Unsupervised machine learning of tornado-producing storms in the southeastern United States

A Thesis Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Morgan R. Steckler

August 2023

Abstract

The east-southeastern US is uniquely affected by storm and tornado-related damages, costs, injuries, and deaths. Based on doppler radar, satellite, and modeled data, previous research sought to understand these different types of storms that produce strong tornadoes. Many approaches to storm classification are time intensive, complex, and vary significantly across the literature. The purpose of this work is to (1) explore the radar-derived data structure and spread of strong tornado-producing mesoscale storms in the east-southeastern US; (2) use K-Means unsupervised machine learning methods to elucidate clusters (storm types) and clustering attributes; and (3) assess the utility of K-Means as a storm typing algorithm. Convective and stratiform strength, length, width, shape, and area, as well as the position of stratiform rain relative to convection, were evaluated using principal component analysis and K-Means clustering. The results show that convective strength and length attributes best explained variance in the dataset and minimized cluster overlap. While other attributes wielded explanatory power, they did not separate into distinct clusters. After testing K-Means on 2 through 8 clusters, the intensity and shape attributes generated three strong tornado-producing storm types: Weaker Linear (n=146), Stronger Cellular (n=222), and Weaker Cellular (n=292). These storm types were most like those in the literature and yielded the highest K-Means Silhouette score (0.41). It was found that 54% of storms were placed with less than 80% certainty, which emphasizes that storm types belong on a continuum and are not easily divided into groups. However, climatological analyses of the storms from 2000–2020 reveal that Weaker Linear, Stronger Cellular, and Weaker Cellular storm types produce similar spatial and temporal signals to those of seminal storm classification literature. This work can provide a framework for future storm taxonomies and highlights the value of machine learning in place of subjective, manual storm classification.

Table of Contents

CHAPTER 1. INTRODUCTION AND LITERATURE REVIEW 1
1.1 Storm Environments 2
1.2 Storm Types 4
1.3 Classification 5
1.3.1 Supervised 5
1.3.2 Unsupervised 9
1.4 Research Questions 11
CHAPTER 2. DATA AND METHODS 12
2.1 Data 12
2.2 Methods 15
2.2.1 Pre-Processing 15
2.2.2 Clustering 18
CHAPTER 3. RESULTS AND DISCUSSION 22
3.1 K-Means Clustering 22
3.2 C-Means Analysis 30
3.3 Threshold Approach 32
3.4 Storm Type Climatology 34
3.4.1 Spatial climatology 34
3.4.2 Seasonal and diurnal climatology 36
CHAPTER 4. CONCLUSION 44
LIST OF REFERENCES 47
VITA 56

List of Tables

Table 1.1 Literature summary of storm classification 6

Table 2.1 Radar summary statistics 16

Table 2.2 Top-ranking attributes for each PC (first PCA run) 19

Table 2.3 Top-ranking attributes for each PC (second PCA run)..... 20

Table 3.1 Strength descriptive statistics for K-Means storm types (z-scores)..... 29

Table 3.2 Length descriptive statistics for K-Means storm types (z-scores)..... 29

Table 3.3 FCM results confusion matrix 31

Table 3.4 Strength descriptive statistics for threshold storm types (z-scores)..... 33

Table 3.5 Length descriptive statistics for threshold storm types (z-scores)..... 33

Table 3.6 Chi-square and Cramer's phi for storm type, season, and time of day 41

Table 3.7 Chi-square test observed, expected, and standardized residual *R* values 41

List of Figures

Figure 2.1 Study area, radar locations, and tornado touchdowns 13

Figure 3.1 Three-cluster results using five attributes..... 23

Figure 3.2 Three-cluster results using four attributes 24

Figure 3.3 Three-cluster results using three attributes..... 25

Figure 3.4 Three-cluster results using two attributes..... 27

Figure 3.5 Random image samples of the three K-Means storm types 28

Figure 3.6 Storm type KDE maps of the East-Southeast US (2000–2020)..... 35

Figure 3.7 Smith et al. (2012) comparison of KDE maps 37

Figure 3.8 Seasonal and diurnal counts of the three K-Means storm types..... 38

Figure 3.9 Monthly peaks of the three K-Means storm types..... 40

Figure 3.10 Smith et al. (2012) comparison of annual peaks 42

CHAPTER 1. INTRODUCTION AND LITERATURE REVIEW

Since 1980, severe storms are the third costliest hazard behind hurricanes and drought. In general, the southern, central, and southeastern US experiences a higher frequency of billion-dollar events and subsequent fatalities. In the US, ten or more billion-dollar events have occurred for the past seven years in a row and have been increasing over time (NOAA, 2022; NOAA, 2023). Social factors, such as old infrastructure, dated building codes, and rapidly expanding urban spaces, play a significant role in this recent upward spike in expenses, and also deaths, from natural hazards (Ashley, 2007; Strader et al., 2017). Coupled with 21st century socio-economic changes are shifts in the area and frequency of tornado-producing convective storms and outbreaks in the central and southeastern US.

Tornado-producing storms occur most often east of the Rocky Mountains and particularly in the Great Plains, an area where atmospheric conditions are most frequently favorable for tornado production. Such conditions include abundant low-level moisture availability, atmospheric instability (rising air), and moderate to high vertical and directional shear (MRCC, n.d.). However, portions of the Southeast experience more tornado fatalities than the Great Plains (Ashley, 2007). Additionally, recent research points to a spatial shift in the occurrence of tornadoes across the US. Specifically, findings suggest that tornadoes (Cao et al., 2021), and the conditions that favor them (Gensini & Brooks, 2018), are shifting away from the Great Plains toward other regions, such as the Southeast. Beyond individual events, an increasing occurrence of tornado outbreaks—many tornadoes being produced by a single storm system—is contributing to a decrease in the spatiotemporal dispersal of tornadoes (Moore & McGuire, 2019).

In addition to spatial changes in tornadoes are possible changes in tornado frequency. Several researchers found that tornado days are decreasing in frequency, but when there is a tornado day, we are more likely to see an outbreak, which is when a system produces several tornadoes (Brooks et al., 2014; Elsner et al., 2015; Moore & McGuire, 2019; Tippett et al., 2016). Tippett et al. (2016) also found that the number of tornadoes occurring in an outbreak are increasing, especially for “extreme” outbreaks of 12 or more tornadoes, though Elsner et al.

(2014) suggested that an average outbreak consists of 12 to 13 tornadoes. As for why these changes could be occurring, research has attempted to link it to decadal oscillations (M.C. Brown & Nowotarski, 2020; V.M. Brown et al., 2016; S. Lee et al., 2013; Nouri et al., 2021) and anthropogenic climate change (Balling & Cerveny, 2003; Diffenbaugh et al., 2013; C.C. Lee, 2012; Strader et al., 2017) with no final agreement on the leading causes. Regardless, the storms that produce tornadoes and the environments under which they develop are changing spatially and temporally.

Thus, researchers seek to better understand tornado-producing storms across space and time. Certain storms exhibit predictable behavior, evolution, and hazards. These characteristics are observable using extensive Doppler radar networks such as WSR-57, WSR-74, and currently WSR-88D, as well as satellite imagery and ground observations (Whiton et al., 1998). Thus, researchers have long sought to classify predominant storm patterns. However, classifying storms manually—i.e., looking at several radar-derived images of a storm and identifying it by name—is time consuming, inconsistent, and subject to the knowledge and experience of the meteorological professional. As a result, researchers are increasingly turning toward machine learning to classify thousands of remotely sensed storms quickly and more objectively. The goal of this thesis is to understand the underlying data characteristics of storms that produced a rank of 2 or higher on the destruction-based Enhanced Fujita (EF) tornado rating scale in the eastern and southeastern US, and to determine the most useful attributes for clustering these storms into unique groups. Results from this work will build on current climatological studies of tornadic storms in the east-southeastern US and, secondly, will contribute to future decision-making when attempting to classify strong, tornado-producing storms.

1.1 Storm Environments

The basic environmental requirements for storm formation and thus tornadogenesis are moisture; atmospheric instability, which is the tendency for air to rise; and vertical wind shear, which are changes in the directionality of wind with height (Brandes, 1978; Jackson & Brown, 2009; Lemon & Doswell, 1979; Markowski & Richardson, 2013; Schumacher & Rasmussen,

2020; R. K. Smith & Leslie, 1978). Some ways to quantify instability include vertical lapse rates, convective available potential energy (CAPE) at different heights, the lifted condensation level, and the level of free convection. Clear days are generally characterized by little CAPE, light winds, and low shear, while high CAPE and shear suggest stormy weather, and these properties may be a good discriminator between non-tornadic storms and tornadic storms (Craven & Brooks, 2004; Jackson & Brown, 2009). Diurnal, seasonal, annual, and decadal changes in temperature and precipitation can also drive tornado formation. Most tornadoes spawn around or after sundown due to the lag time in daily heating from the sun, and though tornadoes can occur all year, there are known seasonal peaks and declines. On a larger time-scale, teleconnections that influence tornado variability in the US include the Southern, Atlantic, North Atlantic, Atlantic Multidecadal, El Niño, and Pacific Decadal oscillations (M.C. Brown and Nowotarski, 2020; V.M. Brown, 2016; S. Lee et al., 2013; Nouri et al., 2021). Some specific moisture sources for the east-southeastern US include the Gulf, Caribbean, and the North Atlantic, and their contributions also vary inter- and intra-annually and can be linked to teleconnections (Molina & Allen, 2020). Studies have also explored the role of anthropogenic climate change in the modern shifts of tornado frequency, dispersion, and intensity (Balling & Cerveny, 2003; Diffenbaugh et al., 2013; C.C., Lee 2012; Strader et al., 2017).

The majority of reported tornadoes in the US occur across the entire central and eastern US, from the foothills of the Rockies to the Atlantic seaboard. However, storm types, tornado characteristics, activity, and impacts vary across such a large region. In the Great Plains, tornado-producing storms are characterized by having substantial CAPE and moderate to high shear, the latter being capable of producing significant tornadoes (Brooks et al., 2003; Thompson et al., 2003). In contrast, the east-southeastern US experiences more high-shear low-CAPE storms that produce tornadoes than the Plains (Schneider et al., 2006; Sherburn & Parker, 2014; Smith et al., 2012). Also in contrast to the Plains, the Southeast experiences a peak in tornado activity in both spring and fall rather than only spring (Ashley, 2007; Gagan et al., 2010; Jackson & Brown, 2009; Long et al., 2018). And though the Great Plains has the highest frequency of strong and violent tornadoes according to Ashley (2007) and Concannon et al. (2000), the Southeast experiences a higher number of injuries, and fatalities, as well as higher costs (Ashley, 2007;

Gagan et al. 2010; Grazulis et al., 1993; Strader et al., 2017). Thus, it is important for researchers to understand the storms that are producing such tornadoes in the east-southeastern US.

1.2 Storm Types

Rainfall can be generalized into two categories, stratiform and convective, which may both produce weather-related hazards (Anagnostou, 2004). Lower elevation clouds that are usually horizontally developed produce stratiform rain events that are lighter and longer sustained than their convective counterparts (NOAA, 2009b). Convective rain often accompanies vertically developed cloud structures that may produce thunderstorms, heavy bursts of precipitation, lightning, and other related severe weather. On radar images, convective areas can be distinguished as groups of pixels at or above a threshold of decibels of Z (dBZ), which is a logarithmic scale of radar reflection intensity (NOAA, 2009a).

Further, meteorologists can divide storms into groups based on the organization of cells, which are groups of high-reflectivity pixels loosely representing convection. They might be alone or single-celled, multi-celled and linear, or multi-celled and non-linear (Baldwin et al., 2005; Jergensen et al., 2019; B.T. Smith et al., 2012). Some researchers incorporate precipitation attributes into their classification schemes, such as high or low precipitation or where the mass of precipitation is positioned—next to, in front of, or behind the convection (Doswell & Burgess, 1993; Gagne et al., 2009; Gallus et al., 2008; Pacey et al., 2021; Parker & Johnson, 2000; Schiesser et al., 1995). One might also consider the intensity of a storm via the presence of a mesocyclone, a storm-scale region of rotation that can be relatively small or large, and weak or intense (Edwards et al., 2012; Gagne et al., 2009; Hu et al., 2019; Johnson et al., 1998; Lack & Fox, 2012; B.T. Smith et al., 2012). A storm can also be classified by its length or shape, whether that is defined by how long its lifetime is, how long its major and minor axes are, how elliptical or eccentric the precipitation is, and the direction or speed of propagation (Gallus et al., 2008; Haberlie & Ashley, 2018; B.T. Smith et al., 2012). Lastly, some storm classifications incorporate the near-storm environment instead of—or in addition to—radar-derived information (Agee & Jones, 2009; Jergensen et al., 2019; Kolodziej Hobson et al., 2012; Lack & Fox, 2012;

Miller & Mote, 2018; Shield & Houston, 2022). Essentially, storms can be classified in a multitude of ways that are dependent on the professional opinion and scope of research.

According to Smith et al. (2012), across the US, single cells are most common in the plains, linearly shaped storms occur most often in the Mississippi valley and Tennessee, and particularly long, persistent, linearly shaped storms—often referred to as quasi-linear convective systems (QLCSs)—are persistent across the Ohio, Tennessee, and Mississippi valleys. In general, linear storms often occur in the fall, clustered storms or single cells occur more frequently in the summer, and linear systems are most prevalent in the winter (V.M. Brown et al., 2016; Smith et al., 2012). Tornadoes often occur at night, and of those nighttime, tornado-producing storms, QLCSs are 3 times more likely to be the culprit (V.M. Brown et al., 2016; Ellis et al., 2020).

1.3 Classification

Machine learning can be subdivided into three major categories: supervised, unsupervised, and reinforcement learning. From my review of 18 relevant articles (Table 1.1), most researchers approaching storm and storm environment classification rely on no machine learning at all or supervised machine learning approaches, while unsupervised learning falls to the wayside. Of 18 papers reviewed, 7 were manual, heuristic classifications from experts only, 9 were a mix of manual classification and supervised machine learning, and 4 papers incorporated unsupervised machine learning. The use of the unsupervised K-Means method to identify and/or track individual cells has been the ongoing influence of the Lakshmanan et al. (2003) paper and is the most popular use of unsupervised learning in storm classification that I have identified in the literature.

