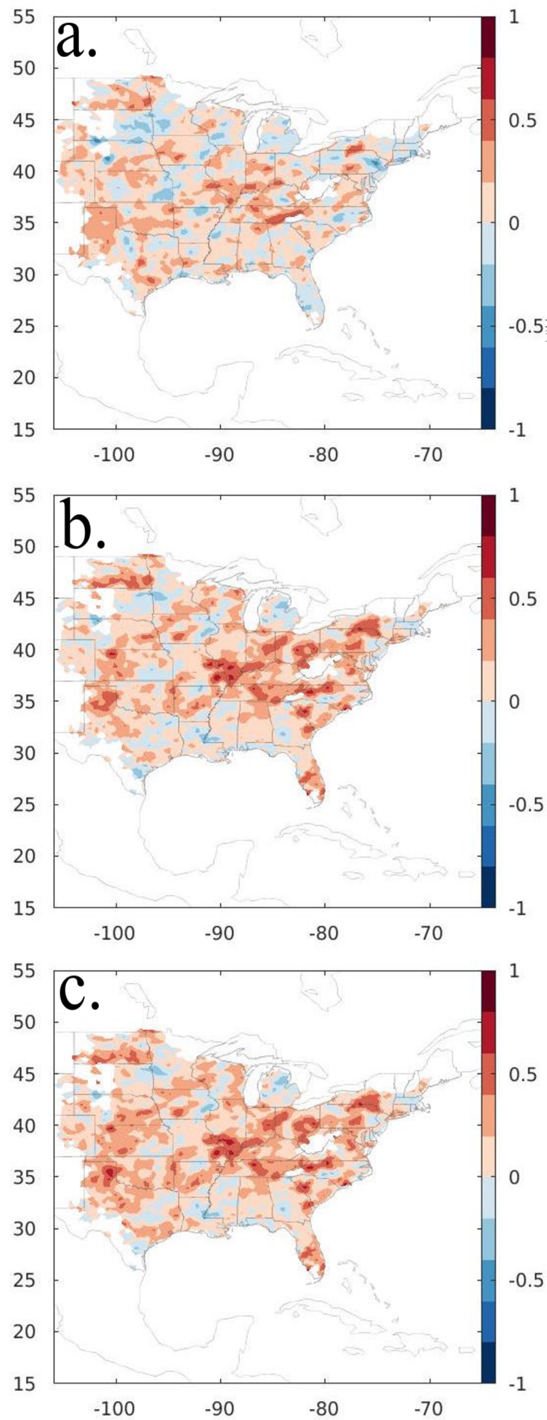

Figure 3.17. As in Figure 3.9, but for the interannual variations.

