Spring 5-11-2024

# A Weather Radar Analysis of the Effects of the Albuquerque Urban Area on Thunderstorms from 2001 to 2020

Alexandra Jane Vivier
*University of New Mexico*

5-2024

# A Weather Radar Analysis of the Effects of the Albuquerque Urban Area on Thunderstorms from 2001 to 2020

Alexandra Jane Vivier

Alexandra J. Vivier

*Candidate*

Department of Geography and Environmental Studies

*Department*

This thesis is approved, and it is acceptable in quality and form for publication:

*Approved by the Thesis Committee:*

Christopher D. Lippitt, Chairperson

Xi Gong

David S. Gutzler

# A WEATHER RADAR ANALYSIS OF THE EFFECTS OF THE ALBUQUERQUE URBAN AREA ON THUNDERSTORMS FROM 2001 TO 2020

**BY**

**ALEXANDRA J. VIVIER**

**B.S., Aerospace Studies, Embry-Riddle Aeronautical University, 2003**
**A.A.S., Geographic Information Technology, Central New Mexico Community College, 2017**

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

**Master of Science**

**Geography**

The University of New Mexico
Albuquerque, New Mexico

**May 2024**

## ACKNOWLEDGMENTS

I wish to thank my advisor, Dr. Chris Lippitt, for his guidance and patience as I worked on my data and thesis. I also wish to thank my committee members, Dr. Xi Gong and Dr. David Gutzler, for their invaluable input on my work.

I am also grateful to Dr. Karl Benedict for our discussions about thunderstorms and urban climates, and his help with setting up RStudio scripts for the analysis.

I would also like to thank Dr. Steve Verzi for his help with configuring the Weather and Climate Toolkit scripts to be able to download radar data efficiently. I thank him and all of my Chinese Culture Center friends for their interest and encouragement as I worked on my thesis.

And finally, a special thank you to my Mom and Dad; without their support this effort would not have been possible.

# A WEATHER RADAR ANALYSIS OF THE EFFECTS OF THE ALBUQUERQUE URBAN AREA ON THUNDERSTORMS FROM 2001 TO 2020

**by**

**Alexandra J. Vivier**

**B.S., Aerospace Studies, Embry-Riddle Aeronautical University, 2003**

**A.A.S., Geographic Information Technology, Central New Mexico Community College, 2017**

**M.S., Geography, University of New Mexico, 2024**

## ABSTRACT

Urban areas have been found to affect the atmosphere locally through changes in emissions, thermal properties, the addition of moisture, and changes in airflow patterns. This study assessed the relationship between urban areas in central New Mexico and thunderstorm intensity and duration between 2001 and 2020. Weather radar data was used to derive thunderstorm events, as it is well-suited to fast-changing thunderstorm conditions. Regression analyses were used to determine the relationship between developed land cover and thunderstorm intensity and duration, while controlling for factors such as elevation, aspect, and near-surface temperatures. It was found that developed land cover was statistically significant and positively correlated with thunderstorm intensity and duration. Additionally, although developed land cover and the population of central New Mexico increased, it was found that maximum intensity, mean intensity, and duration did not increase between 2001 and 2020. Maximum intensity in the $95^{th}$ percentile did increase during that time.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

## 1.0 Introduction

Thunderstorms are a powerful combination of moisture, instability, and vertical motion in the atmosphere. Lightning and flash flooding caused by thunderstorms are the primary causes of weather-related fatalities during the summer in New Mexico (Sullivan et al., 2018), in addition to producing damaging winds and hail. Urbanization has the potential to change thunderstorms that occur in the vicinity (Bentley et al., 2010; Bentley et al., 2012; Shem & Shepherd, 2009; Shepherd, 2006). As city populations grow worldwide (Seto & Shepherd, 2009), increasing numbers of people are further exposed to thunderstorm hazards such as "large hail, high winds, tornadoes, and lightning" (Sullivan et al., 2018, p. 15). New Mexico's population has also seen growth in recent decades. Between 2001 and 2020, the population of the Albuquerque metropolitan area, New Mexico's largest urban center, grew by about 160,000 people, an increase of 30% (US Census). During the same time period, the city of Rio Rancho, within the northwestern extent of the Albuquerque metropolitan area, grew from a population of 54,893 to 104,257, an increase of 90% (US Census).

Urban environments can make a number of changes to the local atmosphere. Paved surfaces and buildings have distinct thermal properties, emissions from vehicles and industry add particulates to the air, irrigated lawns and sources of exhaust emit moisture into the atmosphere, and buildings change the pattern of airflow over the surface (Oke et al., 2017). Clouds form when water vapor condenses onto particles, called cloud condensation nuclei (CCN). Urban aerosols and air pollution can potentially provide a source of CCN (Oke et al., 2017, p. 270). In addition, higher temperatures and surface

roughness over cities can provide uplift to create clouds (Oke et al., 2017). Precise measurements of the effects of these factors can be difficult as the resulting changes in precipitation may not occur directly over the urban area but move beyond the city with the prevailing winds (Shepherd, 2006; Shem & Shepherd, 2009). Changes in cloud cover and precipitation indicate changes in atmospheric properties. Thunderstorm activity can also be presumed to reflect these changes.

Both thunderstorm activity and urban land cover relate to the temperature profile of the atmosphere, airflow across the surface, concentrations of air particulates, and atmospheric moisture. As such, it is noted that the output and characteristics of urbanization may correspond to the inputs for the initiation and life cycles of thunderstorms. As the population of central New Mexico has grown, these elements of urbanization have increased, while exposing more people to summer weather hazards.

This study assessed the relationship between the growth of the Albuquerque metropolitan area, including Rio Rancho and other neighboring communities, between 2001 and 2020, and the intensity and duration of thunderstorms in central New Mexico during that time. Regression analyses controlling for factors known to affect thunderstorms, such as the elevation of the terrain and near-surface temperatures, were used to assess the impact of urbanization. Twenty years of data from ground-based weather radar were used to represent the intensity and duration of thunderstorms. These data were compared to

surface land cover to determine if and how changes in the urban environment have affected thunderstorm activity in central New Mexico.

## 2.0 Background

A number of studies (Bentley et al., 2010; Bentley et al., 2012; Shem & Shepherd, 2009; Shepherd, 2006) have related changes in thunderstorms and precipitation to the urban heat island (UHI) effect and other results of urbanization, including increases in aerosols over and near cities. There are still debates about the extent to which these atmospheric changes may be directly linked to urban environments, or if other factors, including climate change, may be responsible. Much of the existing research (Bentley et al., 2010; Bentley et al., 2012; Shem & Shepherd, 2009; Shepherd, 2006) consists of case studies, as each city exists in a different topographic and climatic regime. Much of this work has focused on large cities such as Atlanta, Georgia (Shem & Shepherd, 2009; Haberlie et al., 2015); Indianapolis, Indiana (Niyogi et al., 2011); and Phoenix, Arizona (Shepherd, 2006). The current study presents analysis of a 20-year dataset of thunderstorm activity in central New Mexico, a location not previously studied.

### 2.1 Properties of Thunderstorms

In the late 1940s the Thunderstorm Project was conducted in Florida and Ohio to study "ordinary thunderstorms," and resulted in the creation of an idealized model of these storms and their life cycle. Ordinary thunderstorms, also known as airmass thunderstorms, or single-cell thunderstorms, were found to have a three-stage life cycle: the cumulus stage, the mature stage, and the dissipating stage. The cumulus stage is defined primarily by updrafts throughout the cloud. Downdrafts, created by drag from falling rain drops, signify the mature stage. The dissipating stage consists mainly of downdrafts and can be identified by its anvil shape as the top of the thunderstorm reaches

the upper limits of the troposphere and spreads out horizontally (Wallace & Hobbs, 2006, pp. 350 – 351). This is an idealized model; "single cell thunderstorms are rare; almost all thunderstorms are multicell. A multicell cluster thunderstorm … consists of a cluster of cells at various stages of their life cycle. With an organized multicell cluster, as the first cell matures, it is carried downwind, and a new cell forms upwind to take its place" (FAA, 2016, pp. 19-2 – 19-3).

Thunderstorms can form when the "ingredients" of conditional instability, atmospheric moisture, and a lifting mechanism are present. These atmospheric conditions, or their proxies, influenced the selection of regression variables used in this study.

### 2.1.1 Conditional Instability

There are a number of measures of change in air temperature with an increase in height above the ground. These include the *environmental lapse rate*, which is the actual change in air temperature with height, as measured by sensors carried aloft by weather balloons.

Atmospheric instability can be caused or increased by several mechanisms. Increased instability is based on the environmental lapse rate becoming steeper, "that is, as the air temperature drops rapidly with increasing height. This circumstance may be brought on by either air aloft becoming colder or the surface air becoming warmer" (Ahrens, 2007, p. 145). The air aloft may become colder due to wind bringing in colder air. Air near the surface may become warmer due to "daytime solar heating of the surface," "an influx of

warm air brought in the by the wind," or "air moving over a warm surface" (Ahrens, 2007, p. 146).