1.3.1 Supervised

Supervised machine learning classification is essentially tasking a computer to learn how to categorize objects by looking at examples. Specifically, an expert will select the labels and

Table 1.1 Literature summary of storm classification

	Supervised	Clustering	Manual	Attributes	N Types	Attribute Types
Williams & Abernethy (2008)	x	x	x	radar	4	not severe, isolated supercell, convective line, pulse
Agee & Jones (2009)			x	radar, sounding	15	supercells, QLCS, other
Anagnostou (2004)	x			radar, satellite	2	convective, stratiform
Cui et al. (2021)	x		x	radar, satellite	5	trailing/leading/parallel linear, trailing line, back building, nonlinear
Gagne et al. (2009)	x	x	x	radar	6	isolated pulse, isolated strong, multicell, trailing/leading/parallel linear
Gallus et al. (2008)			x	radar	9	isolated cells, clusters of cells, broken line, none/trailing/leading/parallel squall
Jergensen et al. (2019)	x		x	radar, sounding	3	supercell, QLCS, disorganized
Karpinski (2020)			x	radar	6	supercell, QLCS, tropical cyclone
Kolodziej Hobson et al. (2012)	x	x	x	radar, sounding	4	short-lived convective, supercell, ordinary, convective
Lack & Fox (2012)	x		x	radar, sounding	8	air mass, pulse, small-scale rotating, severe small-scale rotating, linear convective, rotating linear convective, supercell, general
Li et al. (2000)	x		x	radar	2	hail cell, tornado cell
Miller & Mote (2018)	x		x	radar, sounding	2	pulse, non-severe weakly forced thunderstorm
Parker & Johnson (2000)			x	radar	3	trailing/leading/parallel linear, trailing line, back building, nonlinear

Table 1.1 Continued

	Supervised	Clustering	Manual	Attributes	N Types	Attribute Types
Potvin et al. (2022)			x	radar, satellite	6	QLCS, QLCS mesoscale, QLCS ordinary, supercell, supercell cluster, ordinary, other
Shield & Houston (2022)	x		x	sounding	2	supercell vs non-supercell environment
Smith et al. (2012)			x	radar	13	QLCS line, QLCS bow echo, hybrid QLCS, hybrid cell in line, discrete cell/cell in cluster/cell in line, marginal discrete cell/cell in cluster/cell in line, disorganized discrete cell/cell in cluster/cell in line
Thompson et al. (2003)			x	sounding	4	significant tornadic supercell, weak tornadic supercell, non-tornadic supercell, discrete non-supercell
Baldwin et al. (2005)		x	x	radar, ground	5	stratiform, cellular 1 and 2, linear 1 and 2

relevant defining characteristics (attributes) of a population of objects, and manually compile this information into a "training" dataset. The training dataset is then cross validated and tested for accuracy using a set-aside testing dataset. Once the algorithm is properly trained, it can apply the model to the testing dataset. This broad machine learning approach can be highly accurate and fast once the training dataset is developed by hand. However, if the training dataset includes incorrect, missing, or irrelevant data, the algorithm may produce poor results (Kotsiantis, 2007). Training a dataset can also become very time consuming and is dependent on the opinion and knowledge of the trainer. In the case of storm classification, each expert may have a different opinion on how to classify storms based on their personal and professional experiences.

Parker and Johnson (2000) is a heavily cited paper on the organization of mesoscale convective systems in the central US. Their work suggests that linearly shaped, persistent and >100 kilometers storms can be subclassed into three categories based on the position of stratiform rain relative to convective areas using composited radar data (the max pixel value out of every elevation in a volume scan) from May 1996 and 1997. Their dataset was manually classified and heavily influenced future work that incorporates QLCS-type storms (Gagne et al., 2009; Gallus et al., 2008). Years later, another influential paper by Smith et al. (2012) manually classified storms into 13 types, though they did not subdivide QLCS into multiple classes based on stratiform rain position. Their work focused on storms that produced significant weather for the entire US from 2003 to 2011 based on radar-derived attributes like length, shape, organization, and velocity. They noted that storms did not always fit neatly into a particular category.

Using the (simplified) manually classified dataset created by Smith et al. (2012), Jergensen et al. (2019) classed significant weather producing storms into only three categories: supercell, QLCS, and disorganized. They tested logistic regression, random and gradient-boosted forests, and support vector machines–all supervised machine learning methods. They considered hundreds of predictors based on radar attributes, environmental soundings, and shape, of which they found shape to be the most important predictor. In contrast, based on the QLCS subcategories suggested by Parker and Johnson (2000), Cui et al. (2021) aimed to understand

linear mesoscale convective systems in the US by categorizing MCSs into one nonlinear class and 5 linear classes. They manually developed a training dataset for their supervised machine learning approach to storm classification based on shape, movement, and radar properties, with a particular focus on how those different storms pose flooding risks. Both studies used similar attributes to classify their storms into completely different categories with high accuracy.

So, according to the literature, what is the correct storm taxonomy for training supervised classification models? In short, there does not appear to be one agreed upon taxonomy, though similar attributes and names are shared among the literature. From the 18 articles reviewed, the top ten most used attributes (and the respective number of articles) include: reflectivity (9), shear (7), eccentricity (7), aspect ratio (6), hail (5), length (5), CAPE (5), vertically integrated liquid (4), orientation (3), and height (3). Some storm types often mentioned in the literature include supercell, disorganized, non-linear, discrete, linear and/or QLCS with subtypes. Often, supercells are aptly named for their strength and vorticity; QLCSs for their longer duration and 1:3 aspect ratio; and linear, non-linear, cluster, and discrete storms for the shape and organization of their cells. Some less common storm types, and some of which are characteristically similar to the aforementioned types but under less common names, include pulse, isolated, general, non-supercell, marginal, hybrid, broken line, high-precipitation or low-precipitation, and severe or non-severe. How these storms are typed varies among papers.

1.3.2 Unsupervised

Unsupervised classification, also known as clustering or class discovery, is when data points are grouped together based on their shared attributes. According to Nagy (1968), clustering is used to better understand the structure of a dataset, and is dependent on the data's dimensionality, sample size, and the presence of "mavericks," or noise. There are many different clustering algorithms, but they can be broadly categorized into hierarchical (top-down or ground-up) and partitioning. Clustering is useful because there is no need for a training dataset, thus expert opinion on labeling is not required. This methodology can reveal previously unknown categories and show how the data and attributes are grouped, dispersed, and related. However,

when a good validation dataset is available, clustering results are often not as accurate as supervised approaches. Some known challenges to clustering include when groups are connected by noise, the groups are non-spherical, the sample size is small, or the cluster populations are different (Nagy, 1968).

Clustering methods are often applied to storms when teasing out areas of convection versus non-convection. Pacey et al. (2021) used the K-Means clustering method to differentiate high-shear low-CAPE and low-shear low-CAPE environments, but they followed up with a manual classification method of the storms. Conversely, Gagne et al. (2009) used K-Means clustering to segment reflectivity data into clusters and followed with a supervised decision tree method to label the clustered reflectivity data. Kolodziej Hobson et al. (2012) improved upon the Baldwin et al. (2005) watershed threshold technique by preceding the watershed method with K-Means clustering to identify levels of reflectivity. They argue that watershed segmentation alone (though better than basic thresholding) could leave out areas of weaker or smaller convection. All of the aforementioned works used unsupervised methods as a precursor to supervised storm type classification, but not as a standalone method. Only Baldwin et al. (2005) used a clustering algorithm to define classes in combination with labeled data. Similar to the work here, they used intensity and shape attributes to build a rule-based classification system. However, rule-based classification does not scale well to outside data like machine learning classification does.

Unsupervised learning, in general, can be used to discover patterns in massive amounts of data (Celebri & Aydin, 2016). In the Northern Hemisphere, convective storms can occur any time of year and at any time of day, which can result in very large, complex datasets. Supervised classification, though more accurate, requires training datasets that can take extensive time to develop. Most notably, nearly every supervised classification scheme is based on the professional meteorological opinion of a person or group of people with varying backgrounds, perspectives, and research goals. Thus, I think it is important to expand unsupervised clustering methods beyond separating convective versus non-convective components of storms, and instead use K-Means unsupervised learning to explore what the most natural, data-focused classification

schemas are for strong tornado-producing storms in the east-southeast. The research questions below are focused around this goal.

1.4 Research Questions

1. What are the radar-derived attributes and underlying data structure of EF2+ tornado-producing mesoscale storms in the east-southeastern US?
2. How are EF2+ tornado-producing mesoscale storms clustered using K-Means and Fuzzy C-Means unsupervised machine learning methods?
3. What is the spatial and temporal climatology of the storm types identified using clustering?

CHAPTER 2. DATA AND METHODS

2.1 Data

Tornado touchdown times and locations were derived from the SPC tornado report SVRGIS dataset, which includes tornado track beginning and end points, injuries, deaths, descriptions, EF-scale ratings, and more, from 1950 until a year or two before present (SPC, 2022). I selected data from 2000 until 2020 so that I might later climatologically compare the 2000s to the 2010s, and so that my dataset is large enough for reliable clustering. Tornado touchdowns and corresponding attributes were selected based on their spatial intersection with NEXRAD locations downloaded from the Department of Homeland Security's coordinates file (Homeland Infrastructure Foundation, 2016). I further narrowed down the dataset by removing tornado touchdowns within 1 hour of each other to reduce the impact of tornado outbreaks from one storm system. For this project, I want to emphasize strong tornado-producing storms outside of the Great Plains and deemphasize tornadoes produced from hurricanes—which was mostly achieved by excluding radar sites in the Plains and the majority of Florida, respectively (Figure 2.1).

Storms have been classified using a range of remote sensing sources, such as satellite-based infrared readings of cloud tops, satellite-derived lightning estimation, and doppler radar, as well as modeled information about the atmosphere. For this work, I use data acquired from the Next-Generation Weather Radar network (NEXRAD), which is a doppler radar array of 160 S-band sensors operated by the US National Weather Service, the US Federal Aviation Administration, and the US Air Force. There are two data suites available for use, Level II and Level III. I acquired NEXRAD Level II data stored remotely by Amazon Web Services (Unidata, n.d.). In one volume scan, or several 360-degree sweeps at different altitudes, there are three base products (moments) including reflectivity, mean radial velocity, and spectrum width. Mid-2009, NEXRAD was updated for dual-polarization, which added three more moments: differential reflectivity, differential phase, and correlation coefficient. The Level III suite contains over 75 data products, but only a limited amount of data is stored remotely–data from the previous couple years. The rest of archived Level III data is coarser resolution and must be requested and

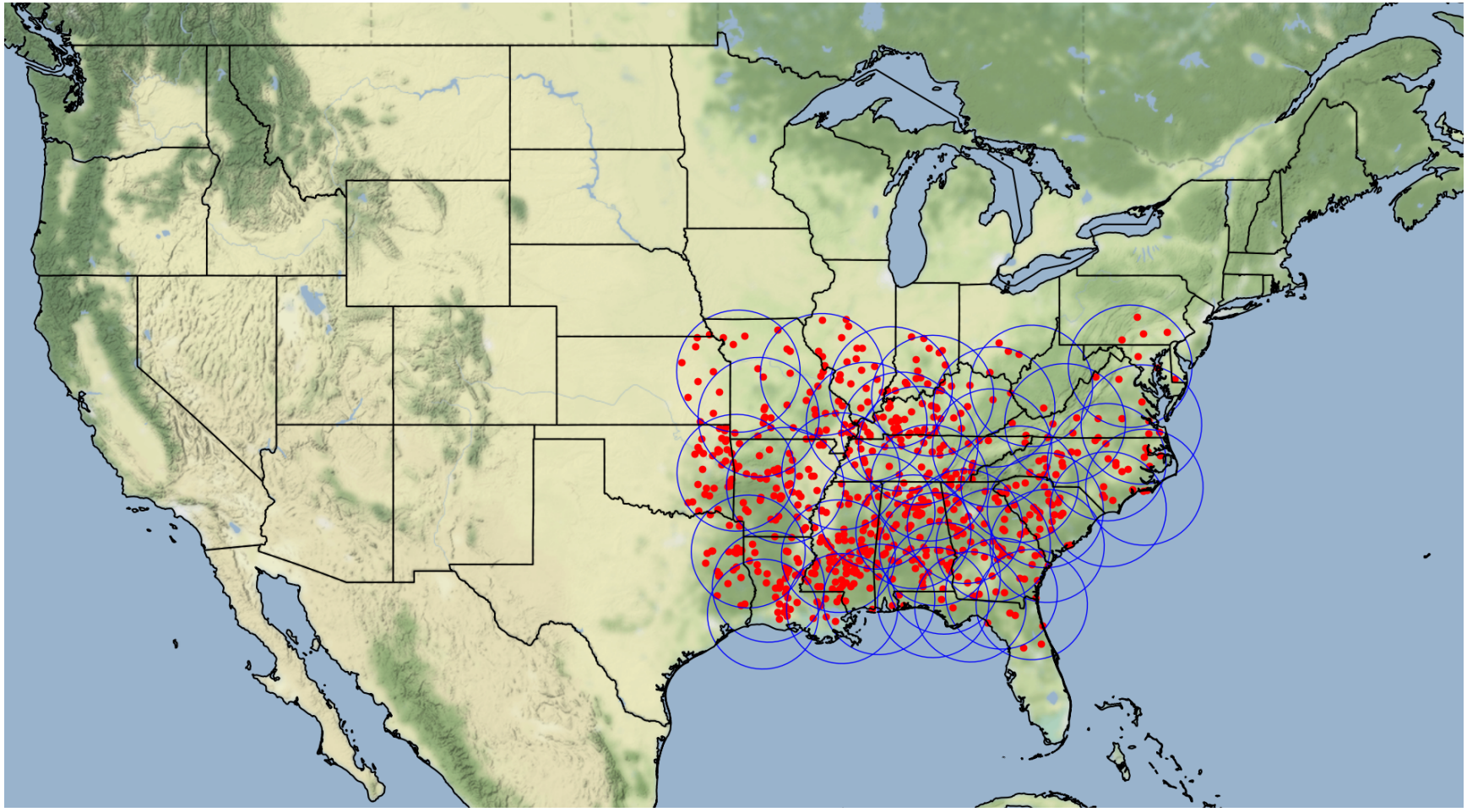


Figure 2.1 Study area, radar locations, and tornado touchdowns

downloaded in bulk, whereas Level II data is finer resolution and can be accessed on the cloud, filtered, cleaned, and then downloaded.