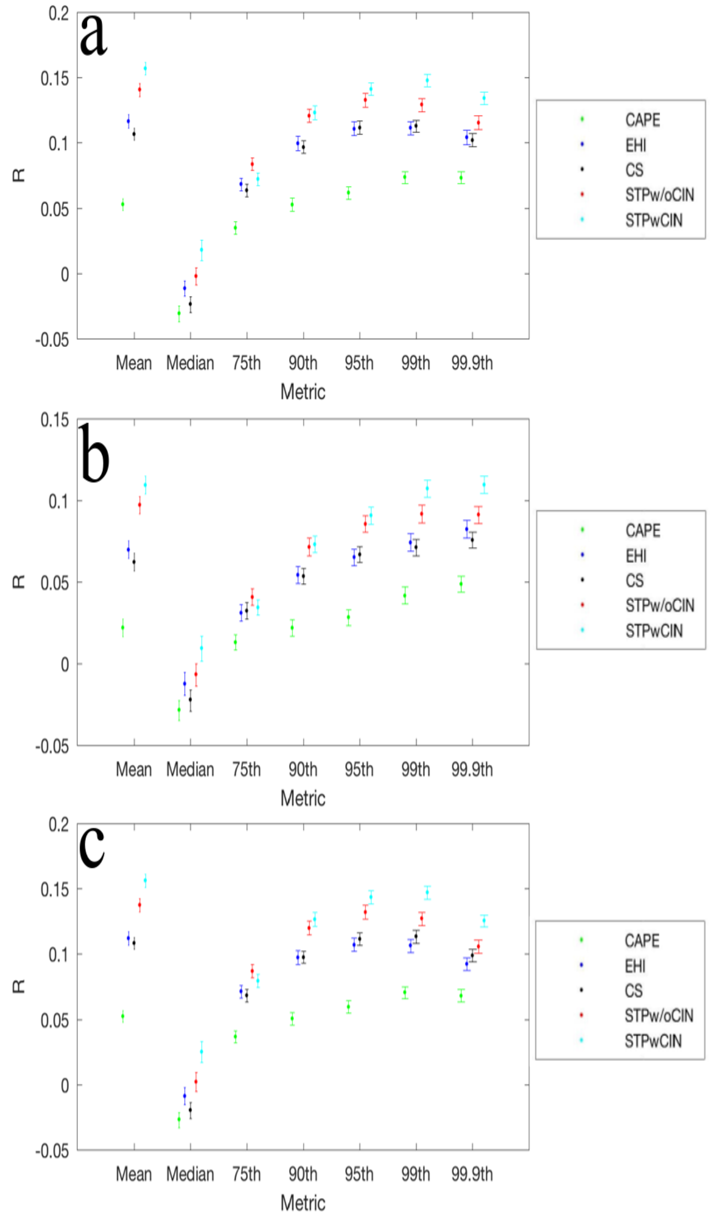

Figure 3.18. Scatter plots between interannual proxy metrics and correlation coefficients using different methods for counting tornadoes for R = 50 km for: a. all tornadoes rated EF/F 1 and greater, b. all significant tornadoes (EF/F2+), and c. all tornado days. CAPE is green, EHI is blue, CS is black, STP without using CIN is red, and STP including CIN is turquoise. Error bars represent the 95 percent confidence intervals derived from bootstrapped re-sampling of the data.