Air that is unsaturated, meaning that its relative humidity is less than 100 percent, cools with an increase in height at a rate of 10°C/km. When saturated air, which has a relative humidity of 100 percent, is lifted and cooled, it releases latent heat as the moisture in the air condenses. This causes the lapse rate of saturated air to vary, but it is approximated to 6°C/km. Unsaturated air, once it has been lifted sufficiently to cool to its dew point, will then become saturated and cool at the saturated lapse rate (Ahrens, 2007, p. 140). *Conditional instability*, one of the ingredients necessary for thunderstorm formation, "occurs whenever the environmental lapse rate is between the [saturated] adiabatic lapse rate and the dry adiabatic lapse rate" (Ahrens, 2007, p. 145). Further, conditional instability "depends on whether or not the rising air is saturated. When the rising parcel of air is unsaturated, the atmosphere is stable; when the parcel of air is saturated, the atmosphere is unstable. Conditional instability means that, if unsaturated air could be lifted to a level where is becomes saturated, instability would result" (Ahrens, 2007, p. 145). The depth of the conditionally unstable layer, above the LCL, determines the vertical extent of any cumulus clouds that have formed, and whether they are able to develop into thunderstorms (Ahrens, 2007, pp. 151 – 152).

Instability is described by several different metrics. One of the most common is *convective available potential energy* (CAPE), which describes the difference between

the measured air temperature and the temperature of a theoretical parcel of air rising and cooling at the saturated lapse rate (Wallace & Hobbs, 2006, p. 345). CAPE "is the maximum amount of energy available to an ascending air parcel for convection.… CAPE is directly related to the maximum potential vertical speed within an updraft" (FAA, 2016, p. 12-10).

### 2.1.2 Atmospheric Moisture

Sufficient atmospheric moisture must be present to form clouds. Moisture, in the form of water vapor, can enter the atmosphere from the surface as evaporation from areas of open water or transpiration from vegetation (FAA, 2016, p. 3-2).

Winds can bring atmospheric moisture into New Mexico as "low-level southeast flow ushers in moisture from the Gulf of Mexico" and "mid- and high-level moisture from the Pacific Ocean [is] most common with west to southwest flow aloft" (Sullivan et al., 2018, p. 14). However, "during the summer monsoon season, enhanced atmospheric moisture moves into the state from both the southwest and southeast" (Sullivan et al., 2018, p. 14) and "remnants of tropical storms, from both the Pacific Ocean and the Gulf of Mexico, can bring locally intense tropical rainfall. Such tropical events peak in September" (Sullivan et al., 2018, p. 15).

In multicell thunderstorms, surface outflow from one storm may enhance the intake of moisture into next storm cell (Ahrens, 2007, p. 371). Once the thunderstorm reaches the

mature stage, “precipitation descends through the cloud and drags adjacent air downward, creating a strong downdraft alongside the updraft” (FAA, 2016, p. 19-2). This precipitation is measured by weather radar (discussed in detail in Section 2.2) and thus, thunderstorms are not well detected by radar during their initial formation (Vasquez, 2015, p. 120; Fabry, 2015). Once the downdraft contacts the surface and spreads ahead of the storm as a gust front, “uplift … may trigger the formation of new [thunderstorm] cells, sometimes well ahead of the parent cell” (FAA, 2016, p. 19-2).

### 2.1.3 Lifting Mechanisms

Thunderstorms are a visible indicator of convective activity in the atmosphere, as “convection is the upward transfer of thermal energy by air motion” (Barry & Blanken, 2016, p. 109). Convection can be initiated by heating of the surface, creating regions of lower air density, as well as when air “flow over the surface is perturbed by inertial forces due to surface irregularities forcing it to rise” (Barry & Blanken, 2016, p. 109). More specifically, lifting mechanisms that begin to move air upward can include cold fronts or surface lows, but most commonly in New Mexico lifting is caused by uneven heating of the earth’s surface or by wind moving over rising terrain (FAA, 2016; Bowen, 1996). As previously mentioned, lifting may also be caused by outflow from nearby thunderstorms.

### 2.1.4 Spatial and Temporal Distribution of Thunderstorms

The spatial and temporal distribution of thunderstorms in New Mexico, Arizona, Texas, and central Mexico has been studied a number of times and discussed relative to topographic features in each location (Sullivan et al., 2018; Bowen, 1996; Fosdick & Watson, 1995; Novo & Raga, 2013; Wagner et al., 2006; Maddox et al., 2002; Watson et al., 1994; and López et al., 1997). A 10-year study of rain gauge and other meteorological data collected near Los Alamos, New Mexico, in the Jemez Mountains northwest of Santa Fe, showed that the number of thunderstorm days peaked in July and August and about 35% of the annual rainfall occurred during those months (Bowen, 1996). Consistently, precipitation occurred at high elevations in early afternoons, and at lower elevations later in the afternoon and evening, with lower elevations receiving more precipitation overall (Bowen, 1996).

Additionally, a study using cloud-to-ground lightning flash data archived by the Bureau of Land Management (BLM) was used to map the distribution of lightning flash densities and the number of thunderstorm days in New Mexico for June through September of 1985 to 1990. The study found that "high flash densities correlate closely with the elevated terrain found in the mountainous areas…. An excellent example of the strong correspondence between high flash densities and elevated terrain was the high lightning activity … directly over the Jemez Mountains, which are comprised of isolated peaks rising over 11,000 feet (3,353 m)" (Fosdick & Watson, 1995, p. 19). The authors also found that lightning activity generally started in the mountainous areas of New Mexico, particularly the Sangre de Cristo Mountains immediately east and northeast of Santa Fe,

at midday, and ended in the plains southeast of the mountains at night (Fosdick & Watson, 1995). This movement was surmised to be due to "mid-level northwest flow [that] may be steering the mountain-generated thunderstorms southeastward across the Plains later in the day," and that mid-level northwest flow corresponded with the maximum "frequency of thunderstorm genesis along the eastern slopes of the Sangre de Cristo Mountains," where "leeside convergence" occurred (Fosdick & Watson, 1995, p. 22). The authors concluded that "this suggests that terrain features, such as slope and aspect, play a more critical role in the development of deep convection than elevation alone" (Fosdick & Watson, 1995, p. 22).

Mexico, Arizona, New Mexico, and other western states experience the North American monsoon (NAM) season in the summer months (Adams & Comrie, 1997). In New Mexico, the NAM occurs mainly in July and August, and during that time, most of New Mexico receives about half of its yearly rainfall, much of it coming from "high-intensity, short-duration, thunderstorm events" (Sullivan et al., 2018, p. 14). Sullivan et al. (2018) add, "Although thunderstorms become more numerous statewide during the monsoon months of July and August, the number of severe thunderstorms decreases from spring to summer" (p. 16).

### 2.2 Weather Radar

The NEXRAD (Next Generation Weather Radar) WSR-88D (Weather Surveillance Radar – 1988, Doppler) is a ground-based weather radar system that was developed

through the 1970s and 1980s. WSR-88D was first installed in Oklahoma City as a prototype in 1990, then deployed at 154 locations across the United States between 1993 and 1997. WSR-88D remains the current system used in the United States (Vasquez, 2015, p. 7) and there are 160 sites in operation as of 2022 (NOAA, 2022b). The data used in this study were generated by the Albuquerque, New Mexico, NEXRAD WSR-88D weather radar; its identifier is KABX.

### 2.2.1 WSR-88D Fundamentals

Each WSR-88D unit consists of a transmitter, antenna and pedestal, receiver, and signal processor. These four components comprise the radar data acquisition (RDA) unit. The RDA unit sends data to the radar product generator (RPG), which does additional data processing to create Level III products (not used in this study) (Vasquez, pp. 39 – 40, 2015).

The WSR-88D transmitter sends out very short pulses of electromagnetic energy in the microwave portion of the electromagnetic spectrum. The energy transmitted and received by a WSR-88D unit has a peak wavelength (λ) of 10cm and is classified as S-band. When the pulses of energy strike particles in the atmosphere, they are generally scattered in all directions. The energy that returns to the antenna, called backscatter, is interpreted and presented as analog data (Level I), by the RDA. The signal processer in the RDA performs additional data processing to convert the analog Level I data to digital Level II

data products. These products include reflectivity, mean radial velocity, and spectrum width (Vasquez, p. 39, 2015). This study used Level II Reflectivity data.