Reflectivity data is arguably the most important information used for classifying storms and has been most extensively used in the literature, including 16 of 18 reviewed storm classification articles (Table 1.1). The most important moment, reflectivity, reveals the size of the storm, the shape, the strength, and other precipitation-related information. Certain visual cues, like hook and bow echoes, can indicate the presence of a tornado or the strength of the storm, respectively. Another useful moment, radial velocity, can be fed into the NEXRAD tornado detection algorithm, which identifies tornadic vortex signatures and mesocyclone presence, all of which are available only in the Level III suite. Spectrum width can be used in tandem to identify turbulence via the distribution of velocity within a radar pixel, though it is most often used to find the leading edge of convection. There are several other ways to use information from these moments, but the aforementioned applications are useful for identifying potentially tornadic storms.

There are several known issues with doppler radar data. Besides radar malfunctions, some common problems with NEXRAD data include atmospheric attenuation leading to the underestimation of storm intensity, beam spreading that lowers the resolution significantly with distance from radar, range folding that produces a "haze" on the radar feed, anomalous propagation from ground clutter and ducting, as well as a "cone of silence" where the radar cannot look directly overhead (Duda, 2009). While some of these issues are resolved before the Level II suite is produced, they can remain as artifacts and require interpolation or removal, and in any case should be understood as a basic limitation of the data source.

Initially, I acquired 1,614 volume scans for 824 tornado touchdowns from 2000 to 2020 in the eastern and southeastern US for 40 radar sites. Summary statistics of reflectivity, velocity, and spectrum width were used to identify normal versus outlier data found within volume scans. I also considered the number of rays (radar pulses), the number of sweeps (360-degree rotations at various elevation angles), elevation angles (vertical tilt of the radar), and sweep azimuths (orientation of the radar relative to North) to understand the normal behavior of NEXRAD radars

touchdowns within the context of the 1,614 volume scans (Table 2.1). The number of rays, sweeps, and elevation angles are a product of the chosen volume coverage pattern (VCP) selected by a meteorologist at the time of a precipitation event. There are 7 different precipitation mode VCPs and 2 clear air mode VCPs (National Weather Service, 2019).

I would expect the radars to adhere to the VCPs, so any distinct outliers were identified and removed from the radar summary statistics table. For example, though there is no official elevation angle range, generally radars should be within 0.5 and 19.5 degrees (NEXRAD Technical Information, 2022). However, some coastal radars may tilt below 0.5 degrees. So, I did not set a minimum elevation angle threshold of 0 and instead relied on outlier detection using the absolute value of Z-scores beyond 3, which only removed three extreme outliers. Twenty-seven radar volume scans with all zero rows were also removed from the summary statistics table, which left 1,584 volume scans and 811 storms. Lastly, some of the oldest volume scan files were missing essential VCP information and were removed, leaving enough volume scans for 708 storms. Thus, the final dataset contains gridded radar mosaics of 708 EF2+ tornado-producing storms within 6 minutes of tornado touchdown from 2000 until 2020 across 40 radar sites in the east-southeastern US.

2.2 Methods

2.2.1 Pre-Processing

Polar gridded NEXRAD Level II data within range of each tornado touchdown were mosaicked together using functions within the Python library, Py-Art (Helmus & Collis 2016), converted to a single cartesian grid, and were exported locally as a NetCDF file. For each grid, I extracted relevant shape, reflectivity, velocity, and precipitation-related attributes. The attributes were created on a >20 dBZ grid to represent stratiform rain areas and a >40 dBZ grid to represent likely areas of convection. Shape attributes describe the organization and linearity of the storms. Reflectivity attributes describe the storm's intensity. To mimic the search for velocity couplets, which are possible indicators of rotation, I used a Sobel edge detection algorithm

Table 2.1 Radar summary statistics

	Min dBZ	Max dBZ	Mean dBZ	Min Velocity	Max Velocity	Mean Velocity	Min Sweep Elevation	Max Sweep Elevation	N Sweeps	N Gates	N Rays
Mean	-29	67	13	-33	33	0	0	18	17	1835	7897
Std	5	7	7	2	2	2	0	4	3	4	2162
Min	-32	46	-11	-48	21	-8	0	0	2	1832	696
25%	-32	62	8	-34	32	-1	0	19	16	1832	6120
50%	-31	66	13	-34	34	0	0	20	17	1832	8280
75%	-28	70	18	-32	34	1	0	20	18	1840	8297
Max	-2	94	31	-21	36	12	1	20	25	1840	14040

implemented in the Python package, Scikit-Image (van der Walt et al., 2014), to identify areas of steep velocity gradients in convective portions of the storm. High Sobel edge values will correspond to steep gradients and thus possible velocity couplets. Lastly, to identify the positionality of stratiform rain relative to convection, I implemented the mean stratiform distance d_s equation created by Gagne et al. (2009):

$$d_s = |m_1|(x_s - x_c) + (y_c - y_s)$$

where m_1 is the slope of the convective major axis, *s* values are centroid coordinates of the stratiform areas, and *c* values are centroid coordinates of the convective areas. Assuming storms most often propagate eastward, higher positive values indicate precipitation in front of the storm, near zero suggest in-line precipitation, and lower negative values indicate precipitation behind the storm.

For each storm, I created two attribute tables—one for stratiform storm areas and one for convective storm areas. All attributes were considered for convective storm areas, but only the area attributes were selected for stratiform rain regions, resulting in 29 total attributes to consider for clustering. In clustering algorithms, too many attributes can introduce redundancy and noise, also known as the curse of dimensionality. Too many dimensions can confound results, increase computing complexity, and later, make it difficult for human interpretation. Humans are visual pattern seekers, but in the words of Nagy (1968), "What a pity our power fails us in hyperspace!" Further, for high dimensional data, Jain et al. (2000) suggested it is good practice to have ten times as many data points per class as the number of attributes. So, for example, if I expected 4 classes, then I would have needed a sample size of 1,160 storms for each class if I used 29 attributes, or dimensions. With a total sample size of 708 storms, I should extract *at most* 3 to 5 attributes. It is also relevant in this work to narrow down which attributes are most useful for creating classes and which ones are not.

There are two main ways to reduce dimensionality. The first, feature extraction, is to select a subset of the original attributes. The second, feature reduction, is to select an algorithm that merges existing attributes into fewer ones. I chose principal component analysis (PCA)—a dimensionality reduction technique. PCA is an unsupervised algorithm that fits new dimensions

within existing ones that represent the maximum variance in all the original data. Then, users can cluster on the new, reduced number of principal components (PC) instead of the high dimensional original dataset. While PCA is often used as a dimensionality reduction technique, my dimensions did not seem appropriately high to use the PCA matrix for clustering, and many of my attributes were correlated. Instead, I used PC eigenvalues to narrow down attributes for extraction, and thus minimizing dimensionality by removing redundant, collinear, or noisy attributes.

For the first PCA I conducted, I generated 6 PCs which explained 80% of the variance in my data. For each PC, the original attributes were assigned a value of importance—an eigenvalue. Then, I extracted the highest eigenvalue attributes from each of the 6 PCs (Table 2.2). I found that the highest eigenvalues for the first PC were convective (c) shape attributes. For the second PC, convective reflectivity attributes were ranked highest. Some PCs were a conglomeration of many kinds of attributes, but the 4th PC highly ranked convective eccentricity attributes, the 5th PC highly ranked convective Sobel attributes, and the 7th PC highly ranked stratiform (s) area attributes. This emphasizes the importance of convective shape and reflectivity variables, and it affirms that the Sobel, eccentricity, and stratiform area attributes are potentially useful. I repeated the PCA on a subset of the original attributes, resulting in 4 PCs and 5 final attributes to use for clustering (Table 2.3).

2.2.2 Clustering

Once I selected my five key attributes, I used K-Means clustering to sort the storms. The K-Means method assumes that the variance of each attribute distribution is spherical and uniform across all attributes. Each cluster should have roughly the same number of observations. If the data are non-normal, it may still perform well but will fail if there is excess noise in the data. In short, K-Means is unfriendly to outliers, non-spherical distributions, and variations in the density of data spread. K-Means is a hard clustering algorithm that only allows a data point to belong to one cluster. The “softer” version of K-means is Fuzzy C-Means (FCM), which incorporates the likelihood that a point belongs to a particular cluster. Thus, a data point could belong to multiple

Table 2.2 Top-ranking attributes for each PC (first PCA run)

	1	2	3	4	5
PC 1	**Mean area (c)***	**Max major axes (c)***	Max area (c)	Mean major axes (c)	Std of area (c)
	0.28	0.28	0.27	0.27	0.26
PC 2	**Max of max dBZ (c)***	**Mean of max dBZ (c)***	Mean of mean dBZ (c)	Median of max dBZ (c)	Max of mean dBZ (c)
	0.36	0.34	0.33	0.31	0.29
PC 3	Median major axes (c)	Median area (c)	Count of objects (c)	Mean eccentricity (c)	Std of eccentricity (c)
	0.36	0.34	0.28	0.27	0.24
PC 4	**Median eccentricity (c)***	**Mean eccentricity (c)***	Max eccentricity (c)	Std of mean dBZ (c)	Max of mean dBZ (c)
	0.43	0.41	0.40	0.32	0.28
PC 5	**Mean of max sobel (c)***	**Std of max sobel (c)***	Max of max sobel (c)	Median of max sobel (c)	Mean area (s)
	0.45	0.40	0.39	0.35	0.28
PC 6	Median area (c)	Std of mean dBZ (c)	Median minor axes (c)	Mean area (s)	Std of area (s)
	0.31	0.29	0.28	0.27	0.25
PC 7	**Median area (s)***	**Mean area (s)***	Std of eccentricity (c)	Mean of max sobel (c)	Std of max sobel (c)
	0.57	0.41	0.25	0.23	0.22

*attributes used for second PCA run

Table 2.3 Top-ranking attributes for each PC (second PCA run)

	1	2	3	4	5
PC 1	**Max major axes (c)*** 0.51	Mean area (c) 0.48	Mean of max dBZ (c) 0.40	Mean area (s) 0.37	Mean eccentricity (c) 0.34
PC 2	**Max of max dBZ (c)*** 0.54	Mean of max dBZ (c) 0.46	Mean of max sobel (c) 0.43	Std of max sobel (c) 0.41	Mean area (c) 0.27
PC 3	**Mean of max sobel (c)*** 0.55	Std of max sobel (c) 0.51	Max of major axes (c) 0.37	Mean area (c) 0.35	Max of max dBZ (c) 0.34
PC 4	**Mean eccentricity (c)*** 0.71	**Mean area (s)*** 0.65	Mean of max dBZ (c) 0.16	Mean of max sobel (c) 0.13	Mean area (c) 0.10

*attributes used for clustering

clusters. This method is useful for data with considerable overlap, and I will later use probability-based FCM results to evaluate K-Means results. Both clustering methods begin with the calculation of a cluster center and iterate until convergence. According to scikit-learn documentation (Pedregosa et al., 2011): "The k-means algorithm divides a set of N samples X into K disjoint clusters C, each described by the mean μ_i of the samples in the cluster. The means are commonly called the cluster 'centroids.' Note that they are not, in general, points from X, although they live in the same space. The K-Means algorithm aims to choose centroids that minimize the inertia, or within-cluster sum-of-squares criterion."

$$\sum_{i=0}^{n} \min_{\mu_j \epsilon C} \left(\|x_i - \mu_j\|^2 \right)$$

The K-Means algorithm was performed four times to further reduce dimensionality and optimize clustering results. FCM was performed using the Fuzzy-C-Means package in Python (Dias, 2019). FCM also divides a set of N samples into C fuzzy clusters. Each sample w_{ij} tells the degree to which element x_i belongs to a cluster c_j. The algorithm aims to minimize the following function:

$$argmin\ C \sum_{i=1}^{n} \sum_{j=1}^{c} w_{ij}^{m} \|x_i - c_j\|^2$$

CHAPTER 3. RESULTS AND DISCUSSION

3.1 K-Means Clustering

The attributes used for clustering include the max of major axes (c), the max of max dBZ (c), the mean of max Sobel (c), the mean of eccentricity (c), and mean area (s). These represent length, strength, velocity, shape, and size, respectively. 158 storms with missing attribute data were dropped because K-Means does not have built-in functionality for handling them, and 37 non-normal storms with absolute z-scores greater than 3 were removed because K-Means does not behave well with outliers. Thus, the first round of K-Means clustering had n=515 storms to cluster from out of the original 708. K-Means clustering requires the number of clusters to be pre-selected, so I tested 2, 3, 4, 5, 6, 7, and 8 clusters *k*, which ran on average in 0.029 seconds, compared to the weeks or months long process of manually classifying storms. The highest silhouette score (0.23) suggested that 3 clusters would yield the best results, followed by 2 (0.21), and 4 (0.20).