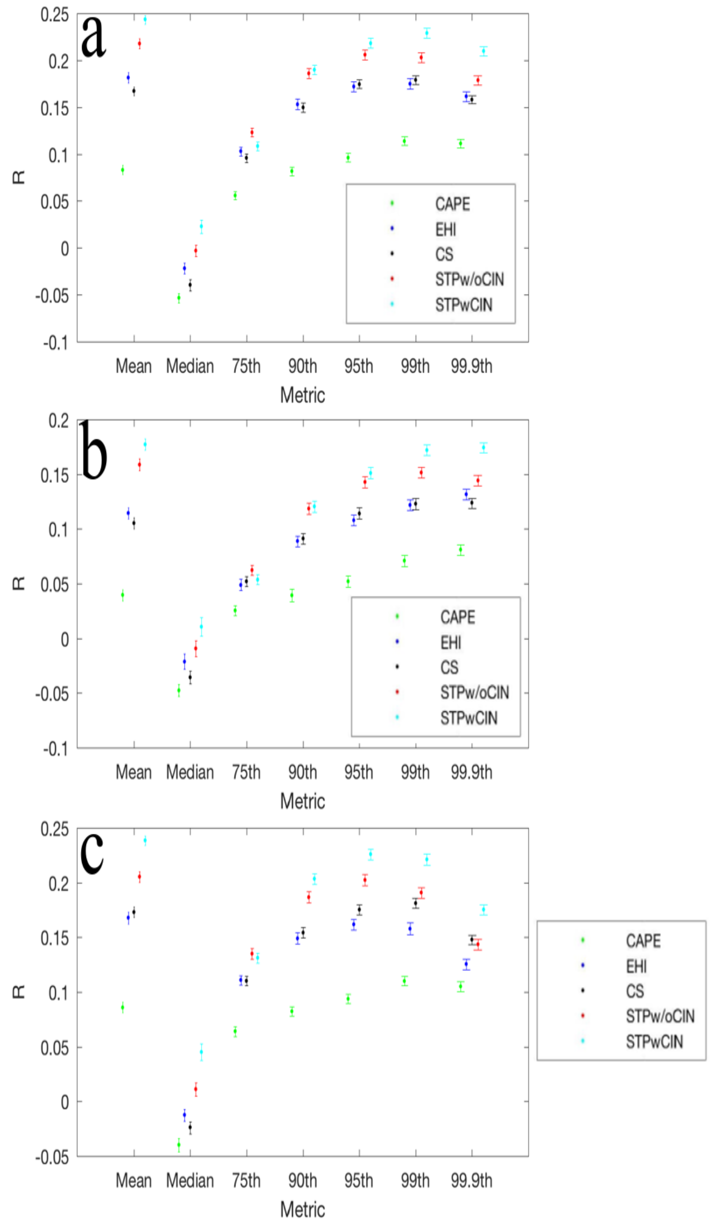

Figure 3.19. As in Figure 3.18, but for R = 120 km.

perform the other two. STP outperforms the other two proxies at three of our four timescales: full climatology, the seasonal cycle, and for interannual variations. It is significantly outperformed by both of the other proxies in the diurnal cycle.

CS and CAPE outperform EHI and both forms of STP at the diurnal cycle. This could be due to the SRH component of EHI and STP. From the map of correlation coefficients over the United States (Figure 3.12), both STP and EHI have low and negative correlation coefficients over the Great Plains. This could be due to the diurnal cycle of the low level jet producing higher SRH values over night, which could boost EHI and STP values during any nocturnal convection events. In the seasonal cycle, the proxies perform comparably worse in the Southeast. This could be explained by cold season tornadoes being more common to the southeastern United States [39]. Childs et al. [39] also found that these tornadic storms are typically accompanied by high shear, low CAPE environments. High shear, low CAPE environments tend to be adverse to calculations of high values of the three proxies used in this study. The correlation coefficients for interannual variability all fall below a value of 0.25. This could possibly be explained by the peak areas in yearly proxies staying confined mainly to the great planes from year-to-year, but tornado spatial distributions show much more variability on an interannual basis.

The use of significant tornadoes lowered the correlation coefficient in every sensitivity test. This could be caused by the number of significant tornadoes. Significant tornadoes are much less common. Removing EF/F1s removes 69.5% of tornadoes from the dataset. In most cases, correlation coefficients increase when using tornado days. The outlier is in interannual variations where the correlation coefficients are statistically similar.

3.2 Conclusions

Using 37 years of reanalysis data and tornado data, we calculate spacial distributions of tornado density and of several statistics: two standard central tendency statistics (mean and median) and five extreme percentiles (75^{th}, 90^{th}, 95^{th}, 99^{th}, and 99.9^{th}), of the three severe weather environmental proxies (CS, EHI, and STP) for the full climatology and across three dominant climate timescales: diurnal, seasonal, interannual. We tested each statistical measure of these common environmental proxies for severe weather favorability to assess their capacity to predict climatological tornado activity over the continental United States by performing spatial correlations between the different distributions.

We found the following key results:

1. The mean is a suitable statistic when compared to the extreme percentiles when used with the full climatology for EHI and STP calculated without CIN, the diurnal cycle for CAPE and CS, and in interannual variations for all proxies except CAPE.

2. The mean is outperformed by the 99^{th} and 99.9^{th} percentiles in the seasonal cycle by 36 and 40% for CAPE, 45 and 53% for CS, 21 and 30% for EHI, 8 and 10% for STP calculated without CIN, and 8 and 5% for STP calculated with CIN.

3. STP is the preferred proxy for the seasonal cycle and interannual variations.

4. STP calculated with 0-1 km SRH and CIN either outperformed or showed marginal change to using STP calculated with 0-3 km SRH and without CIN.