If the diameter of a particle is one-fifth of the wavelength or less, Rayleigh scattering will occur, and the energy will be scattered in all directions, including back toward the antenna. If the particle is larger, Mie scattering will occur, and the energy will be forward-scattered, meaning that very little energy will return to the antenna. In the case of the WSR-88D, with a wavelength of 10cm, the threshold between Rayleigh and Mie scattering is a diameter of 2cm (Vasquez, p. 21, 2015).

The electromagnetic pulse emitted by the transmitter lasts for 1.57 microseconds. Because the pulse is traveling at the speed of light, this means that the pulse of energy extends over a length of 470 meters (Vasquez, 2015, p. 25; NEXRAD Tech Info). Additionally, the energy transmitted has a beamwidth of 1° (NEXRAD Tech Info). Therefore, each pulse encompasses and samples a 3-dimensional volume of the atmosphere as it moves outward from the antenna. As it moves outward, this volume of space expands, meaning that with an angular beamwidth of "1 degree a pulse 100 miles from the radar will be two miles in height and width. As a result, features in a large volume like this that are much smaller than the beam width tend to show a weaker power return and might give false indications of intensity" (Vasquez, 2015, p. 26). Also, because even the lowest tilt angle of the antenna is greater than zero degrees, the volume

that is being sampled is much higher than the elevation of the antenna at far distances (Vasquez, 2015, p. 30).

Between each pulse the antenna is in a listening period. While the pulse is only 1.57 microseconds long, the listening period is 998 microseconds long. This means that the "total time the radar is actually transmitting a signal (when the duration of transmission of all pulses, *each hour*, are added together), is for a little over 7 seconds each hour. The remaining 59 minutes and 53 seconds are spent listening for any returned signals" (NWS, n.d.b).

### 2.2.2 Volume Coverage Pattern

The antenna is capable of rotating 360° and can tilt between 0.5° and 19.5° above horizontal (NEXRAD Tech Info). The antenna follows a predetermined scanning routine, called the volume coverage pattern (VCP). The VCP describes the number of pulses per second, the number of tilt angles used, and the time it takes to complete one scan. A number of VCPs exist, and each is designed to maximize the efficiency of the radar unit for particular meteorological conditions. For example, VCP 12 scans 14 tilt angles in 4.5 minutes and is most suitable for rapidly changing convective (thunderstorm) conditions, while VCP 32 has only 5 tilt angles and completes its pattern in 10 minutes, so it is best suited for clear conditions (Vasquez, 2015, p. 145). Generally, the two lowest tilt angles are scanned twice to reduce ground clutter and other anomalies (OFCM, 2006, p. 2-2).

This study used the lowest tilt angle of 0.5° in order to analyze conditions closest to urban and non-urban land cover surfaces.

### 2.2.3 Reflectivity

Reflectivity is the radar product used in this study. Reflectivity is a "measure of the efficiency of a target in intercepting and returning radio energy. With hydrometeors, it is a function of the drop size distribution, number of particles per unit volume, physical state (i.e., ice or water), shape, and aspect" (OFCM, 2021, p. 18). The backscattered energy is initially calculated as equivalent reflectivity ($Z_e$), which is determined by the following equation:

$$Z_e = (P_r\, r^2)/(C_r L_a),$$

where $P_r$ is power received, $r$ is range, and $L_a$ is a constant representing attenuation factor (Vasquez, 2015, p. 22). $C_r$ is a "radar constant" for that particular hardware, and includes "transmitter power, antenna gain, beam width, pulse width, dielectric constant, and wavelength" (NWS, n.d.b). Range ($r$) must be accounted for because reflectivity is stronger for near targets and weaker for more distant targets. Once range is accounted for, the primary determinant of equivalent reflectivity is the power received, which is an expression of how much backscattered energy is received by the antenna. This equation assumes that Rayleigh scattering is occurring (Vasquez, 2015, pp. 22, 25).

Equivalent reflectivity is expressed in $mm^6m^{-3}$, which describes the number of drops and their size per cubic meter. This could range from $10mm^6m^{-3}$ in light snow to $10^6mm^6m^{-3}$ in hail (Vasquez, 2015, p. 25). To make reflectivity values easier to use, the following equation applies a logarithmic scale to $Z_e$ to produce decibels of equivalent reflectivity (dBZ) (Vasquez, 2015, p. 25):

$$\mathrm{dBZ} = 10\ \log_{10} Z_e$$

Decibels of equivalent reflectivity (dBZ) are "widely used in radar meteorology and may be described simply as *intensity*" (Vasquez, 2015, p. 25).

Two products derived from radar data that are seen online or in the media are the Base Reflectivity product and Composite Reflectivity product. Base Reflectivity shows the intensity of radar energy received at the lowest tilt angle of 0.5°. Composite Reflectivity displays the "maximum echo intensity (reflectivity) from any elevation angle," and "this product is used to reveal the highest reflectivity in all echoes"(NWS, n.d.a). The public can access either product at radar.weather.gov. Media such as television weather broadcasts usually use the Base Reflectivity product (NOAA, 2023). As previously mentioned, this study used data gathered at the 0.5° tilt angle, so that while the Base

Reflectivity product was not used, the dataset created for this work is similar to the Base Reflectivity product.

### 2.2.4 WSR-88D Limitations

Weather radar is well-suited for studying thunderstorms because of its high temporal resolution. Information about thunderstorms can also be gathered by weather satellites and lightning detection systems; however, while thunderstorms are visible in satellite imagery, their extent may be obscured by surrounding clouds, and their intensity is not directly measured. Lightning data also would not provide a direct measure of intensity.

However, weather radar has some important limitations. It is restricted to line-of-sight, so high terrain obstructs radar energy. Also, weather radar may return data for non-meteorological phenomena, known as clutter. Clutter may be generated by smoke from wildfires; large groups of birds, bats, or insects; and wind turbines at wind farms. Also, when the "radar dish is pointing into the rising or setting sun, a spike may be seen extending in this direction" (Vasquez, 2015, p. 50).

Stationary objects on the ground, such as buildings, trees, and terrain, can be filtered out by the WSR-88D ground clutter algorithm. However, wind turbines remain a problem:

> NEXRAD's clutter filter scheme only removes clutter that is stationary, such as buildings, trees, and terrain. Unfortunately, both precipitation and wind turbine blades are moving and the filter is not applied to them. Trying to filter out moving blades will inevitably alter how the radar sees real precipitation…. A single radar volume sample (gate) at 48 km (26nm) from the radar is approximately a square kilometer. Thus, for a typical wind farm, the radar may receive reflected energy from many turbines within that gate, each with multiple rotating blades. These numerous rotating blades appear similar to precipitation, which is also made up of numerous distributed moving targets…. There are fewer blades than raindrops within a sample volume, but the blades make up for their smaller numbers by reflecting significantly more energy back to the radar. However, the radar has no way to determine the number of targets it is sampling within a particular gate. Also, the reflected energy is constantly changing as the blades change their pitch and orientation relative to radar, with some blades moving towards the radar, some moving away, and some not appearing to move at all (perpendicular). This is analogous to the movement of precipitation within a volume sample. (NOAA, 2022a)

The portion of the atmosphere sampled by each tilt of the radar should also be considered. The volume samples the atmosphere at higher and higher altitudes above the elevation of the antenna as the pulse moves outward. Also, the size of the pulse volume increases with distance from the antenna.

### 2.3 Urban Effects on Local Climate

In their 2016 book, *Microclimate and Local Climate*, Roger G. Barry and Peter D. Blanken define a local climate, or topoclimate, as existing on a scale of hundreds of meters to approximately 10km and encompassing an area less than 100km$^2$. A topoclimate, or "topographical climate," is one that is defined by valleys, slopes, and ridges, as well as urban areas. Their chapter on urban climates summarizes some of the foundational works of T. R. Oke on the causes of the urban heat island. The urban heat island (UHI) is defined as the "difference between air temperature in the urban and adjacent rural areas" (Barry & Blanken, 2016, p. 250), and it involves a number of factors, including changes in the radiation balance due to alterations in atmospheric composition and albedo; the production of heat by buildings and traffic; the reduction of heat diffusion due to surface roughness; and the increase in sensible heat due to the reduction of evapotranspiration (Barry & Blanken, 2016).

Shem and Shepherd (2009) applied Weather Research and Forecast (WRF) numerical simulations to two case studies for Atlanta, GA. The simulation of the two storms, which occurred in August of 2002 and July of 1996, followed the actual events closely. The authors then changed the parameters of the model by removing the urban land cover that represented Atlanta, and replaced it with dryland, cropland, and pasture. When they ran the simulation without the urban land cover, they found that the thunderstorms started at the same time as before, showing that initiation of the storms was not affected; however, significantly more rainfall occurred in the urban model. The authors discussed several limitations, particularly the omission of urban aerosols.