Reflectivity and major axis attributes had the clearest *k* separation, while Sobel, eccentricity, and stratiform area had considerable overlap (Figure 3.1). Mean stratiform area separated the least for every *k* approach, so it was removed as an attribute and clustering was repeated. For the next round of clustering (n=520), without the mean stratiform area attribute, silhouette score confidence overall increased with the best being 3 clusters (0.25), and 2 and 4 tied for second best (0.20). Mean convective eccentricity did not separate into distinct groups (Figure 3.2), with most groupings centered around the mean. Thus, this attribute was removed, and a third round of clustering was performed on max of major axes (c), the max of max dBZ (c), and the mean of max Sobel (c) with n=525. Silhouette scores improved again, with the highest scores centered around 2 (0.29), 3 (0.30), and 4 (0.29) clusters. Out of three attributes, the mean of max Sobel (c) had the most cluster overlap, so the final, simplest clustering approach (Figure 3.3) consisted of reflectivity and length attributes with n=674. Three (0.41) and 6 (0.37) clusters were suggested from the silhouette scores. This final clustering approach also showed the least data loss, with only 34 outlier storms dropped. While some of these storms may have been real, they were not a large portion of the dataset and did not strongly affect the shape of the

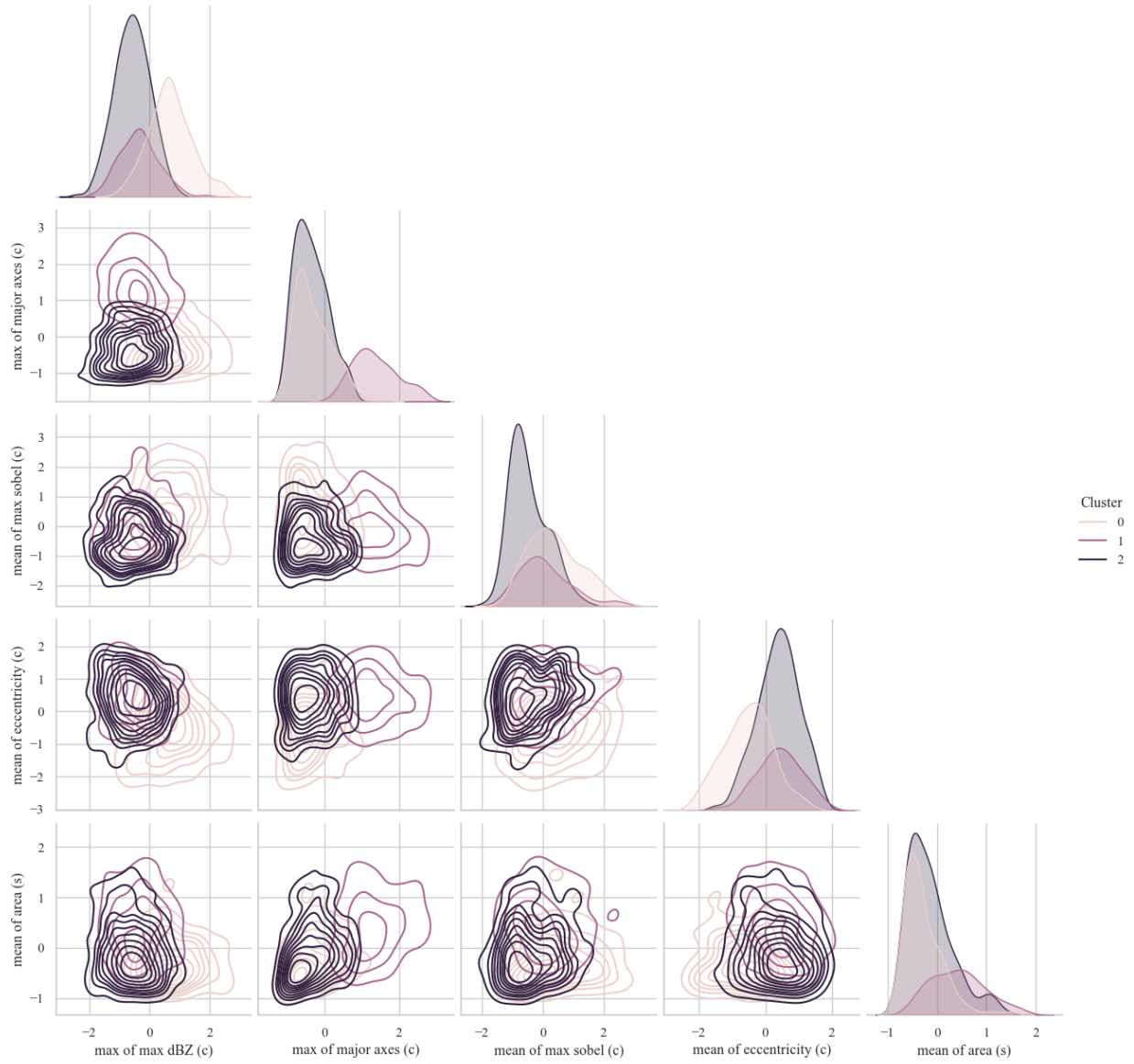


Figure 3.1 Three-cluster results using five attributes

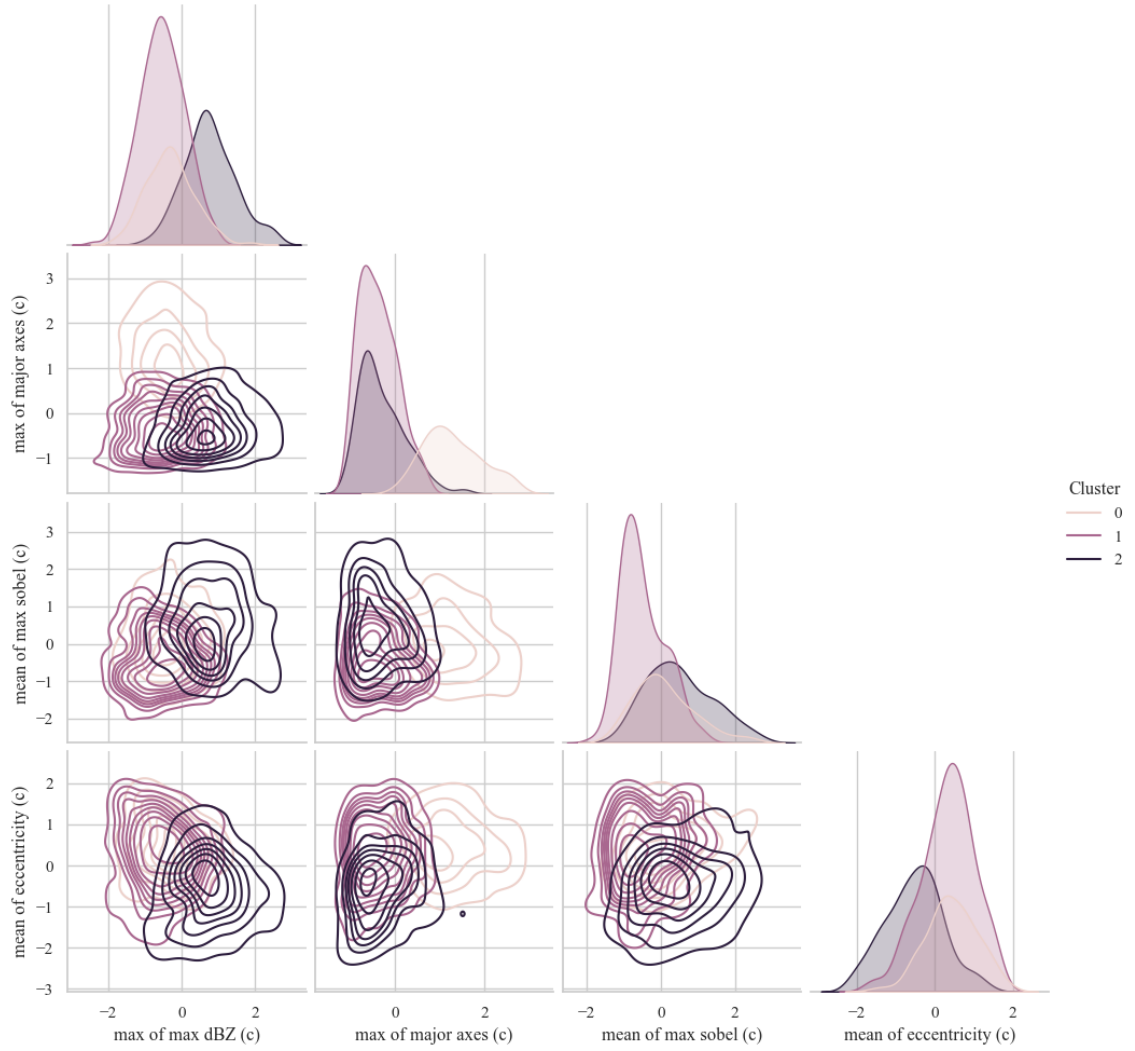


Figure 3.2 Three-cluster results using four attributes

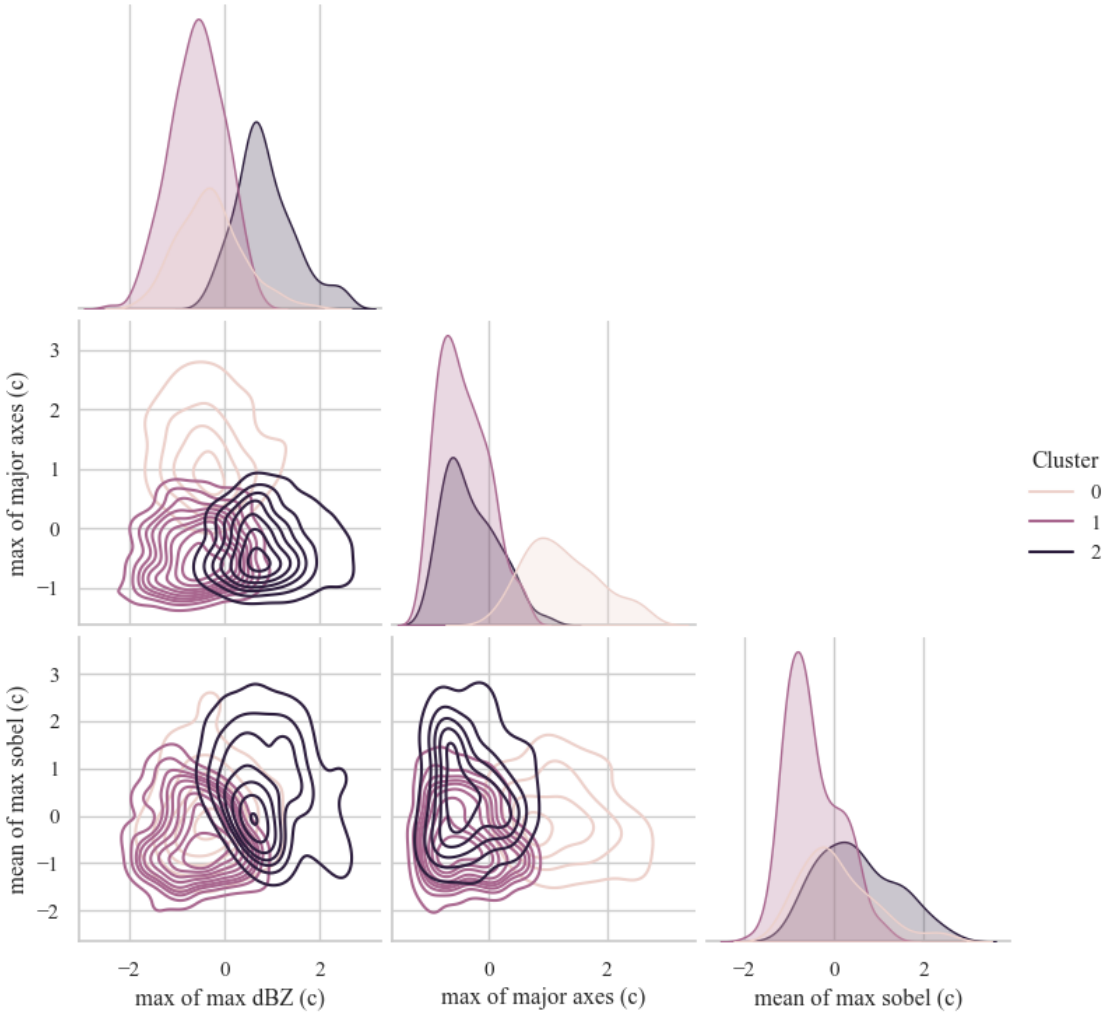


Figure 3.3 Three-cluster results using three attributes

data distribution. In contrast, earlier trials with more attributes may have been more affected by the removal of storms with missing data, which significantly reduced data breadth.

For simplicity, I selected 3 clusters: Cluster 0 (Weaker Linear) represented weaker, longer storms; Cluster 1 (Weaker Cellular) represented weaker, shorter storms; and Cluster 2 (Stronger Cellular) represented stronger, shorter storms (Figure 3.4, Figure 3.5). I generated descriptive statistics for each of the storm clusters (Table 3.1, Table 3.2). First, for Stronger Cellular storms, the mean strength was high around 0.9, while the mean length moderately low at -0.4. Similarly, the Weaker Cellular mean strength was appropriately low around -0.65 while length was also moderately low around -0.5. The Stronger Cellular and Weaker Cellular storms did well separating strength and moderately well separating length. However, I found that for Weaker Linear storms, the average strength was only mildly low around -0.3, but average length was high around one. This shows evidence that length was captured well, but strength was being skewed toward normal because of strong storms being mixed into the Weak Linear cluster.

I found that the K-Means clustering results most closely resembled the clustering schema designed by Baldwin et al. (2005). First, they manually classified storms into a linear-cellular-stratiform schema. Baldwin then used agglomerative clustering to define clusters (based on precipitation and shape attributes) that best resembled the manual schema. The final clusters were stratiform, two cellular groups, and two linear groups. Finally, they sorted storms outside the "training" dataset into these clusters using Euclidean distance. This schema was also developed using an unsupervised machine learning approach and emphasized shape and intensity attributes. Similarly, Agee and Jones (2009), Jergensen et al. (2019), and Karpinski (2020) used a linear-nonlinear-other classification schema that was shape-focused but did not incorporate intensity. Likewise, the main categories in Smith et al. (2012) comprised QLCS, supercell, and disorganized storms—categories that are, again, heavily dependent on shape and strength characteristics. However, like Gagne et al. (2009), Gallus et al. (2008), and Potvin et al. (2022), Smith et al. (2012) incorporated cellular organization as a defining attribute (e.g., clustered, linear, disorganized), which I did not include in my clustering analysis.