5. CS is the preferred proxy for the diurnal cycle.

6. Qualitative results are similar when using significant tornadoes or tornado days, but using significant tornadoes results in lower correlation coefficients and using tornado days results in higher correlation coefficients.

7. The sensitivity tests showed that the use of tornado days results in higher correlation coefficients than using EF/F1+ tornadoes for the full climatology by 10-40%, the seasonal cycle by 3-8%, and the diurnal cycle by 3-4%.

8. Qualitative results are similar when using a larger tornado radius.
9. These qualitative results are robust to the increased radius value and the only significant change is that correlation coefficients increase with the increased radius.
10. Every statistic of each proxy showed very low predictability of the interannual variations in tornado activity.

3.2.1 Potential Limitations

Our datasets possess important limitations. We use reanalysis data which should not be taken as ”real” values, but as the model's best guess for the environmental conditions. The NARR dataset gives us data with high temporal and spatial resolution [55], as opposed to environmental soundings which would be more accurate. However, environmental soundings are only available at discreet locations regularly twice a day: at 0z and 12z [59]. The tornado historical record carries numerous uncertainties [60] that may themselves vary with time, especially before Doppler technology in the 1990s. This stems from tornadoes requiring an observer to be present at the time and location of the event. With increasing population, better technology, and more awareness, the annual count of tornadoes has increased throughout the years [61]. We attempt to minimize this aspect by removing tornadoes rated F/EF 0 as Verbout et al. [61] and Agee et al. [62] found that the number of tornadoes rated greater than F/EF 0 from 1954-2003 was fairly consistent, showing that the increase in count is, in large part, due to the increased reportings of weak (F/EF 0) tornado events. We also only work with proxies for environmental favorability, and we do not account for variability in factors that initiate convection.

3.2.2 Avenues of Future Work

Possible areas of future study are to break these results down by region (northern plains, southern plains, midwest, northeast, southeast). Examinations of the interannual variability associated with the climate time scales used in this study could also be performed, such as how the correlations change based on a year-to-year analysis of the seasonal cycle. Another area of possible research is to further test the smoothing radii. Performing similar analyses while grouping the seasons together (DJF, MAM, JJA, SON) is another possible extension of this research. These results can be used with climate models to determine areas where tornado activity could potentially increase. Figure 2.3 (c, f, and i) shows slight increases in the extreme percentiles of the three proxies over the 37 year period for Lafayette, Indiana. Climate models could be run to investigate how these trends could continue to evolve in the future to see if the increase continues, if the values level off, or after leveling off, if the values begin to decrease. This analysis could be extended over the rest of the continental United States to see if this upward trend applies to more areas. This could be run for different concentrations of CO_2 and using different methods of calculating these proxies as well. These results could be used with global reanalysis data to help improve our understanding of severe weather in other parts of the world. These analyses could be done using other forms of reanalysis data to test how robust these results are. This research could be focused onto smaller areas by applying similar analyses on novel datasets, such as radar data or environmental sounding data. These results could be used with global climate models to investigate how global severe weather tendencies could evolve in a future climate state. Another possible avenue is to apply a gaussian smoother to the tornado distribution similar to the methods by Gensini et al. [42].

3.2.3 Broader Impacts and Utility

This research could prove invaluable to both researchers and society as a whole. It can impact how regional and global climate models predict tornado activity in a warmer, future climate. Better predictions allow society to be more prepared in the future if locations of peak tornado activity are to shift. This could lead to better risk assessment that could protect countless lives and property. One such example is in building preparedness. Construction codes could need to be amended and strengthened in order for buildings to survive an increased threat of severe weather. Or families, when selecting a new home, would differently weigh the benefits of a basement and insurance for severe weather. And likewise, insurance companies can change what policies are offered in areas to match the changing climate.

REFERENCES

REFERENCES

[1] National Weather Service. Severe weather definitions, 2018. [Available online at http://www.spc.noaa.gov/faq/#4.1].

[2] National Weather Service. Summary of natural hazard statistics for 2015 in the united states, 2016. [Available online at http://www.nws.noaa.gov/om/hazstats/sum15.pdf].