Debbage and Shepherd (2015) studied the spatial contiguity and temperature intensity of the fifty most populous cities in the United States. Spatial contiguity in this regard describes the level of fragmentation of urban land cover. Earlier research on the "spatial configuration of cities" had come to conflicting conclusions that "both sprawling and high-density urban development can amplify urban heat island intensities" (Debbage & Shepherd, 2015, p. 181). Sprawling cities cover more land area and impair heat circulation in the city center, while higher-density cities experience heat storage due to urban canyon geometry. Using Ordinary Least Squares (OLS) regression modeling Debbage and Shepherd (2015) determined that the urban heat island is "generally magnified by more contiguous urban development, across a variety of urban intensity levels," suggesting that "sprawling and high-density city configurations both have the propensity to increase UHI intensities if the urban development is highly contiguous" (Debbage & Shepherd, 2015, p. 190). Additionally, they found no significant correlation between UHI intensity and population.

A study using 108-year-old precipitation records, observations from a global climate observing system, and satellite data analyzed changes in precipitation in two arid cities: Phoenix, AZ, and Riyadh, Saudi Arabia (Shepherd, 2006). The study found that during the monsoon season of July through September, the suburbs of Phoenix had a 12 – 14% increase in precipitation between the "pre-urban" time period (1895 – 1949) and "post-urban" time period (1949 on) even though the area was experiencing a severe drought. The author hypothesizes that this is due to "urban-topographic interactions and possibly

irrigation moisture" (Shepherd 2006, p. 607). The results for Riyadh are less clear. There was an increase in precipitation near the city, but there was also an increase 50 to 100km north of the city as well. The author calls for more research to investigate the relationship between urban land use and precipitation, especially the role of aerosols (Shepherd, 2006).

Niyogi et al. (2011) studied the interaction of the Indianapolis urban region with thunderstorms during the summer months of May through August, from 2000 to 2009. Using radar data, the authors studied the occurrences of 91 thunderstorms as they passed the urban environment and moved over rural areas nearby. The "observed base-reflectivity radar plots were subjectively analyzed for storm-structure changes. Storm-composition change was noted when the change occurred (initiated, split, intensified, or dissipated) in and around the urban or rural region" (Niyogi et al., 2011, p. 1131). The researchers found that "more than 60% of storms changed structure over the Indianapolis area as compared with only 25% over then rural regions" (Niyogi et al., 2011, p. 1129). Additionally, simulations of a typical thunderstorm event were conducted with and without the urban land cover and it was found that the storm could not be reproduced without the urban land cover.

Bentley et al. (2010) created a 10-year "radar climatology" for the Atlanta, GA, region, using WSR-88D radar data for June through August of 1997 to 2006, with a 5-minute temporal resolution. They found that a large number of high-reflectivity storm events

occurred near the downtown area, confirming their hypothesis that the UHI of Atlanta affects thunderstorm development. However, the research does not seem to consider conditions outside of the city area. Potentially there are an equal number of storms in rural areas of Georgia as well. In a slightly more recent study, Ashley et al. (2012) used the same concept of "radar climatologies" for major and medium-sized cities in the Southeastern United States. They found that urban areas increased thunderstorm frequency and intensity, and that these changes corresponded to the size of the urban footprint. This study also attempted to overcome past critiques of similar studies by including nearby control regions, in addition to large and medium-sized cities.

Haberlie et al. (2015) further expanded upon the study of thunderstorm activity near Atlanta. The authors used 17 years of weather radar data, for the months of May through September, from 1997 to 2013, to extract isolated convective initiation (ICI) events. They defined ICI events through their selection criteria, analyzing the "evolution of each cell through time, requiring that qualifying events must meet or exceed a 40dBZ threshold, their location must be greater than 30km from existing convection, and the initial cluster of convectively active pixels must develop into a storm that lasts at least 30 min." (Haberlie et al., 2015, p. 666). The authors created two study areas, one centered over the Atlanta urban area and one 150km west of the city as a rural control region. Haberlie et al. (2015)'s results were similar to other studies. ICI events occurred more often over and downwind of Atlanta than over nearby rural areas. Additionally, more ICI events occurred on weekdays than weekends, suggesting that aerosols linked to work commute schedules may also play a role. Aerosols have been found to exist in higher quantities on

weekdays than weekends due to fluctuations in commuting traffic. Haberlie et al. (2015) compared the numbers of ICI events on weekdays to weekends and found that, in the city, there were significantly more ICI occurrences during the week, especially Tuesdays and Wednesdays, than on the weekends. Rural areas showed very little difference in the frequency of ICI events based on day of the week.

Haberlie et al. (2015) also collected the prevailing wind at 700 hPa (pressure altitude approximating 10,000 feet above sea level) at the beginning of each ICI event. They divided both the Atlanta area and the rural control area into two halves, upwind and downwind. Then they divided the ICI events that occurred into upwind and downwind groups. There was no significant difference in the groups for the rural control area, but there were significantly more ICI events downwind of the Atlanta area. Finally, the authors caution that complex terrain can influence thunderstorm activity in ways that may complicate studies of urban effects on convective activity (Haberlie et al., 2015).

Lazzarini et al. (2015) addressed the effect of vegetation in urban areas located in arid and semiarid regions. Compared to more temperate regions, cities in arid regions may have a greater amount of vegetation in urban areas than in undeveloped areas, potentially creating a cooling, “oasis effect” (Lazzarini et al., 2015, p. 9981). Lazzarini et al. (2015) found that in hot desert cities, “downtown areas [are] on average cooler than suburbs during the daytime (urban cool island) and warmer at night (classical UHI)” (p. 9980), due to “the abundance of irrigated vegetation in [hot desert cities], which creates large

areas where surface temperatures are mitigated through evaporative cooling," as well as "shadowing and plant transpiration" (p. 9987). Additionally, moisture that would not normally be present in dry regions is added to the atmosphere by vehicle exhaust, air conditioners, swimming pools, and irrigation of lawns and agriculture (Lazzarini et al., 2015; Oke et al., 2017).

An understanding of the genesis of thunderstorms and the development of a local "radar climatology" (Myers, 1964) should allow for new insights into the characteristics of thunderstorms in the Albuquerque urban area over the last twenty years. The Albuquerque metropolitan area differs from other cities used in studies of urban climate effects on local precipitation and thunderstorms, due to its smaller population and differing climate and topography. Atlanta, GA, has a population of nearly 6 million people, and is located in flat terrain. Phoenix, AZ, has a population of 4.8 million people, and is situated in a hotter, drier climate where irrigated lawns and agriculture provide stark contrast to the natural environment. A study of Albuquerque's unique climate, topography, and city morphology adds to knowledge of thunderstorms and urban effects on local climate.

## 3.0 Methods

To evaluate the effects of urbanization on thunderstorms in central New Mexico, regression analyses were used to compare thunderstorm intensity and duration to measures of urban land cover and relevant control variables (e.g., temperature, elevation, surface aspect, and other land cover characteristics). The following subsections provide details about the study area, the data used, and the analytical methods applied.

### 3.1 Study Area

The study area contains much of New Mexico's Middle Rio Grande region, which "includes the urban environments of Albuquerque, Santa Fe, as well as surrounding small towns, and rural agricultural communities" (Benson et al., 2014, p. 196). The spatial extent of the study area is based on the limitations of Albuquerque NEXRAD WSR-88D weather radar station (KABX). As noted earlier, weather radar is subject to a number of limitations, including line-of-sight. Due to the mountainous terrain of central New Mexico, radar coverage does not extend to its full range in all directions. Figure 1 depicts the study area, KABX, and urban areas in relation to terrain.