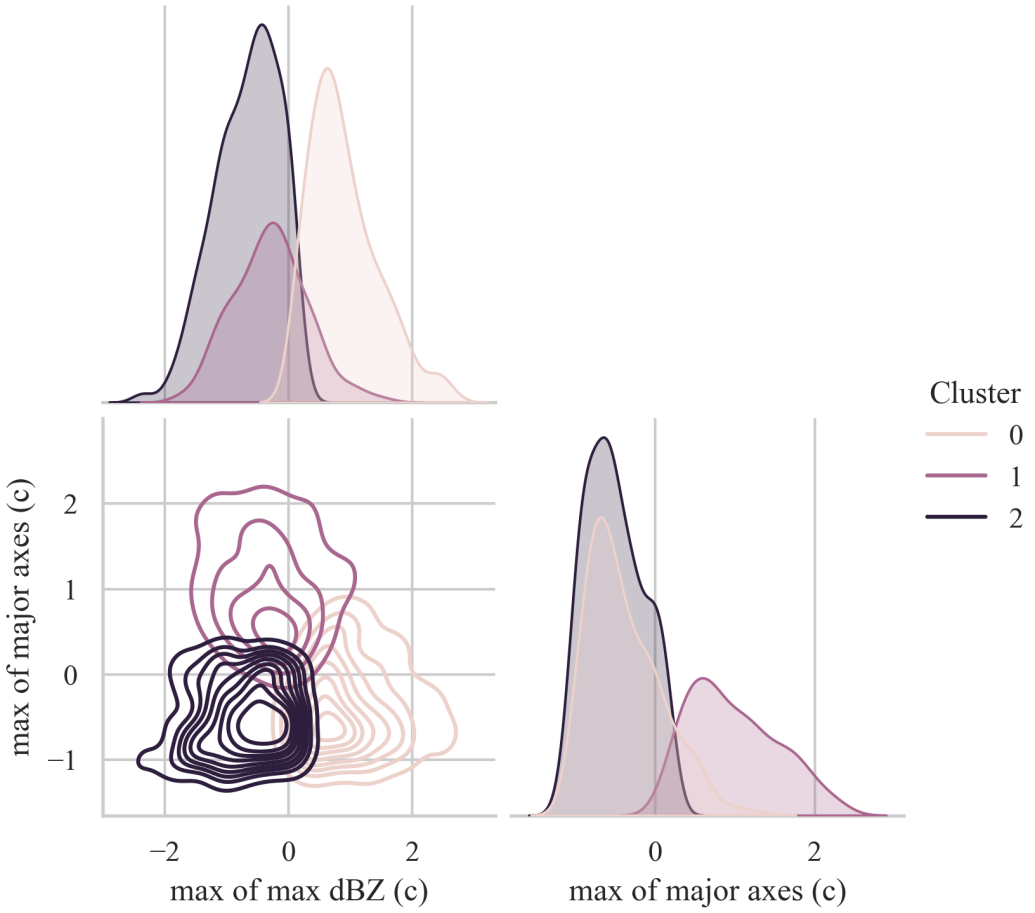


Figure 3.4 Three-cluster results using two attributes

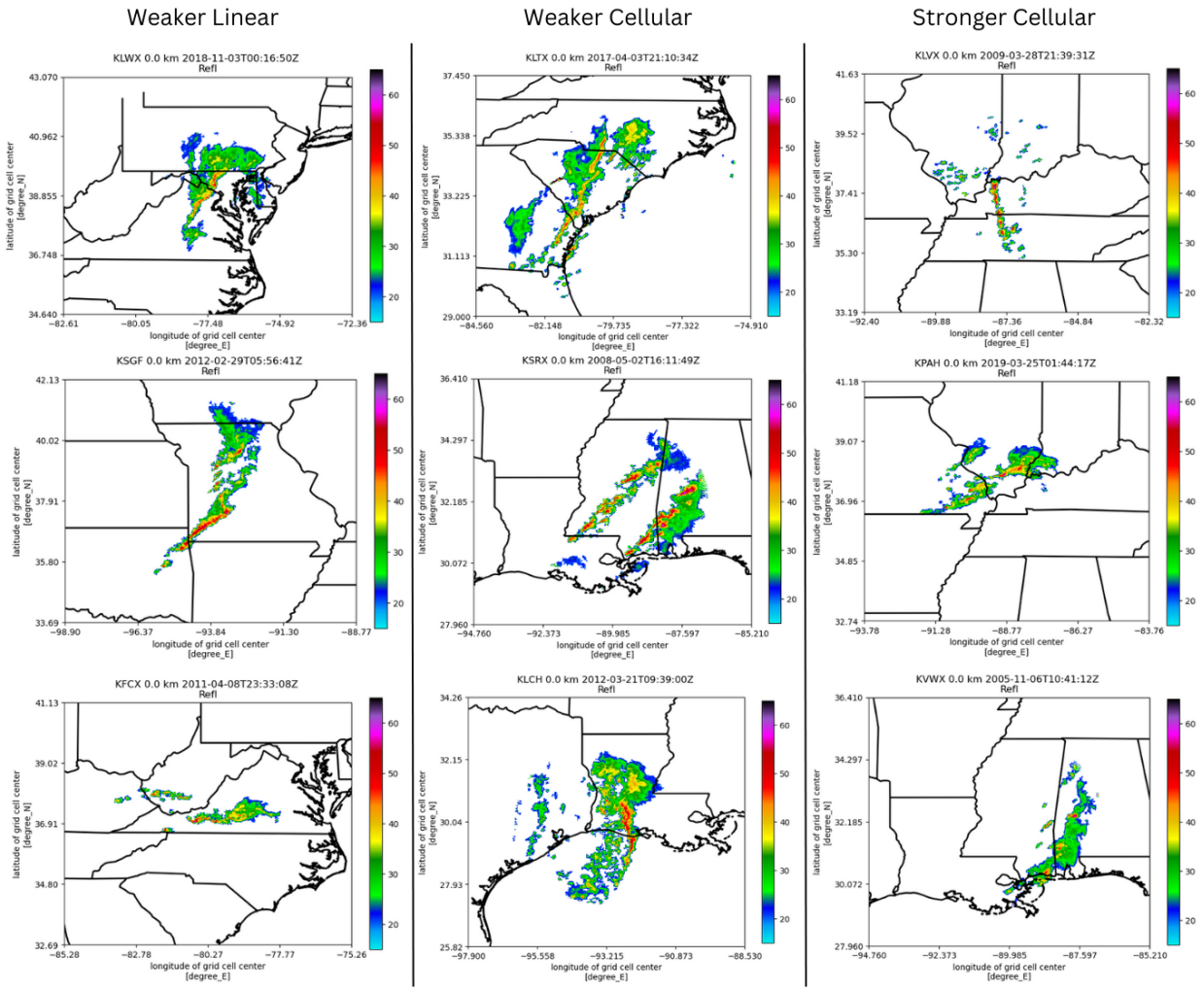


Figure 3.5 Random image samples of the three K-Means storm types

Table 3.1 Strength descriptive statistics for K-Means storm types (z-scores)

	Count	Mean	Std	Min	25%	50%	75%	Max
Weaker Linear	146	-0.3	0.6	-1.7	-0.7	-0.3	0.1	1.6
Weaker Cellular	292	-0.7	0.5	-2.4	-1	-0.6	-0.3	0.1
Stronger Cellular	222	0.9	0.6	0.1	0.5	0.8	1.2	2.6

Table 3.2 Length descriptive statistics for K-Means storm types (z-scores)

	Count	Mean	Std	Min	25%	50%	75%	Max
Weaker Linear	146	1.0	0.6	0.2	0.6	0.9	1.4	2.4
Weaker Cellular	292	-0.5	0.4	-1.2	-0.8	-0.5	-0.2	0.3
Stronger Cellular	222	-0.4	0.4	-1.1	-0.7	-0.5	-0.1	1.0

Several classification schemas also subdivided linear storms by the position of stratiform rain relative to the convective line. Though I incorporated an attribute designed by Gagne et al. (2009) to determine stratiform rain positionality, it was early on removed as a non-explanatory, low eigenvalue attribute during PCA. This, however, does not suggest that linear storms cannot be subdivided based on this attribute. In fact, Gallus et al. (2008), Gagne et al. (2009), Cui et al. (2021), and Parker and Johnson (2000) all subdivided long, persistent linear storms (QLCSs) into those with trailing, leading, or parallel stratiform rain. Lastly, future work could include common attributes like vertically integrated liquid and hail indices (NEXRAD Level III products) and environmental attributes like CAPE and shear, which may improve clustering results.

3.2 C-Means Analysis

After the final dataset was selected for climatological analysis, I performed FCM to identify storms that could belong to multiple clusters. I selected storms that were assigned a cluster with less than 80% certainty to evaluate how these uncertain storms were assigned to one of the 3 clusters. I found that 356 of 658 storms were assigned to a cluster with less than 80% certainty. Of those 356 storms, I divided the storms into four groups based on their relative position to normal reflectivity and normal length—below and above normal for both attributes (Table 3.3). This yielded four storm groups: relatively weak, short storms; relatively strong, short storms; relatively weak, long storms; and relatively strong, long storms. This is very similar to the K-Means clusters, but simplified thresholding has created a fourth group of strong, long storms, which was not identified by the K-Means clustering schema.

First, relatively weak, short storms represented 26% of all uncertain placements. These storms were all correctly placed in the Weaker Cellular cluster. The relatively strong, short storms represented 25% of all uncertain placements. These storms were placed in the Weaker Cellular or Stronger Cellular clusters. This suggests uncertainty separating weak from strong storms, and the potential for an “Average Cellular” category. Next, weaker, linear storms represented 26% of all uncertain placements, while stronger, linear storms represented 23% of all uncertain placements. Weaker, longer storms were placed in the Weaker Cellular cluster or were

Table 3.3 FCM results confusion matrix

	Weaker Cellular	Stronger Cellular	Weaker Linear	Total
weaker/shorter	91	0	0	91
weaker/longer	30	0	63	93
stronger/shorter	20	70	0	90
stronger/longer	2	39	41	82
Total	143	104	109	356

correctly placed in Weaker Linear clusters, suggesting that some uncertain placements were due to difficulties separating cellular from linear. This suggests the potential for a "Weak Average Length" storm cluster. Lastly, stronger, longer storms, which did not have their own cluster, were divided up amongst all three clusters. This may suggest the inclusion of a "Stronger Linear" cluster. K-Means likely did not produce a fourth, Stronger Linear cluster because K-Means divides the data into near-equal sized groups. Groups that are "rare" will not be extracted by K-Means.

3.3 Threshold Approach

After examining the results from K-Means and FCM, I noticed that the algorithms naturally divided the data along normalcy. As such, I wanted to compare creating storm types using thresholds around normalcy to storm types created by K-Means clustering. Similar to the creation of an FCM-based confusion matrix in Section 3.2, I created four categories by sorting storms along lines of normalcy. The four categories represented above/below normal strength and above/below normal length: Stronger Linear, Stronger Cellular, Weaker Linear, Weaker Cellular. In this case, there is no potential for overlap like there is in K-Means, and the "rarer" Stronger Linear storm type was created. There were n=140 Weaker Linear storms, n=88 Stronger Linear storms, n=230 Weaker Cellular storms, and n=202 Stronger Cellular storms. According to descriptive statistics (Table 3.4, Table 3.5), Weaker Linear storms separated out very well in low strength (-0.6) and high length (0.8), and Stronger Cellular storms were particularly strong (0.8). Weaker Cellular and Stronger Linear storms were not quite as distinct. Overall, strength appeared to divide slightly better than length, emphasizing the difficulties in separating short from long storms—i.e., there are more average-length storms than there are average-strength storms.

While investigating attribute distributions and inter-relationships using PCA, K-Means, FCM, and thresholding, I found that it is essential to understand that storm data does not separate into clear, distinct clusters. My analysis shows that storm types are not unequivocal—rather, storms belong on a continuum. For example, while there are linear storms and cellular storms,

Table 3.4 Strength descriptive statistics for threshold storm types (z-scores)

	Count	Mean	Std	Min	25%	50%	75%	Max
Weaker Linear	140	-0.6	0.4	-1.8	-1	-0.5	-0.3	0
Weaker Cellular	230	-0.7	0.5	-2.4	-1	-0.7	-0.4	0
Stronger Cellular	202	0.8	0.6	0	0.3	0.7	1.2	2.6
Stronger Linear	88	0.7	0.5	0	0.4	0.6	1	2.3

Table 3.5 Length descriptive statistics for threshold storm types (z-scores)

	Count	Mean	Std	Min	25%	50%	75%	Max
Weaker Linear	140	0.8	0.6	0	0.3	0.7	1.2	2.4
Weaker Cellular	230	-0.6	0.3	-1.2	-0.8	-0.6	-0.3	0
Stronger Cellular	202	-0.6	0.3	-1.1	-0.8	-0.6	-0.4	0
Stronger Linear	88	0.6	0.5	0	0.2	0.5	0.9	2.2

there are also those in between; while there are strong and weak, there are also average strength. Basic thresholding suggests that dividing Linear storms by their strength is useful, and it may also be of use to create average strength/length categories to generate greater separation between extremes. This separation may also be achieved by incorporating other attributes that were not surveyed in this work, such as orientation, environmental attributes, and more complex Level III-derived radar attributes.

3.4 Storm Type Climatology

3.4.1 Spatial climatology

The final datasets used for climatological analysis included 658 storms divided into 3 K-Means types: Weaker Linear, Weaker Cellular, and Stronger Cellular; as well as 658 storms divided into 4 threshold types: Weaker Linear, Stronger Linear, Weaker Cellular, and Stronger Cellular. Kernel Density Estimation (KDE) maps were created to understand the spatial climatology of the strong tornado-producing storm types. By removing storms within an hour of each other, the impact of tornado outbreaks was reduced but not removed. However, I found that an hour was sufficient time for the storm to appear different, and thus there are no near-exact replicas of individual storms within the dataset. By excluding a large portion of Florida and selecting only EF2+ tornadoes, tornado-producing hurricanes were limited but not removed. Additionally, potential edge effects may skew results in the KDE maps along my study area boundaries, especially along the Central Plains boundary where many strong tornadoes occur.

When comparing the 3 K-Means types to the 4 threshold types (Figure 3.6), the Weaker Linear, Stronger Cellular, and Weaker Cellular types were alike, with slight shifts in the center of each KDE. Stronger Linear storms from the threshold type displayed an intense signal between the Central Plains and Southeast, which is not shown by the K-Means types. In general, Weaker Cellular storms frequent the deep, coastal South, Appalachian Piedmont, and Tennessee Valley, while Stronger Cellular storm occur mostly in the Tennessee and Ohio Valleys. Weaker Linear storms are prevalent in the deep, coastal South and into the Tennessee and Ohio Valley regions.