[3] NOAAs National Climate Data Center. U.s. tornado climatology, 2016. [Available online at https://www.ncdc.noaa.gov/climate-information/extreme-events/us-tornado-climatology].

[4] Morris L Weisman and Joseph B Klemp. The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Monthly Weather Review*, 110(6):504–520, 1982.

[5] Bryan T Smith, Richard L Thompson, Jeremy S Grams, Chris Broyles, and Harold E Brooks. Convective modes for significant severe thunderstorms in the contiguous united states. part i: Storm classification and climatology. *Weather and Forecasting*, 27(5):1114–1135, 2012.

[6] William A Gallus Jr, Nathan A Snook, and Elise V Johnson. Spring and summer severe weather reports over the midwest as a function of convective mode: A preliminary study. *Weather and Forecasting*, 23(1):101–113, 2008.

[7] Joseph M Schoen and Walker S Ashley. A climatology of fatal convective wind events by storm type. *Weather and forecasting*, 26(1):109–121, 2011.

[8] Alan R Moller, Charles A Doswell III, Michael P Foster, and Gary R Woodall. The operational recognition of supercell thunderstorm environments and storm structures. *Weather and Forecasting*, 9(3):327–347, 1994.

[9] Robert H Johns and Charles A Doswell III. Severe local storms forecasting. *Weather and Forecasting*, 7(4):588–612, 1992.

[10] Greg Leander Dial and JP Racy. Forecasting short term convective mode and evolution for severe storms initiated along synoptic boundaries. In *Preprints, 22d Conf. on Severe Local Storms, Hyannis, MA, Amer. Meteor. Soc., CD-ROM A*, volume 11, 2004.

[11] P. Markowski and Y. Richardson. *Mesoscale Meteorology in Midlatitudes*. Advancing Weather and Climate Science. Wiley, 2010.

[12] Robert G Beebe. Tornado proximity soundings. *Bulletin of the American Meteorological Society*, pages 195–201, 1958.

[13] Charles A Doswell. Severe convective stormsan overview. In *Severe convective storms*, pages 1–26. Springer, 2001.

[14] Erik N Rasmussen and David O Blanchard. A baseline climatology of sounding-derived supercell andtornado forecast parameters. *Weather and Forecasting*, 13(4):1148–1164, 1998.

[15] Grant William Petty. *A first course in atmospheric thermodynamics*. Sundog Publishing, 2008.

[16] RL Thompson, R Edwards, and JA Hart. Evaluation and interpretation of the supercell composite and significant tornado parameters at the storm prediction center,[w:] 21st conference on severe local storms. *San Antonio*, 2002.

[17] M.L. Salby, S. MURRYL, R.A. Pielke, and R. Dmowska. *Fundamentals of Atmospheric Physics*. International geophysics series. Academic Press, 1996.

[18] J.R. Holton and G.J. Hakim. *An Introduction to Dynamic Meteorology*. Academic Press. Elsevier, 2013.

[19] Harold E Brooks. Severe thunderstorms and climate change. *Atmospheric Research*, 123:129–138, 2013.

[20] James Espy. Essays on meteorology, no. iv: North east storms, volcanoes, and columnar clouds. *Journal of the Franklin Institute*, 22(4):239–246, 1836.

[21] David M Romps. Exact expression for the lifting condensation level. *Journal of the Atmospheric Sciences*, 74(12):3891–3900, 2017.

[22] Harold E Brooks, James W Lee, and Jeffrey P Craven. The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmospheric Research*, 67:73–94, 2003.

[23] Jeffrey P Craven, Harold E Brooks, and John A Hart. Baseline climatology of sounding derived parameters associated with deep, moist convection. *Natl. Wea. Dig*, 28(1):13–24, 2004.

[24] John Terrence Allen, David J Karoly, and Graham A Mills. A severe thunderstorm climatology for australia and associated thunderstorm environments. 2011.