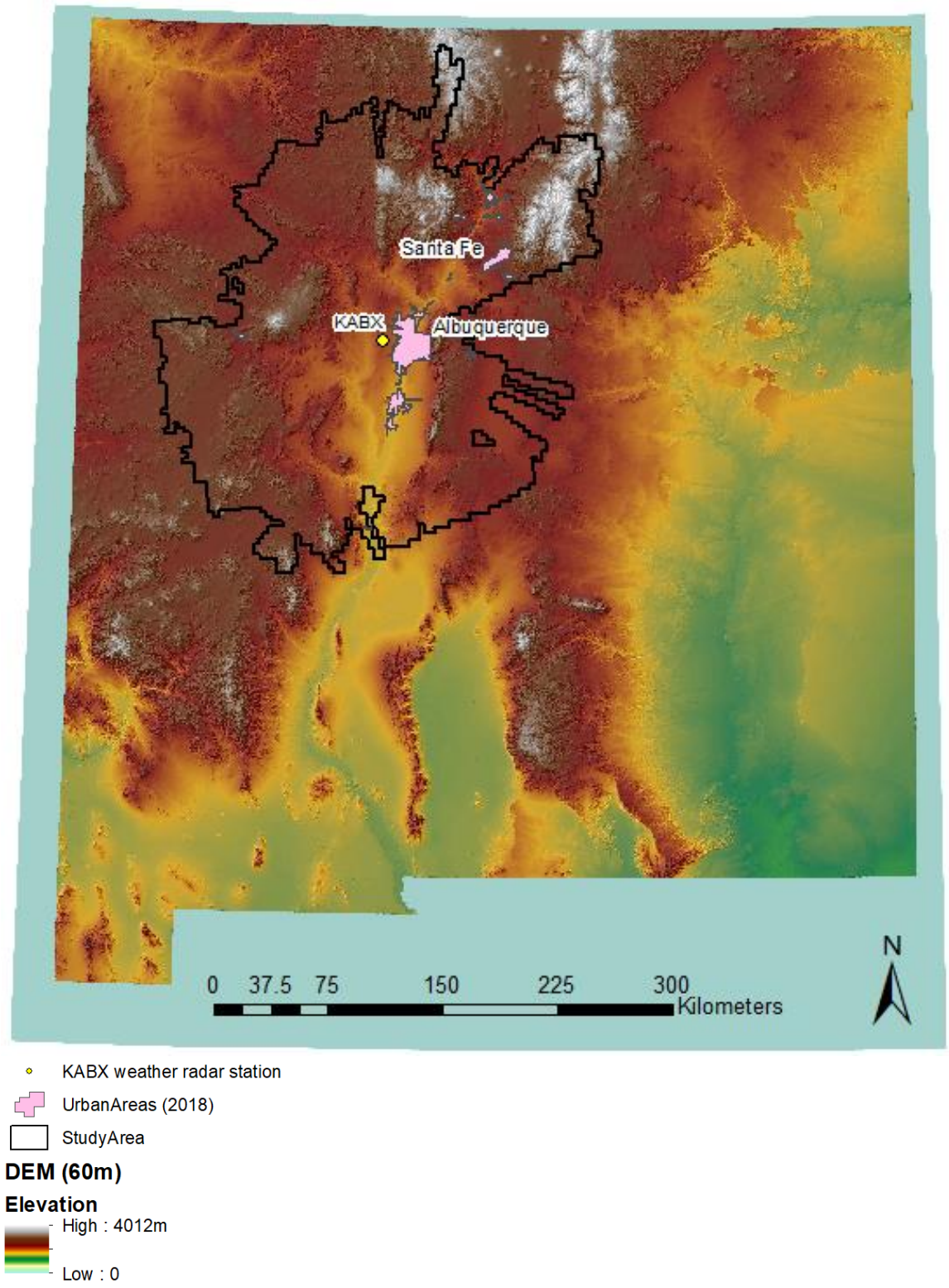



Figure 1. Study area, terrain, urban areas, and the KABX weather radar antenna relative to the state of New Mexico.

### 3.1.1 Topography and Climate of the Study Area

The KABX station is centrally located in New Mexico at 35.14972° N, 106.82333° W, at an elevation of 1,789m (5,870 feet) above sea level. It is located on a mesa immediately west of the city of Albuquerque.

The elevation of the study area varies from 1,363m (4,472 feet) to 3,943m (12,936 feet) above sea level. The Rio Grande, and adjacent farmland and Bosque (i.e., riparian zone), run north-south through the study area. There are several areas of mountainous terrain to note as well. Mount Taylor is approximately 70km west of KABX, and the Sandia and Manzano Mountains are immediately east and southeast of the city of Albuquerque. The southern extent of the Jemez Mountains begins approximately 55km north of KABX, and the Sangre de Cristo Mountains are immediately east and northeast of Santa Fe.

The elevation of the study area varies by more than 2,600m (8,000 feet), and thus the region contains a number of different climate classifications. Within the study area, there are regions classified as arid (BSk and BWk); warm temperate (Cfb) in the mountainous areas; and boreal (Dfb and Dfc) in the high mountain regions. Albuquerque is located in a BSk (arid) region with a warm temperate Cfb region in the Sandia Mountains nearby to the east. Santa Fe is also located in a warm temperate region in close proximity to boreal Dfb and Dfc areas in the mountains to the east and northeast (Kottek et al., 2006). Figure 2 shows the monthly average near-surface temperatures in and around the study area for

June 2018, which was selected as an illustrative example. Figure 3 depicts the land cover in and around the study area as of 2019.

This study focused on the months of May through September, from 2001 to 2020. These five months were selected for a number of reasons. For the most part, these months are free of "the large-scale processes that drive [synoptic-scale cyclones that] may mask any urban effects," and it has been found that "urban influences are greatest during periods of weather dominated by convection, especially during the warm season when heating of the Earth's surface generates uplift" (Oke et al., 2015, p. 282). The North American monsoon occurs in New Mexico in July and August (Sullivan et al., 2018, p. 16), and this study period is meant to capture thunderstorm activity before, during, and after that time. Oke et al. (2015, p. 282), however, warn that both monsoon flow and complex topography may make it difficult to discern urban influences on thunderstorms.

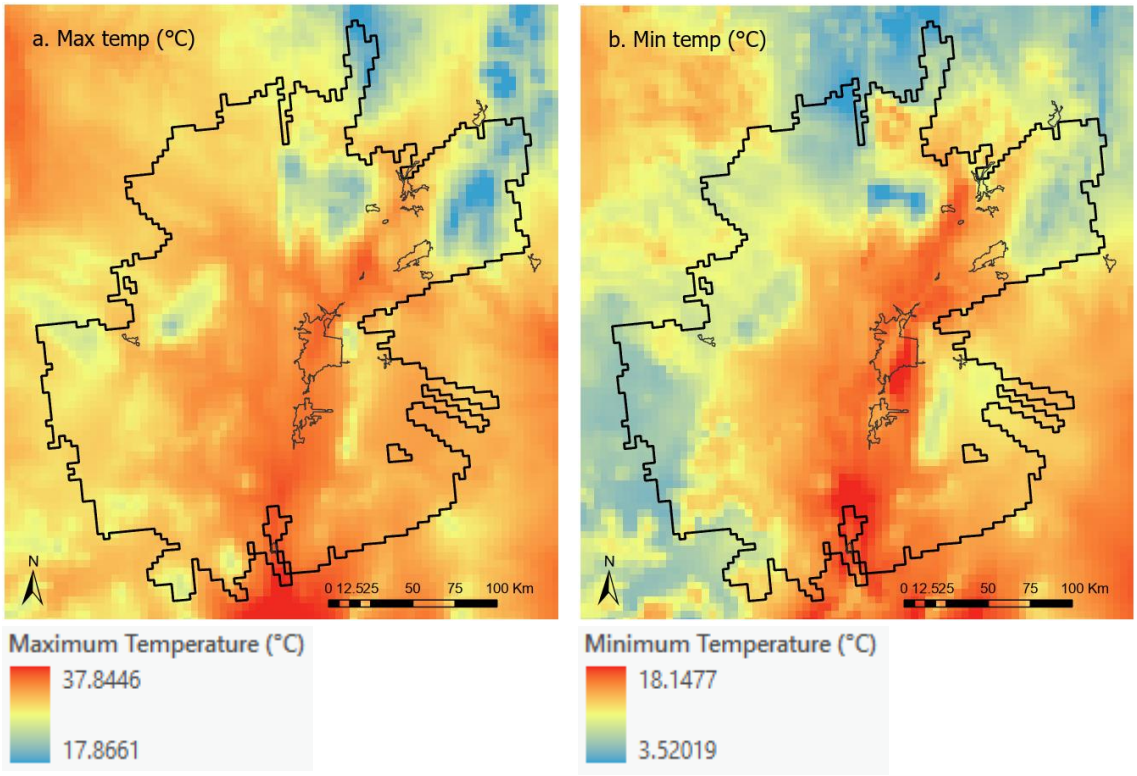



Figure 2. Monthly (a.) maximum and (b.) minimum near-surface temperature in and around the study area for June 2018, with urban areas outlined in grey.