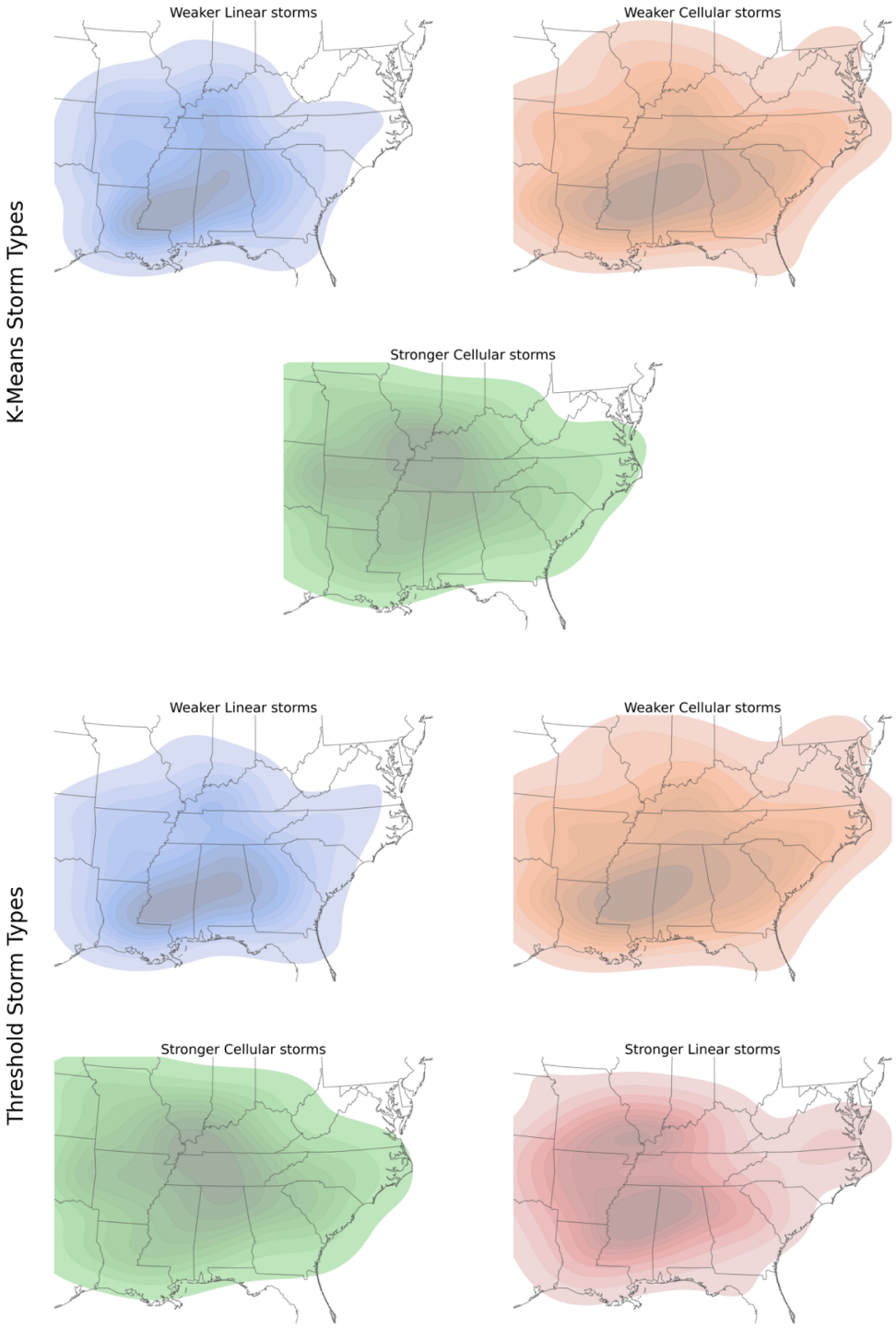


Figure 3.6 Storm type KDE maps of the East-Southeast US (2000–2020)

My KDE maps were visually akin to the KDE maps in Smith et al. (2012), where cellular storms prevailed in the deep, coastal South and central plains, while linear storms frequented the South and northward through the Tennessee and Ohio Valleys (Figure 3.7).

It is important to note that K-Means and thresholding produced similar results to each other with regards to their overlapping storm types: Stronger Cellular, Weaker Cellular, and Weaker Linear. In general, the centroid and orientation of the densest occurrences varied slightly between K-Means and thresholding types, and the Stronger Linear storm type revealed a new pattern not captured by the 3 K-Means clusters. Because K-Means sorts data into near-equal groups, this "rarer" group was not created. While there is nothing in the literature that suggests linear storms should be divided by their strength, this work presents evidence that this additional category could be considered. Depending on the needs of the researcher, both the K-Means storm types and threshold storm types spatially compare to similar storm types delineated in the literature.

3.4.2 Seasonal and diurnal climatology

Next, I assigned the season using equinox dates, and I assigned the time of day based on solar azimuth-derived sunrise/sunset times and location (Figure 3.8). Overall, strong, tornado-producing storms were most frequent during the day (74%) and least at night (26%). At night, most occurred during the winter (n=83), then fall (n=72), and then spring/summer (n=15). In contrast, during the day, most were in spring (n=284) and winter (n=120), with fewer in the fall (n=46) and summer (n=40). Of all Stronger Cellular storms, 91% of them occurred in spring (59%) and winter (32%), with a steep decline in fall (6%) and summer (3%). In contrast, only 33% of Weaker Cellular storms happened in spring, 29% in winter, 26% in fall, and 12% in summer. Thus, Weaker Cellular storms produced tornadoes essentially any time of year with a marginal peak in spring and a deeper minimum in summer. Finally, there were no instances of Weaker Linear storms in the summer. Weaker Linear storms also peaked in spring (47%), less so in winter (32%), and least in fall (20%) and summer (0%), similar to Weaker Cellular storms.

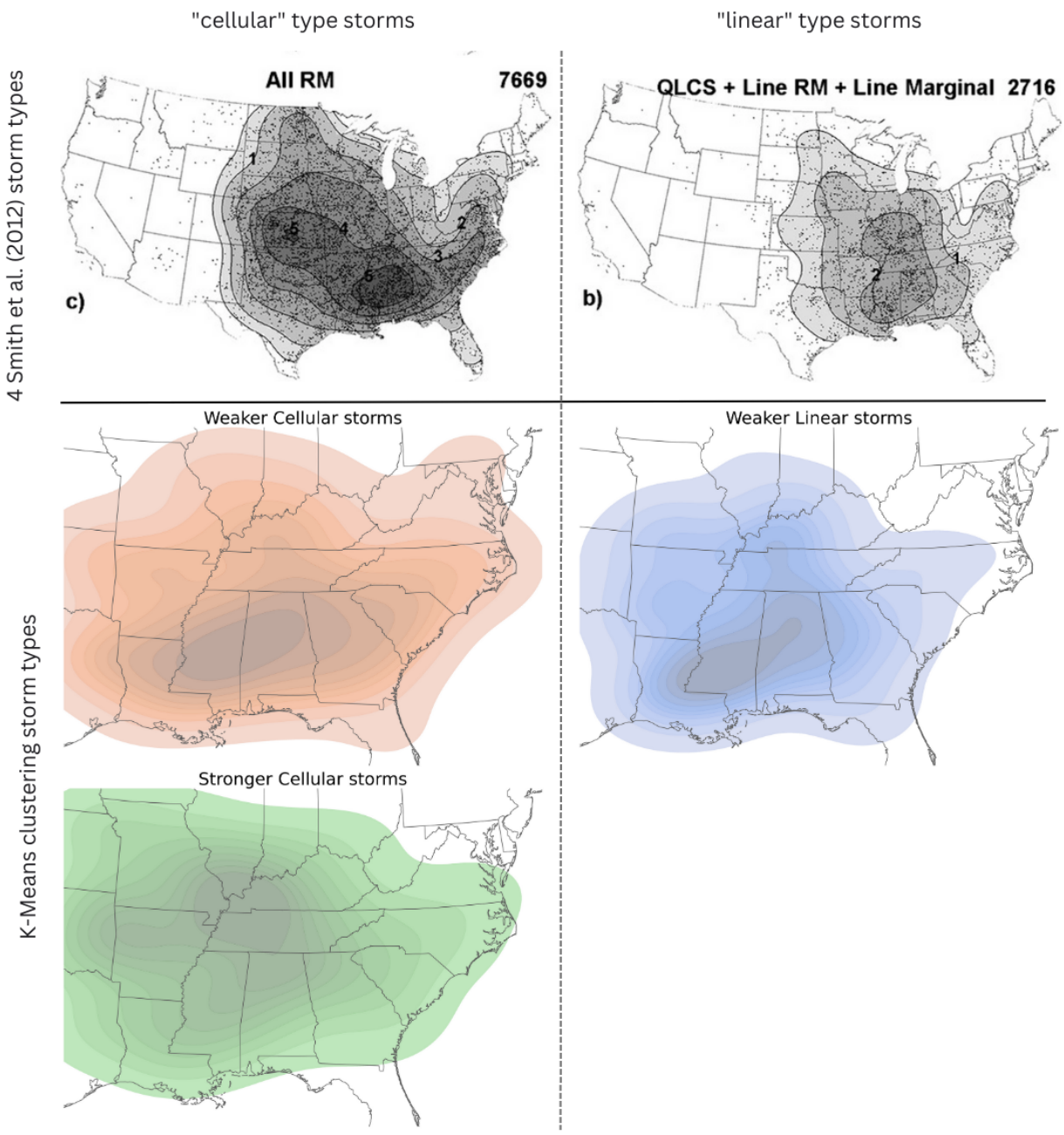


Figure 3.7 Smith et al. (2012) comparison of KDE maps

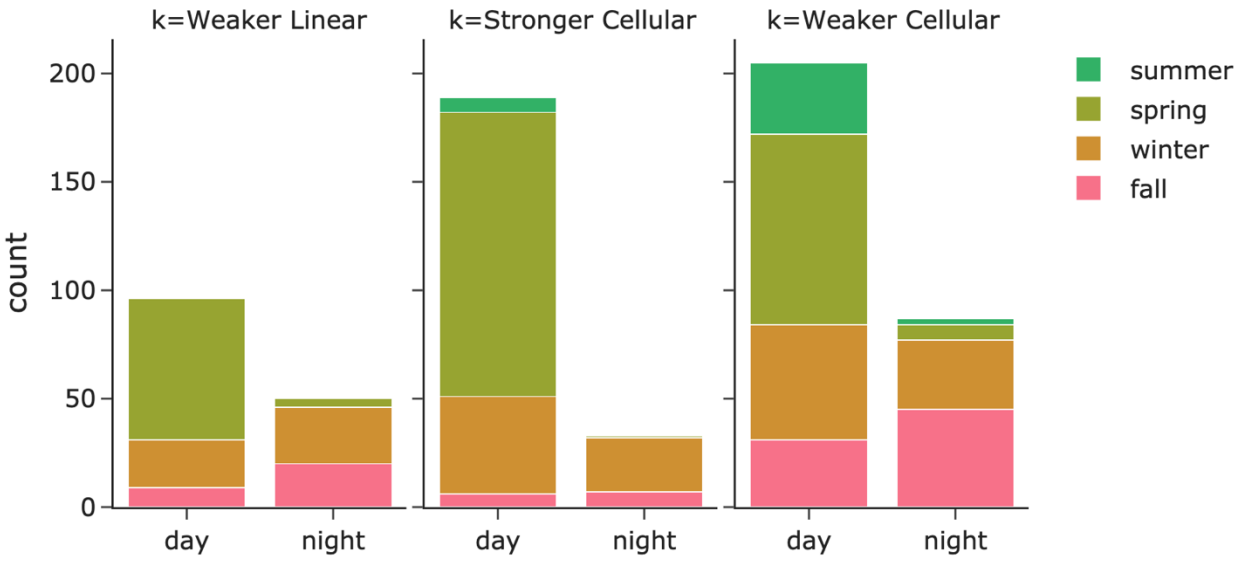


Figure 3.8 Seasonal and diurnal counts of the three K-Means storm types

However, the secondary peak for Weaker Linear storms was later in the year than the secondary peak for cellular storms (Figure 3.9).

Finally, I performed Pearson's Chi-Square tests to quantify the relationships between storm type, season, and time of day. Afterward, I calculated Cramer's phi to determine the strength of the relationships. I found that storm type, season, and time of day were significantly related (Table 3.6). According to Cramer's phi, season and storm type were weakly associated (0.24); time of day and storm type were also weakly associated (0.18); and season and time of day were strongly associated (0.51). When comparing the observed and expected values (Table 3.7), fewer Weaker Linear storms were observed than expected in the fall, winter, and night, while more were observed in the spring and day and a minimal difference in observed and expected values was observed in summer. The opposite occurred for Weaker Cellular storms, where instead, there were more observed than expected in the fall, winter, and night. Opposite to Weaker Linear storms, fewer Weaker Cellular storms were observed in the spring and day, and summer also showed minimal difference. In contrast to the weaker storm types, observed and expected Stronger Cellular values were similar for fall, spring, and summer. More winter and night storms were expected, while fewer day storms were expected.

Finally, I qualitatively compared the Smith et al. (2012) "Total number of tornado events by convective mode and month" figure to my own version with the K-Means storm types (Figure 3.10). The warm season peaks were offset by one month; mine all occurred in April, while the Smith et al. (2012) storms occurred in April for discrete and cluster right-moving (RM) storms, May for line RM and QLCSs, and June for disorganized storms. My figure also does not show as much of a difference in the number of Weaker/Stronger Cellular (discrete RM, cluster RM) and Weaker Linear storms (line RM, QLCS). However, my Weaker/Stronger Cellular storms peak for a second time in the cool season in September, similar to discrete RM and cluster RM storms. Additionally, both the line RM/QLCS and Weaker Linear storms peak later in the cool season than the discrete/cluster RM and Weaker/Stronger Cellular storms (though offset by one month).