[25] Jacob T Seeley and David M Romps. The effect of global warming on severe thunderstorms in the united states. *Journal of Climate*, 28(6):2443–2458, 2015.

[26] JA Hart and W Korotky. The sharp workstation v1. 50 users guide. *NOAA/National Weather Service*, 1991.

[27] Erik N Rasmussen. Refined supercell and tornado forecast parameters. *Weather and Forecasting*, 18(3):530–535, 2003.

[28] Richard L Thompson, Roger Edwards, John A Hart, Kimberly L Elmore, and Paul Markowski. Close proximity soundings within supercell environments obtained from the rapid update cycle. *Weather and Forecasting*, 18(6):1243–1261, 2003.

[29] Richard L Thompson, Roger Edwards, and Corey M Mead. An update to the supercell composite and significant tornado parameters. In *Preprints, 22nd Conf. on Severe Local Storms, Hyannis, MA, Amer. Meteor. Soc. P*, volume 8, 2004.

[30] Richard L Thompson, Bryan T Smith, Jeremy S Grams, Andrew R Dean, and Chris Broyles. Convective modes for significant severe thunderstorms in the contiguous united states. part ii: Supercell and qlcs tornado environments. *Weather and forecasting*, 27(5):1136–1154, 2012.

[31] DL Kelly, JT Schaefer, RP McNulty, CA Doswell III, and RF Abbey Jr. An augmented tornado climatology. *Monthly Weather Review*, 106(8):1172–1183, 1978.

[32] Michael K Tippett, Adam H Sobel, and Suzana J Camargo. Association of us tornado occurrence with monthly environmental parameters. *Geophysical research letters*, 39(2), 2012.

[33] Harold E Brooks, Gregory W Carbin, and Patrick T Marsh. Increased variability of tornado occurrence in the united states. *Science*, 346(6207):349–352, 2014.

[34] James B Elsner, Svetoslava C Elsner, and Thomas H Jagger. The increasing efficiency of tornado days in the united states. *Climate Dynamics*, 45(3-4):651–659, 2015.

[35] Aiguo Dai and Kevin E Trenberth. The diurnal cycle and its depiction in the community climate system model. *Journal of Climate*, 17(5):930–951, 2004.

[36] TN Carlson and FH Ludlam. Conditions for the occurrence of severe local storms. *Tellus*, 20(2):203–226, 1968.

[37] TN Carlson, SG Benjamin, GS Forbes, and YF Li. Elevated mixed layers in the regional severe storm environment: Conceptual model and case studies. *Monthly weather review*, 111(7):1453–1474, 1983.

[38] Harold E Brooks, Charles A Doswell III, and Michael P Kay. Climatological estimates of local daily tornado probability for the united states. *Weather and Forecasting*, 18(4):626–640, 2003.

[39] Samuel J Childs, Russ S Schumacher, and John T Allen. Cold-season tornadoes: Climatological and meteorological insights. *Weather and Forecasting*, 33(3):671–691, 2018.

[40] John A Long, Paul C Stoy, and Tobias Gerken. Tornado seasonality in the southeastern united states. *Weather and Climate Extremes*, 20:81–91, 2018.

[41] Ernest Agee, Jennifer Larson, Samuel Childs, and Alexandra Marmo. Spatial redistribution of us tornado activity between 1954 and 2013. *Journal of Applied Meteorology and Climatology*, 55(8):1681–1697, 2016.

[42] Vittorio A Gensini and Harold E Brooks. Spatial trends in united states tornado frequency. *npj Climate and Atmospheric Science*, 1(1):38, 2018.

[43] Michael K Tippett, Chiara Lepore, and Joel E Cohen. More tornadoes in the most extreme us tornado outbreaks. *Science*, 354(6318):1419–1423, 2016.

[44] Robert J Trapp, Noah S Diffenbaugh, and Alexander Gluhovsky. Transient response of severe thunderstorm forcing to elevated greenhouse gas concentrations. *Geophysical Research Letters*, 36(1), 2009.