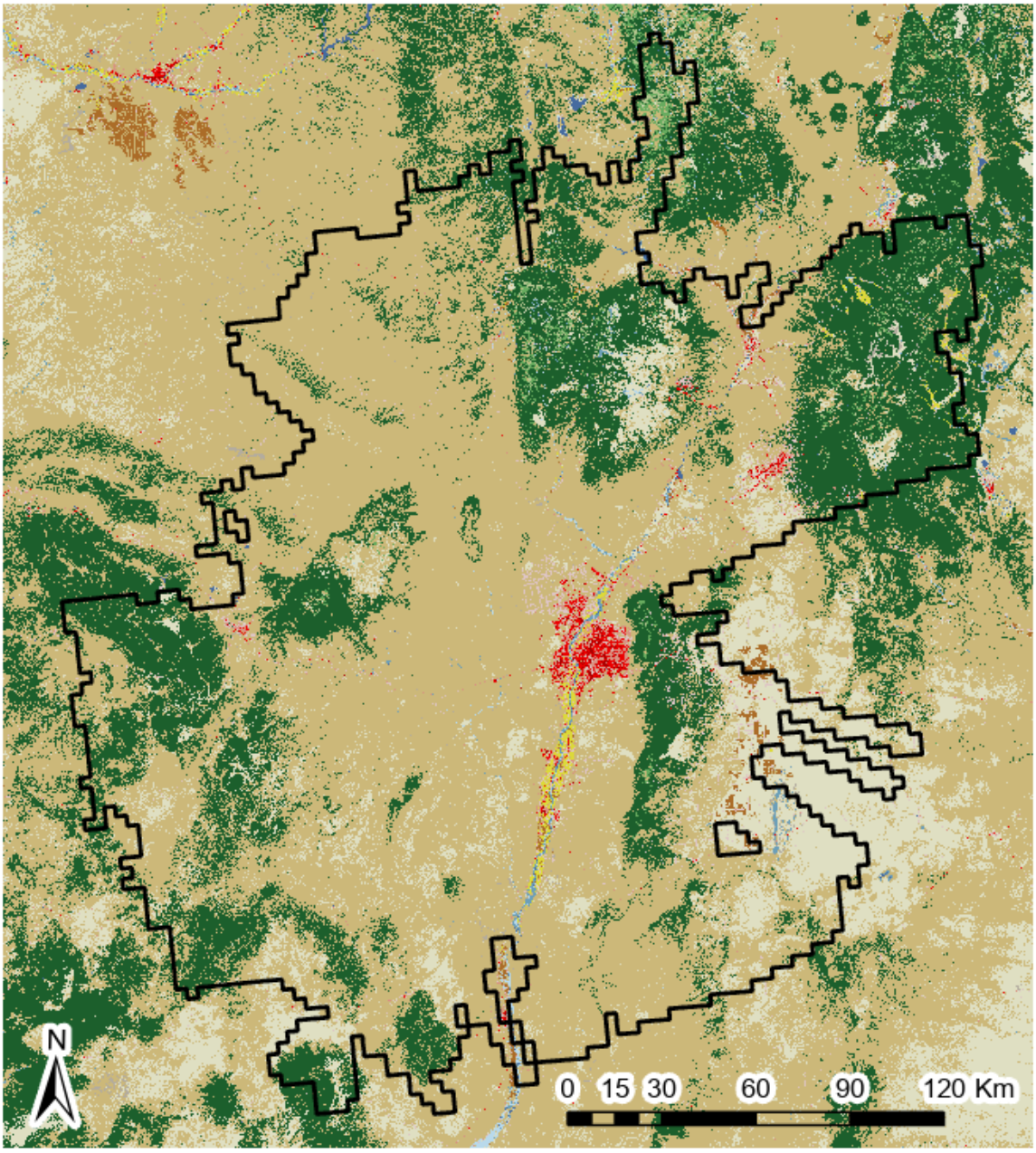


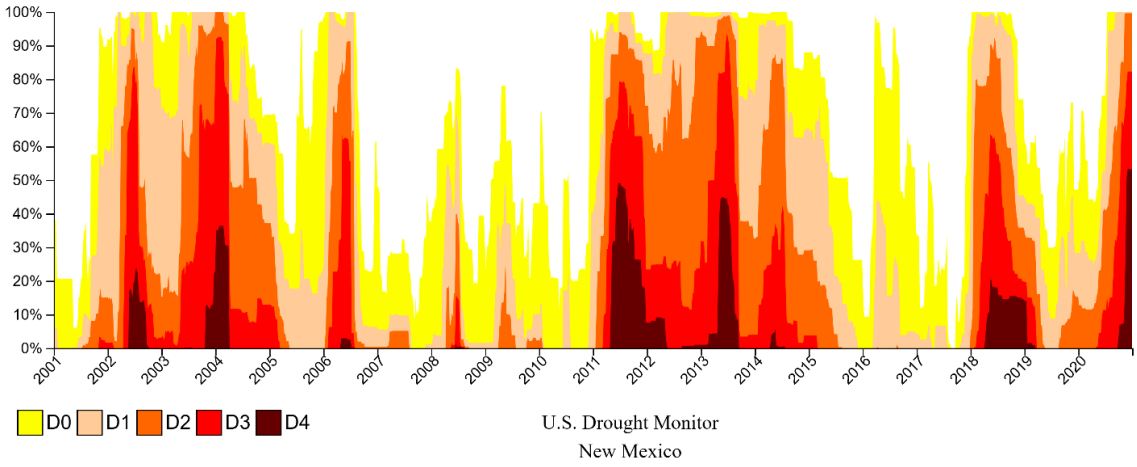

Figure 3. Land cover in and around the study area, as shown in the 2019 NLCD dataset. See Figure 4 for the classification legend and Table 5 Section 3.2.2 for descriptions of each class.


Figure 4. NLCD legend. Source: https://www.mrlc.gov/sites/default/files/NLCD_Colour_Classification_Update.jpg.

Drought may be an additional influence on the occurrence and severity of thunderstorms in New Mexico. The southwestern United States, which includes most of New Mexico, experiences periodic droughts. Shortly before the time period for this study began, the region experienced its "wettest 19-year period (1980 – 1998) in at least 1200 years" (Williams et al., 2020). However, during most of the study period, from 2000 to 2018, drought conditions in the region were "the second driest 19-year period since 800 CE, exceeded only by a late-1500s megadrought" (Williams et al., 2020).

Figure 5 shows drought severity between 2001 and 2020, measured weekly, for New Mexico and Bernalillo County, which is located centrally within the study area and contains most of the Albuquerque metropolitan area. The legend, contained in Table 1, details the impacts of each level of severity, specific to New Mexico. There are a few distinct differences between the conditions averaged over the state and those within Bernalillo County. Specifically, in Bernalillo County the drought was less severe at the end of 2003 and into 2004, and through 2011 and 2018. The drought was more pronounced in Bernalillo County than in the state as a whole during the middle of 2013. A more detailed look at the relationship between drought and thunderstorm activity would require further research, beyond the scope of this study.

a.




b.

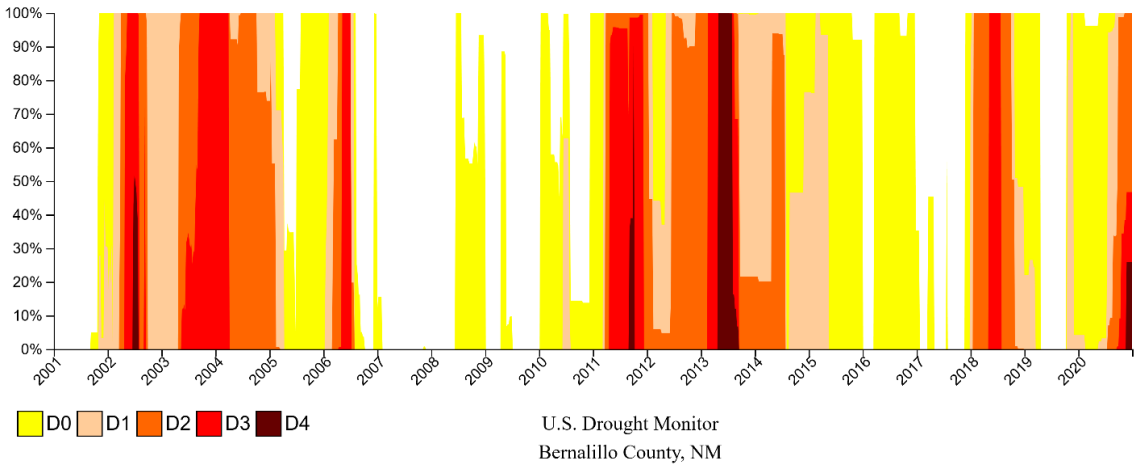



Figure 5. (a.) Drought conditions for the state of New Mexico, 2001 – 2020, measured weekly. Data current as of August 9, 2022. (https://www.drought.gov/states/new-mexico). (b.) Drought conditions for Bernalillo County, 2001 – 2020, measured weekly. Data current as of November 28, 2022. (https://www.drought.gov/states/New-Mexico/county/Bernalillo). See legend in Table 1, which is specific to the state of New Mexico.