In conclusion, the KDE maps and seasonal distribution chart of my three storm types compare qualitatively well to the ones in the seminal Smith et al. (2012) paper. In that work, they

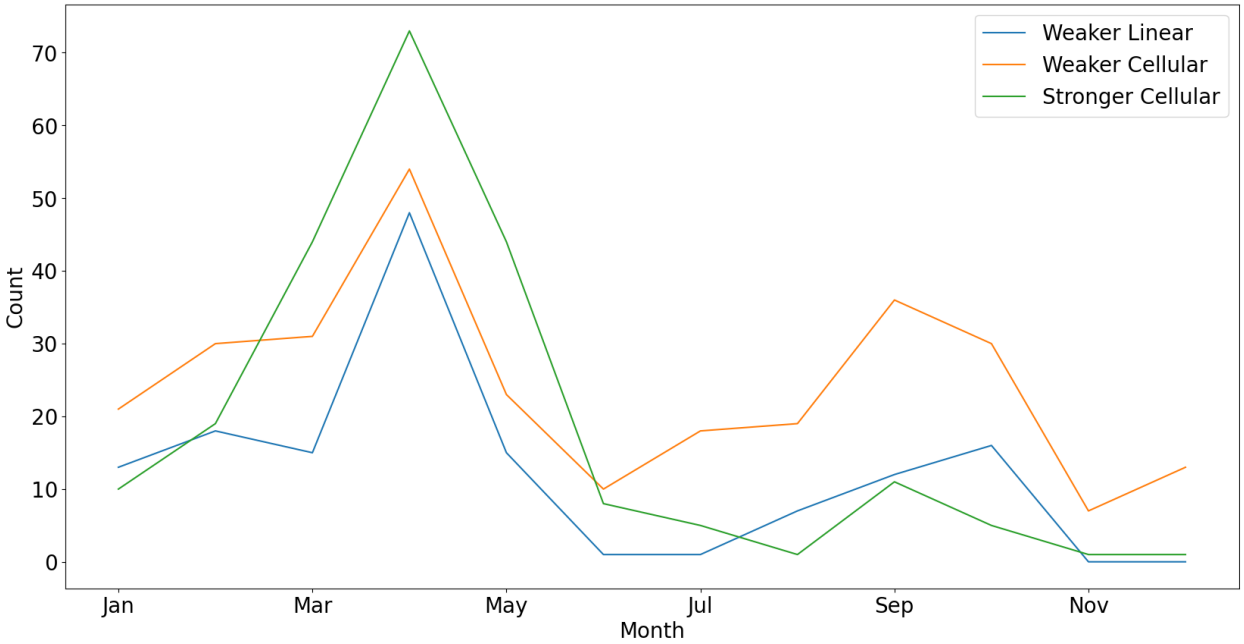


Figure 3.9 Monthly peaks of the three K-Means storm types

Table 3.6 Chi-square and Cramer's phi for storm type, season, and time of day

	Storm Type	Time of Day
Season	78.8**, 0.24	181.9**, 0.52
Time of Day	21.8**, 0.18	

Table 3.7 Chi-square test observed, expected, and standardized residual *R* values

	Stronger Cellular			Weaker Cellular			Weaker Linear		
	Obs	Exp	*R*	Obs	Exp	*R*	Obs	Exp	*R*
fall	13	40	-1.4	76	52	1.3	29	26	0.2
spring	132	100	1.8	95	131	-1.9	69	66	0.2
winter	7	15	-0.4	36	19	0.9	0	10	-0.5
summer	70	68	0.1	85	90	-0.3	48	45	0.2
day	189	165	1.4	205	217	-0.7	96	108	-0.7
night	33	57	-1.4	87	75	0.7	50	38	0.7

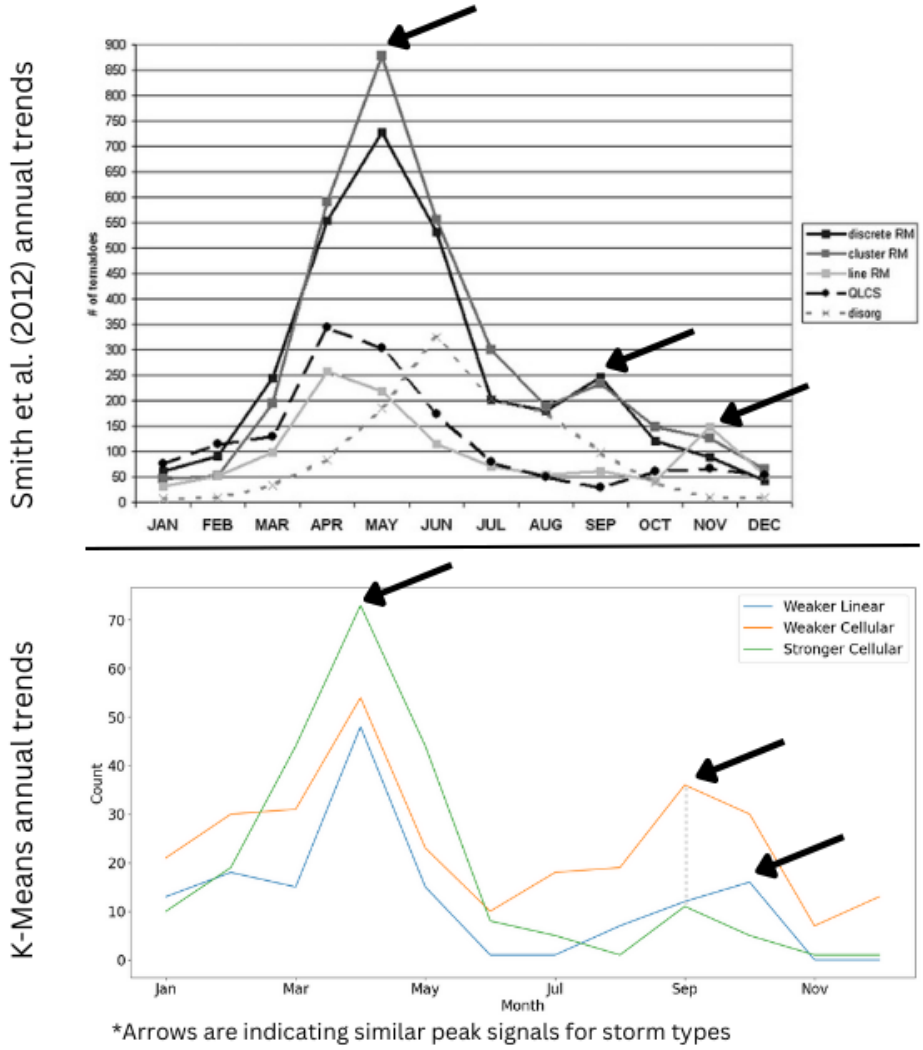


Figure 3.10 Smith et al. (2012) comparison of annual peaks

manually classified 10,724 EF0+ tornado-producing storms from 2003 to 2011. I did not include weaker tornadoes in my work, but the unsupervised method is scalable to any tornado-producing storms where radar data is available—though it may produce different key attributes and storm classes. In Smith et al. (2012), the storms were visually assessed over their lifetimes and were assigned one of 16 storm types, some of which represented storms changing from type to type over its duration. Their approach required expensive software for visualizing storms, while my approach was free and was performed mostly on the cloud. Though pre-processing took 3 to 6 hours, the K-Means results produced three storm types with similar climatological patterns within 0.029 seconds and did not require any manual classification. Manual classification, in contrast, can take weeks to months to complete. Also, the authors of Smith et al. (2012) noted "unambiguous discrimination between clusters and lines, as well as closely spaced cells versus clusters, was not always possible." This stressed the difficulty of classifying storms that did not fit neatly into their predefined categories. Approaching storm classification without such predefined constraints appears to be worthwhile and timesaving, and it is further imperative to understand that storm types belong on a continuum. Overall, this work revealed the potential for unsupervised learning to help narrow down important attributes, to sort both the extreme and "average" storms in a quick, objective way, and to create storm categories that are relevant and comparable to previous studies.

CHAPTER 4. CONCLUSION

One of the earliest references to "storm type" in the literature, by Blake (1933), classifies rainfall events in southern California by their paths, seasonality, origin, and precipitation patterns. Understanding storm types, what hazards they produce, and when, can improve our climatological understanding of them, their forecastability, and how they could be changing. Specifically, Smith et al. (2012) notes that storm type contributes to the potential and type of severe weather, and Guillot et al. (2008) goes further to say that a forecaster's ability to issue warnings is directly related to storm type. Interviews by Ellis et al. (2020), affirmed the complex interactions between time, location, storm type and forecaster behavior. For tornado-producing storms in particular, changes in our human-built environment and social vulnerability compound with changes in storm climatology to create damaging and deadly events (Ashley, 2017). Thus, understanding storm types, their characteristics, lifecycles, and behavior, can improve our understanding of past, present, and future hazard-producing events and their effects on people.

Much of the storm classification literature has focused on categorizing storms using manual methods or supervised machine learning methods. Many of these works fluctuate in the number and kind of storm classes, as well as in the attributes used to develop their classes. No work has used unsupervised machine learning to question, understand, improve, and propose storm taxonomies based on the explanatory and discriminatory ability of radar-derived attributes. Thus, I extracted such attributes from 708 EF2+ tornado-producing mesoscale storms from 2000–2020. Using these data, I sought to (1) understand the structure, explanatory value, and discriminatory abilities of NEXRAD Level II attributes for dividing storms into types; (2) explore potential classification schemas using K-Means and FCM unsupervised methods; (3) and apply my findings to a 20-year spatial and temporal climatology to assess the usability of these storm types in future work.

I found that K-Means could be used to quickly and objectively generate storm type datasets that qualitatively compare to those in the literature and produce familiar climatological trends. I found that the most important radar attributes for building these datasets include maximum convective length, maximum convective strength, mean convective velocity edges,

mean convective shape, and mean stratiform area. Further, the most discriminatory variables for producing distinct clusters with minimal overlap are the strength and length of convective storm areas. However, these clusters were generated mostly along lines of normalcy and were not fully separable, with over half of the storms sorted with less than 80% certainty. This indicated that storms belong on a continuum, and while storms situated on the extreme ends were distinctive, those along the center were more difficult to place. I found that separating strong from weak and short from long storms can pose challenges in classification, and I suggested a thresholding technique or the inclusion of "average" storm types to better separate extreme ends of the storm continuum.

Some important limitations to consider for this work include how data quality degrades rapidly with distance from the radar, so pixels along the farthest edge of the radars may not accurately capture the storm qualities. Also, while several seminal works (Duda & Gallus, 2010; Gallus et al., 2008; Smith et al., 2012) primarily define storms at the two-dimensional level, other works incorporate higher dimensions like vertical and temporal properties (Anagnostou, 2004; Cui et al., 2021; Li et al., 2000). Further, early radar data from the 2000s had VCP issues and were thrown out, which reduced the breadth of the dataset. In the future, works should include more storms and potentially ones that produced weaker tornadoes. They should also broaden the study area to reduce current edge effects and explore the highly tornadic Central Plains. They would also want to assess attributes from environmental soundings, models, and/or NEXRAD Level III moments, which were not included in this study. Lastly, future works may develop a method for calculating the organization of convective cells within a storm, which is a common attribute not expressed in machine learning literature.

Exploring potential tornado-producing storm types and their climatology is essential for understanding the impact of tornadoes on people, particularly those in the east-southeastern US that experience high damage costs and fatalities. With an improved understanding of storm type characteristics and divisibility, we can learn how they may be changing over time in conjunction with infrastructural, socio-political, and cultural knowledge that impacts human vulnerability to extreme weather. Thus, I hope this, and future work, will reduce the variability of mesoscale

storm typing amongst the literature, accelerate the classification process, and emphasize the importance and lack of attention on “average” tornado-producing storms that do not fit neatly within the extrema of the storm type continuum.

LIST OF REFERENCES

Agee, E., & Jones, E. (2009). Proposed conceptual taxonomy for proper identification and classification of tornado events. *Weather Forecast.*, *24*(2), 609–617. https://doi.org/10.1175/2008WAF2222163.1

Anagnostou, E. N. (2004). A convective/stratiform precipitation classification algorithm for volume scanning weather radar observations. *Meteorol. Appl.*, *11*(4), 291–300. https://doi.org/10.1017/S1350482704001409

Ashley, W. S. (2007). Spatial and temporal analysis of tornado fatalities in the United States: 1880–2005. *Weather Forecast.*, *22*(6), 1214–1228. https://doi.org/10.1175/2007WAF2007004.1

Baldwin, M. E., Kain, J. S., & Lakshmivarahan, S. (2005). Development of an automated classification procedure for rainfall systems. *Mon. Weather Rev.*, *133*(4), 844–862. https://doi.org/10.1175/MWR2892.1

Balling, R. C., & Cerveny, R. S. (2003). Compilation and discussion of trends in severe storms in the United States: Popular perception v. climate reality. *Nat. Hazards*, *29*(2), 103–112. https://doi.org/10.1023/A:1023674722282

Blake, D. (1933). Storm types and resultant precipitation in the San Diego area. *Mon. Weather Rev.*, *61*(8), 223–225. https://doi.org/10.1175/1520-0493(1933)61<223:STARPI>2.0.CO;2

Brandes, E. A. (1978). Mesocyclone evolution and tornadogenesis: Some observations. *Mon. Weather Rev.*, *106*(7), 995–1011. https://doi.org/10.1175/1520-0493(1978)106<0995:MEATSO>2.0.CO;2

Brooks, H. E., Carbin, G. W., & Marsh, P. T. (2014). Increased variability of tornado occurrence in the United States. *Science*, *346*(6207), 349–352. https://doi.org/10.1126/science.1257460

Brown, V. M., Ellis, K. N., & Bleakney, S. A. (2016). Tennessee tornado climate: A comparison of three cities. *Southeast. Geogr.*, *56*(1), 118–133. https://doi.org/10.1353/sgo.2016.0008

Brown, M. C., & Nowotarski, C. J. (2020). Southeastern U.S. tornado outbreak likelihood using daily climate indices. *J. Clim.*, *33*(8), 3229–3252. https://doi.org/10.1175/JCLI-D-19-0684.1

Cao, Z., Cai, H., & Zhang, G. J. (2021). Geographic shift and environment change of U.S. tornado activities in a warming climate. *Atmosphere*, *12*(5), 567. https://doi.org/10.3390/atmos12050567

Celebri, M. E., & Aydin, K. (Eds.). (2016). *Unsupervised Learning Algorithms*. Springer International Publishing. https://doi.org/10.1007/978-3-319-24211-8

Concannon, P. R., Brooks, H. E., & Doswell, C. A. (2000). Climatological risk of strong and violent tornadoes in the United States. *Second Symposium on Environmental Applications*, 212–219.

Craven, J.P., & Brooks, H.E. (2004). Baseline climatology of sounding derived parameters associated with deep moist convection. *Nat. Weather Dig.*, *28*, 13–24.

Cui, W., Dong, X., Xi, B., & Feng, Z. (2021). Climatology of linear mesoscale convective system morphology in the United States based on the random-forests method. *J. Clim.*, *34*(17), 7257–7276. https://doi.org/10.1175/JCLI-D-20-0862.1

Dias, M. L. D. (2019). *fuzzy-c-means: An implementation Fuzzy of C-means clustering algorithm*. Zenodo. https://doi.org/10.5281/zenodo.3066222

Diffenbaugh, N. S., Scherer, M., & Trapp, R. J. (2013). Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), 16361–16366.