[45] Noah S Diffenbaugh, Robert J Trapp, and Harold Brooks. Does global warming influence tornado activity? *Eos, Transactions American Geophysical Union*, 89(53):553–554, 2008.

[46] Cameron C Lee. Utilizing synoptic climatological methods to assess the impacts of climate change on future tornado-favorable environments. *Natural hazards*, 62(2):325–343, 2012.

[47] Vince Agard and Kerry Emanuel. Clausius–clapeyron scaling of peak cape in continental convective storm environments. *Journal of the Atmospheric Sciences*, 74(9):3043–3054, 2017.

[48] Kimberly A Hoogewind, Michael E Baldwin, and Robert J Trapp. The impact of climate change on hazardous convective weather in the united states: Insight from high-resolution dynamical downscaling. *Journal of Climate*, 30(24):10081–10100, 2017.

[49] Mengqian Lu, Michael Tippett, and Upmanu Lall. Changes in the seasonality of tornado and favorable genesis conditions in the central united states. *Geophysical research letters*, 42(10):4224–4231, 2015.

[50] Victor A Gensini, Thomas L Mote, and Harold E Brooks. Severe-thunderstorm reanalysis environments and collocated radiosonde observations. *Journal of Applied Meteorology and Climatology*, 53(3):742–751, 2014.

[51] John T Allen and David J Karoly. A climatology of australian severe thunderstorm environments 1979–2011: inter-annual variability and enso influence. *International Journal of Climatology*, 34(1):81–97, 2014.

[52] Austin T King and Aaron D Kennedy. North american supercell environments in atmospheric reanalyses and ruc-2. *Journal of Applied Meteorology and Climatology*, (2018), 2018.

[53] John T Allen, David J Karoly, and Kevin J Walsh. Future australian severe thunderstorm environments. part i: A novel evaluation and climatology of convective parameters from two climate models for the late twentieth century. *Journal of Climate*, 27(10):3827–3847, 2014.

[54] Fedor Mesinger, Geoff DiMego, Eugenia Kalnay, Kenneth Mitchell, Perry C Shafran, Wesley Ebisuzaki, Dušan Jović, Jack Woollen, Eric Rogers, Ernesto H Berbery, et al. North american regional reanalysis. *Bulletin of the American Meteorological Society*, 87(3):343–360, 2006.

[55] NOAA ESRL Physical Sciences Division: Data Management. Ncep north american regional reanalysis: Narr, 2018. [Available online at `https://www.esrl.noaa.gov/psd/data/gridded/data.narr.html`].

[56] NOAA/NWS/SPC. Actual tornadoes, 2016. accessed from NOAA/NWS/SPC website at http://www.spc.noaa.gov/wcm/.

[57] Lori A Schultz and Daniel J Cecil. Tropical cyclone tornadoes, 1950–2007. *Monthly Weather Review*, 137(10):3471–3484, 2009.

[58] Landsea, C. W., and Franklin, J. L. Atlantic hurricane database, 2013. accessed from NOAA/NWS/NHC website at https://www.nhc.noaa.gov/data/.

[59] National Weather Service Forecast Office, Reno Nevada. Upper-air program virtual tour national weather service weather forecast office - reno , nv, 2018. [Available online at `https://www.wrh.noaa.gov/rev/tour/UA/introduction.php`].

[60] Charles A Doswell III and Donald W Burgess. On some issues of united states tornado climatology. *Monthly Weather Review*, 116(2):495–501, 1988.

[61] Stephanie M Verbout, Harold E Brooks, Lance M Leslie, and David M Schultz. Evolution of the us tornado database: 1954–2003. *Weather and Forecasting*, 21(1):86–93, 2006.

[62] Ernest Agee and Samuel Childs. Adjustments in tornado counts, f-scale intensity, and path width for assessing significant tornado destruction. *Journal of Applied Meteorology and Climatology*, 53(6):1494–1505, 2014.