Table 1. Drought conditions legend. The categories of historically observed impacts are specific to New Mexico.

| Category | Historically observed impacts |
|---|---|
| D0 | Soil moisture is low |
| | Fire danger increases |
| D1 | Livestock need supplemental feed and water |
| | Burn bans and firework restrictions begin |
| D2 | Pasture yield is limited; producers sell livestock |
| | Irrigated crops are stunted; dryland crops are brown |
| | Dust storms occur |
| | Abundance and magnitude of wildfires may increase; fuel mitigation practices are in effect |
| | Wildlife feeding patterns change |
| | Well water decreases |
| D3 | Livestock are suffering; producers are selling herds; feed costs are high; emergency CRP grazing is authorized; crop yields are low |
| | Fire danger is extreme |
| | Irrigation allotments decrease |
| | Vegetation and native trees are dying |
| D4 | Federal lands begin to close for fire precautions; burn bans increase |
| | Bears encroach on developed areas; migratory birds change patterns |
| | No surface water is left for agriculture, farmers use private wells |
| | Rio Grande and other large rivers are dry |

Source: https://droughtmonitor.unl.edu/DmData/StateImpacts/ImpactExamples.aspx

### 3.1.2 Population Change in the Study Area

Based on US Census estimates, the population of the Albuquerque metropolitan area has grown from 532,100 people in 2001 to 692,387 people in 2020 (30% increase). This total includes Albuquerque, Rio Rancho, Corrales, the town of Bernalillo, and Los Ranchos de Albuquerque.

Albuquerque itself grew from 457,627 in 2001 to 564,648 in 2020 (23% increase). Especially of note, Rio Rancho grew from a population of 54,893 in 2001 to 104,257 in 2020, a 90% increase, making it the second largest urban area in New Mexico by the end of the study period. Additionally, Santa Fe grew from 65,699 people in 2001 to 87,684 people in 2020 (33% increase). Population changes within the study area are detailed in Table 2.

Table 2. Population and percent population change between 2001 and 2020 in urban centers within the study area (does not include sparsely populated areas). Source: US Census.

| Name | | 2001 Population | 2020 Population | Percent change from 2001 to 2020 |
|---|---|---|---|---|
| **Albuquerque metropolitan area** | | | | |
| Albuquerque | | 457,627 | 564,648 | 23% |
| Rio Rancho | | 54,893 | 104,257 | 90% |
| Corrales | | 7,277 | 8,517 | 17% |
| Town of Bernalillo | | 6,893 | 9,084 | 32% |
| Los Ranchos de Albuquerque | | 5,410 | 5,881 | 9% |
| | **Subtotal** | 532,100 | 692,387 | 30% |
| | | | | |
| **Population centers outside of Albuquerque metropolitan area** | | | | |
| North of Albuquerque | Santa Fe | 65,699 | 87,684 | 33% |
| | Espanola | 10,159 | 10,514 | 3% |
| South of Albuquerque | Los Lunas | 10,401 | 17,370 | 67% |
| | Bosque Farms | 3,946 | 4,053 | 3% |
| | Peralta | 3,485 | 3,380 | -3% |
| | Belen | 7,252 | 7,386 | 2% |
| | Socorro | 8,871 | 8,549 | -4% |
| West of Albuquerque | Grants and Milan | 11,343 | 11,672 | 3% |
| | **Subtotal** | 121,156 | 150,608 | 24% |
| | | | | |
| | **Total** | 653,256 | 842,995 | 29% |

### 3.1.3 Land Cover Change in the Study Area

In addition to an increasing population, changes in land cover have also occurred within the study area. The Multi-Resolution Land Characteristics (MRLC) Consortium releases the National Land Cover Database (NLCD) 30m raster dataset every two to three years (this dataset is discussed in more detail in Section 3.2.2). The 2001 dataset corresponds with the beginning of the study period and the 2019 dataset corresponds most closely with the end of the study period. See Table 5 in Section 3.2.2 for detailed descriptions of each land cover class.

The following noteworthy changes occurred within the study area between the 2001 and 2019 datasets, and are summarized in Table 3. All developed land cover classes combined increased by 13.94%, high-intensity developed land cover increased by 46.94%, and medium-intensity development increased by 38.56%. Both evergreen and deciduous forest land cover decreased, especially in mountainous regions where wildfires had occurred, and were typically replaced by grassland/herbaceous or shrub/scrub vegetation classes. Especially of note is the increase in developed land cover on the northwest periphery of the Albuquerque metropolitan area, as well as the infill in the center of the city, as seen in Figure 6 and Figure 7.

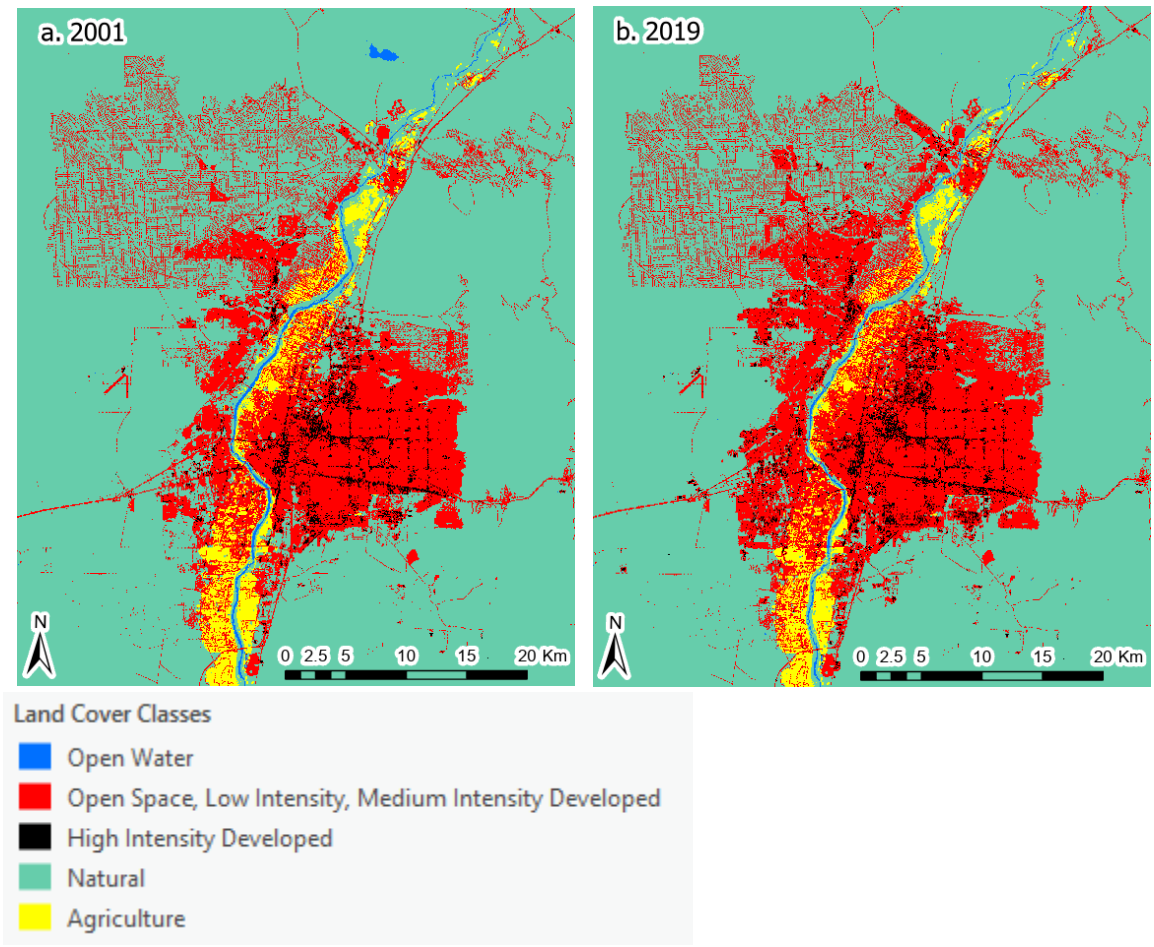



Figure 6. Developed land cover, open water, agricultural land cover, and natural land cover in the Albuquerque metropolitan area, in (a.) 2001 and (b.) 2019.