Doswell, C. A., III, & Burgess, D. W. (1993). Tornadoes and tornadic storms: A review of conceptual models. In *The Tornado: Its Structure, Dynamics, Prediction, and Hazards* (pp. 161–172). American Geophysical Union (AGU). https://doi.org/10.1029/GM079p0161

Duda, J. (2009). *How to use and interpret Doppler weather radar*. Iowa State University. https://meteor.geol.iastate.edu/~jdduda/portfolio/How%20to%20read%20and%20interpret%20weather%20radar.pdf

Duda, J. D., & Gallus, W. A. (2010). Spring and summer midwestern severe weather reports in supercells compared to other morphologies. *Weather Forecast.*, *25*(1), 190–206. https://doi.org/10.1175/2009WAF2222338.1

Edwards, R., Dean, A. R., Thompson, R. L., & Smith, B. T. (2012). Convective modes for significant severe thunderstorms in the contiguous United States. Part III: Tropical cyclone tornadoes. *Weather Forecast.*, *27*(6), 1507–1519. https://doi.org/10.1175/WAF-D-11-00117.1

Ellis, K. N., Burow, D., Gassert, K. N., Mason, L. R., & Porter, M. S. (2020). Forecaster perceptions and climatological analysis of the influence of convective mode on tornado climatology and warning success. *Ann. Assoc. Am. Geogr.*, *110*(4), 1075–1094. https://doi.org/10.1080/24694452.2019.1670042

Elsner, J. B., Jagger, T. H., Widen, H. M., & Chavas, D.R. (2014). Daily tornado frequency distributions in the United States. *Environ. Res. Lett, 9*(2), 024018. https://doi.org/10.1088/1748-9326/9/2/024018

Gagan, J. P., Gerard, A., & Gordon, J. (2010). A historical and statistical comparison of "tornado alley" to "dixie alley*." Natl. Weather Dig.*, *34*(2), 145–155.

Gagne, D. J., McGovern, A., & Brotzge, J. (2009). Classification of convective areas using decision trees. *J. Atmos. Ocean. Technol.*, *26*(7), 1341–1353. https://doi.org/10.1175/2008JTECHA1205.1

Gallus, W. A., Snook, N. A., & Johnson, E.V. (2008). Spring and summer severe weather reports over the Midwest as a function of convective mode: A preliminary study. *Weather and For.*, *23*(1), 101–113. https://doi.org/10.1175/2007WAF2006120.1

Gensini, V. A., & Brooks, H. E. (2018). Spatial trends in United States tornado frequency. *NPJ Clim. Atmos. Sci.*, *1*, 38. https://doi.org/10.1038/s41612-018-0048-2

Grazulis, T. P., Schaefer, J. T., & Abbey, R. F., Jr. (1993). Advances in tornado climatology, hazards, and risk assessment since Tornado Symposium II. *Washington DC American Geophysical Union Geophysical Monograph Series*, *79*, 409–426. https://doi.org/10.1029/GM079p0409

Guillot, E. M., Smith, T. M., Lakshmanan, V., Elmore, K. L., & Stumpf, G. J. (2008). *Tornado and severe thunderstorm warning forecast skill and its relationship to storm type.* 24th Conference on IIPS, New Orleans, LA, US.

Guyer, J. L., Imy, D. A., Kis, A., & Venable, K. (2006). *Cool season significant (F2-F5) tornadoes in the gulf coast states*. 23rd Conf. on Severe Local Storms. St. Louis, MO, US.

Haberlie, A. M., & Ashley, W. S. (2018). A method for identifying midlatitude mesoscale convective systems in radar mosaics. Part I: Segmentation and classification. *J. Appl. Meteorol. Climatol.*, *57*(7), 1575–1598. https://doi.org/10.1175/JAMC-D-17-0293.1

Helmus, J.J. & Collis, S.M. (2016). The Python ARM Radar Toolkit (Py-ART), a library for working with weather radar data in the Python programming language. *J. Open Res. Softw*., *4*(1), e25. http://doi.org/10.5334/jors.119

Homeland Infrastructure Foundation. (2016). *Weather Radar Stations* [Data set]. data.world. https://data.world/dhs/weather-radar-stations

Hu, J., Rosenfeld, D., Zrnic, D., Williams, E., Zhang, P., Snyder, J. C., Ryzhkov, A., Hashimshoni, E., Zhang, R., & Weitz, R. (2019). Tracking and characterization of convective cells through their maturation into stratiform storm elements using polarimetric radar and lightning detection. *Atmos. Res.*, *226*, 192–207. https://doi.org/10.1016/j.atmosres.2019.04.015

Jackson, J. D., & Brown, M. E. (2009). Sounding-derived low-level thermodynamic characteristics associated with tornadic and non-tornadic supercell environments in the southeast United States. *Natl. Weather Dig.*, *33*(1), 12.

Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Trans. Pattern Anal. Mach. Intell.*, *22*(1), 4–37. https://doi.org/10.1109/34.824819

Jergensen, G. E., McGovern, A., Lagerquist, R., & Smith, T. (2019). Classifying convective storms using machine learning. *Weather Forecast.*, *35*(2), 537–559. https://doi.org/10.1175/WAF-D-19-0170.1

Johnson, J. T., MacKeen, P. L., Witt, A., Mitchell, E. D. W., Stumpf, G. J., Eilts, M. D., & Thomas, K. W. (1998). The storm cell identification and tracking algorithm: an enhanced WSR-88D algorithm. *Weather Forecast.*, *13*(2), 263–276. https://doi.org/10.1175/1520-0434(1998)013<0263:TSCIAT>2.0.CO;2

Karpinski, M. R. (2020). Typing tornadoes by storm mode in the southeast United States. *LSU Master's Thesis*, 5239.

Kolodziej Hobson, A. G., Lakshmanan, V., Smith, T. M., & Richman, M. (2012). An automated technique to categorize storm type from radar and near-storm environment data. *Atmos. Res.*, *111*, 104–113. https://doi.org/10.1016/j.atmosres.2012.03.004

Kotsiantis, S. B. (2007). Supervised machine learning: a review of classification techniques. *Informatica*, *31*(3), 249–268.

Lack, S. A., & Fox, N. I. (2012). Development of an automated approach for identifying convective storm type using reflectivity-derived and near-storm environment data. *Atmos. Res.*, *116*, 67–81. https://doi.org/10.1016/j.atmosres.2012.02.009

Lakshmanan, V., Rabin R., & DeBrunner V. (2003): Multiscale storm identification and forecast. *J. Atmos. Res.*, 67–68, 367–380. https://doi.org/10.1016/S0169-8095(03)00068-1

Lee, C. C. (2012). Utilizing synoptic climatological methods to assess the impacts of climate change on future tornado-favorable environments. *Nat. Hazards*, *62*(2), 325–343. https://doi.org/10.1007/s11069-011-9998-y

Lee, S., Robert Atlas, Enfield, D., Wang, C., & Liu, H. (2013). Is there an optimal ENSO pattern that enhances large-scale atmospheric processes conducive to tornado outbreaks in the United States? *J. Clim.*, *26*(5), 1626–1642. https://doi.org/10.1175/JCLI-D-12-00128.1

Lemon, L. R., & Doswell, C. A. (1979). Severe thunderstorm evolution and mesocyclone structure as related to tornadogenesis. *Mon. Weather Rev.*, *107*(9), 1184–1197. https://doi.org/10.1175/1520-0493(1979)107<1184:STEAMS>2.0.CO;2

Li, P. C., Pizzi, N., Pedrycz, W., Westmore, D., & Vivanco, R. (2000). Severe storm cell classification using derived products optimized by genetic algorithms. *2000 Canadian Conference on Electrical and Computer Engineering*, *1*, 445–448. https://doi.org/10.1109/CCECE.2001.933663

Long, J. A., Stoy, P. C., & Gerken, T. (2018). Tornado seasonality in the southeastern United States. *Weather Clim. Extrem.*, *20*, 81–91. https://doi.org/10.1016/j.wace.2018.03.002

Markowski, P., & Richardson, Y. (2013). How to make a tornado. *Weatherwise*, *66*(4), 12–19. https://doi.org/10.1080/00431672.2013.800413

Miller, P. W., & Mote, T. L. (2018). The algorithmic detection of pulse thunderstorms within a large, mostly non-severe sample. *Meteorol. Appl.*, *25*(4), 629–641.

Molina, M. J., & Allen, J. T. (2020). Regionally-stratified tornadoes: Moisture source physical reasoning and climate trends. *Weather Clim. Extrem.*, *28*, 100244. https://doi.org/10.1016/j.wace.2020.100244

Moore, T. W., & McGuire, M. P. (2019). Using the standard deviational ellipse to document changes to the spatial dispersion of seasonal tornado activity in the United States. *NPJ Clim. Atmos. Sci.*, *2*(1), 21. https://doi.org/10.1038/s41612-019-0078-4

MRCC. (n.d.). *MRCC Living With Weather - Tornadoes*. Retrieved January 25, 2023, from https://mrcc.purdue.edu/living_wx/tornadoes/index.html

Nagy, G. (1968). State of the art in pattern recognition. *Proc. IEEE. Inst. Electr. Electron. Eng.*, *56*(5), 836–863. https://doi.org/10.1109/PROC.1968.6414

National Weather Service. (2019). *NWS JetStream MAX - Doppler radar volume coverage patterns (VCPs)*. https://www.weather.gov/jetstream/vcp_max

NOAA. (2009a). *Glossary - NOAA's National Weather Service*. National Weather Service. https://w1.weather.gov/glossary/index.php?word=dBZ

NOAA. (2009b). *NOAA's National Weather Service - Glossary*. https://forecast.weather.gov/glossary.php?word=CONVECTION

NOAA. (2022). *2021 U.S. billion-dollar weather and climate disasters in historical context*. NOAA Climate.gov. https://www.climate.gov/news-features/blogs/beyond-data/2021-us-billion-dollar-weather-and-climate-disasters-historical

Nouri, N., Devineni, N., Were, V., & Khanbilvardi, R. (2021). Explaining the trends and variability in the United States tornado records using climate teleconnections and shifts in observational practices. *Sci. Rep.*, *11*(1), 1741. https://doi.org/10.1038/s41598-021-81143-5

Pacey, G. P., Schultz, D. M., & Garcia-Carreras, L. (2021). Severe convective windstorms in Europe: Climatology, preconvective environments, and convective mode. *Weather Forecast.*, *36*(1), 237–252. https://doi.org/10.1175/WAF-D-20-0075.1

Parker, M. D., & Johnson, R. H. (2000). Organizational modes of midlatitude mesoscale convective systems. *Mon. Weather Rev.*, *128*(10), 3413–3436. https://doi.org/10.1175/1520-0493(2001)129<3413:OMOMMC>2.0.CO;2

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., & Louppe, G. (2012). Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, *12*(85), 2825–2830.

Potvin, C. K., Gallo, B. T., Reinhart, A. E., Roberts, B., Skinner, P. S., Sobash, R. A., Wilson, K. A., Britt, K. C., Broyles, C., Flora, M. L., Miller, W. J. S., & Satrio, C. N. (2022). An iterative storm segmentation and classification algorithm for convection-allowing models and gridded radar analyses. *J. Atmos. Ocean. Technol.*, *39*(7), 999–1013.

Schiesser, H. H., Houze, R. A., & Huntrieser, H. (1995). The mesoscale structure of severe precipitation systems in Switzerland. *Mon. Weather Rev.*, *123*(7), 2070–2097. https://doi.org/10.1175/1520-0493(1995)123<2070:TMSOSP>2.0.CO;2

Schneider, R. S., Dean, A. R., Weiss, S. J., & Bothwell, P. D. (2006). *Analysis of estimated environments for 2004 and 2005 severe convective storm reports*. 23rd Conference on Severe Local Storms. St. Louis, MO, US.

Schumacher, R. S., & Rasmussen, K. L. (2020). The formation, character and changing nature of mesoscale convective systems. *Nat. Rev. Earth Environ.*, *1*(6), 300–314. https://doi.org/10.1038/s43017-020-0057-7

Sherburn, K. D., & Parker, M. D. (2014). Climatology and ingredients of significant severe convection in high-shear, low-CAPE environments. *Weather Forecast.*, *29*(4), 854–877. https://doi.org/10.1175/WAF-D-13-00041.1

Shield, S. A., & Houston, A. L. (2022). Diagnosing supercell environments: A machine learning approach. *Weather Forecast.*, *37*(5), 771–785. https://doi.org/10.1175/WAF-D-21-0098.1

Smith, B. T., Thompson, R. L., Grams, J. S., Broyles, C., & Brooks, H. E. (2012). Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. *Weather Forecast.*, *27*(5), 1114–1135. https://doi.org/10.1175/WAF-D-11-00115.1

Smith, R. K., & Leslie, L. M. (1978). Tornadogenesis. *Q. J. R. Meteorol. Soc.*, *104*(439), 189–198. https://doi.org/10.1002/qj.49710443914

SPC. (2022). *SVRGIS* [Dataset]. https://www.spc.noaa.gov/gis/svrgis/

Strader, S. M., Ashley, W. S., Pingel, T. J., & Krmenec, A. J. (2017). Projected 21st century changes in tornado exposure, risk, and disaster potential. *Clim. Change*, *141*(2), 301–313. https://doi.org/10.1007/s10584-017-1905-4

Thompson, R. L., Edwards, R., Hart, J. A., Elmore, K. L., & Markowski, P. (2003). Close proximity soundings within supercell environments obtained from the rapid update cycle. *Weather Forecast.*, *18*(6), 1243–1261. https://doi.org/10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2

Tippett, M. K., Lepore, C., & Cohen, J. E. (2016). More tornadoes in the most extreme U.S. tornado outbreaks. *Science*, *354*(6318), 1419–1423. https://doi.org/10.1126/science.aah7393

Unidata. (n.d.). *NEXRAD on AWS* [Data set]. Retrieved April 20, 2023, from https://registry.opendata.aws/noaa-nexrad/

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., Yu, T., & Contributors, T. S.-I. (2014). Scikit-image: Image processing in Python. *PeerJ*, *2*, e453. https://doi.org/10.7717/peerj.453

Whiton, R. C., Smith, P. L., Bigler, S. G., Wilk, K. E., & Harbuck, A. C. (1998). History of operational use of weather radar by U.S. Weather Services. Part I: The pre-NEXRAD era. *Weather Forecast.*, *13*(2), 219–243. https://doi.org/10.1175/1520-0434(1998)013<0219:HOOUOW>2.0.CO;2

Williams, J. K., & Abernethy, J. (2008). Using random forests and fuzzy logic for automated storm type identification [presentation]. In *Sixth Conference on Artificial Intelligence Applications to Environmental Science*. American Meteorological Society: New Orleans, LA, US.

VITA

Morgan Steckler grew up in Evansville, Indiana, and moved to Tennessee in 2017 to attend the University of Tennessee in Knoxville for her undergraduate B.A. in Geography. She explored research interests in paleoclimatology, natural hazards, computer science, and remote sensing. After graduating, she worked for several months as an intern at Oak Ridge National Laboratory (ORNL) using her skills in remote sensing and computer science. She returned to the University of Tennessee to acquire her M.S. in Geography, and she will also return to ORNL after graduation.