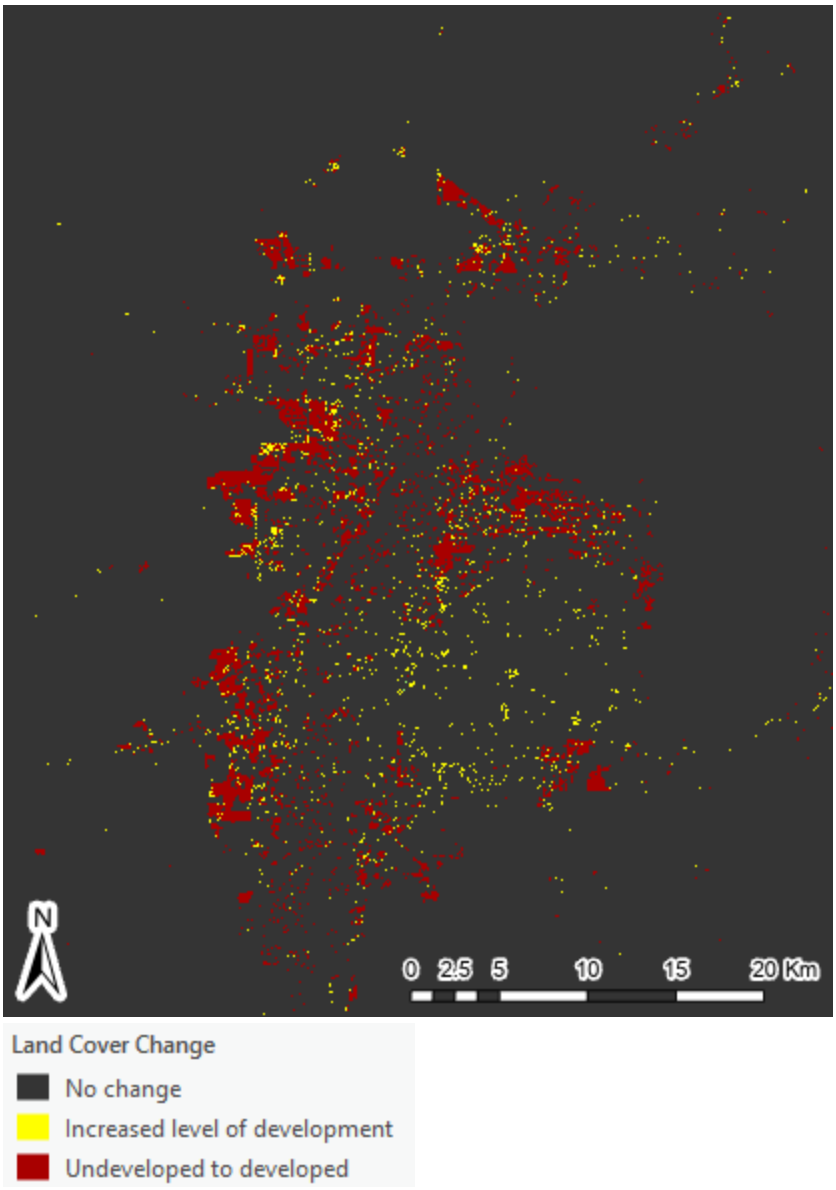



Figure 7. Changes in land cover between 2001 and 2019 in the Albuquerque metropolitan area are shown. Changes from undeveloped land cover to developed land cover are shown in red. Changes from one class of developed land cover to a higher level of development are shown in yellow.

Table 3. Area of each land cover class within the study area, percent of the study area covered by each land cover class, and percent land cover change, based on the 2019 NLCD product. See Table 5 for descriptions of each land cover class.

| Land Cover Class (NLCD classification code) | Area of land cover (ha) 2001 | Percent of total hectares 2001 | Area of land cover (ha) 2019 | Percent of total hectares 2019 | Percent change from 2001 to 2019 |
|---|---|---|---|---|---|
| Note: the study area is 5,508,056 hectares (ha). | | | | | |
| Developed, high intensity (24) | 4,564 | 0.08% | 6,706 | 0.12% | 46.94% |
| Developed, medium intensity (23) | 18,642 | 0.34% | 25,830 | 0.47% | 38.56% |
| Developed, low intensity (22) | 29,568 | 0.54% | 32,338 | 0.59% | 9.37% |
| Developed, open space (21) | 63,833 | 1.16% | 67,987 | 1.23% | 6.51% |
| Medium, low, and open space developed classes (21, 22, 23) | 112,042 | 2.03% | 126,156 | 2.29% | 12.60% |
| All developed land cover classes (21, 22, 23, 24) | 116,606 | 2.12% | 132,862 | 2.41% | 13.94% |
| | | | | | |
| Barren land (rock/sand/clay) (31) | 8,299 | 0.15% | 8,893 | 0.16% | 7.15% |
| | | | | | |
| Deciduous forest (41) | 41,898 | 0.76% | 38,974 | 0.71% | -6.98% |
| Evergreen forest (42) | 1,436,596 | 26.08% | 1,367,753 | 24.83% | -4.79% |
| Mixed forest (43) | 18,544 | 0.34% | 18,929 | 0.34% | 2.08% |
| All forest land cover classes (41, 42, 43) | 1,497,038 | 27.18% | 1,425,657 | 25.88% | -4.77% |
| | | | | | |
| Grassland/herbaceous (71) | 542,084 | 9.84% | 570,914 | 10.37% | 5.32% |
| Shrub/scrub (52) | 3,266,480 | 59.30% | 3,289,202 | 59.72% | 0.70% |
| | | | | | |
| Pasture / hay (81) | 27,743 | 0.50% | 25,006 | 0.45% | -9.87% |
| Cultivated crops (82) | 14,334 | 0.26% | 19,244 | 0.35% | 34.25% |

Table 3 (cont.)

| Land Cover Class (NLCD classification code) | Area of land cover (ha) 2001 | Percent of total hectares 2001 | Area of land cover (ha) 2019 | Percent of total hectares 2019 | Percent change from 2001 to 2019 |
|---|---|---|---|---|---|
| Woody wetlands (90) | 12,844 | 0.23% | 14,146 | 0.26% | 10.14% |
| Emergent herbaceous wetlands (95) | 16,748 | 0.30% | 17,557 | 0.32% | 4.83% |
| Open water (11) | 5,881 | 0.11% | 4,577 | 0.08% | -22.17% |

## 3.2 Data

This study used two primary types of data: (1) thunderstorm activity represented by a time series of weather radar data and (2) the National Land Cover Database (NLCD) dataset. Thunderstorm activity was sampled by NEXRAD WSR-88D weather radar and used to derive a "radar climatology" (Bentley et al., 2010; Myers, 1964) of the Middle Rio Grande. Weather radar data were used to determine what effects developed urban land cover has had on thunderstorm activity in the study area. Table 4 summarizes the data sources used in the study and the variables derived from those data.

Weather radar data were clustered by time and location into polygons representing thunderstorm events, which were used as the unit of analysis. These polygons were then used to collect land cover data within their boundaries, as well as other surface and near-surface data including temperature, elevation, and aspect.

Table 4. Data formats and sources.

| Primary Data | Format | Use | Source | Derived Variables |
|---|---|---|---|---|
| NEXRAD WSR-88D weather radar data | Vector shapefiles | Represent thunderstorm events with ≥40dBZ radar reflectivity | NOAA's Weather and Climate Toolkit (WCT) https://www.ncdc.noaa.gov/wct/ | Maximum intensity of thunderstorm events |
| | | | | Mean intensity of thunderstorm events |
| | | | | Duration of thunderstorm events |
| National Land Cover Database (NLCD), 2019 release | Raster | Land cover typology for 2001 through 2020 | Multi-Resolution Land Characteristics (MRLC) Consortium https://www.mrlc.gov/data | Land cover typology |
| **Secondary Data** | **Format** | **Use** | **Source** | **Derived Variables** |
| PRISM Monthly Temperature | Raster | Measure of the condition of the atmosphere near the surface | PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu | Mean monthly temperature |
| | | | | Maximum monthly temperature |
| | | | | Minimum monthly temperature |
| Digital Elevation Model (DEM) | Raster | Measure of elevation, slope, and aspect | RGIS https://rgis.unm.edu/ | Elevation |
| | | | | Slope |
| | | | | Aspect |
| **Auxiliary Data** | **Format** | **Use** | **Source** | **Derived Variables** |
| Metropolitan Statistical Areas (MSAs) | Vector shapefiles | Define boundaries of urban areas | US Census https://www.census.gov/cgi-bin/geo/shapefiles/index.php | For user reference |

### 3.2.1 NEXRAD Weather Radar Data

The study area boundary was created based on the area sampled by the Albuquerque NEXRAD WSR-88D weather radar station. No GIS-friendly layer existed which depicted that boundary; however, an ungeoreferenced image showing NEXRAD radar coverage was available from the National Oceanic and Atmospheric Administration (NOAA), National Weather Service (NWS) Radar Operations Center (ROC) (NOAA, 2022c). This coverage image was georeferenced and the outline of the service area for KABX was digitized to delineate the extent of the study area as seen in Figure 1. The digitized study area was created using the NAD 1983 UTM Zone 13N (EPSG: 26913) spatial reference system.

NEXRAD weather radar data, like many other types of earth science data, is available online and for free to the public. NEXRAD data can be downloaded manually, through the Weather and Climate Toolkit (WCT; https://www.ncdc.noaa.gov/wct/), or by creating a script that will access, download, and convert the format of the data. The WCT is free to download from NOAA's National Centers for Environmental Information (NCEI) and it includes a graphical user interface that allows users to select and download the data they require. The Toolkit's limitation, in regard to this study, is that data can be downloaded for only one day at a time and downloading subsequent days' data must be started manually.

An example Python script for downloading and converting radar data to shapefiles was found in the supplementary materials in Ansari et al. (2018). The script allowed for data to be downloaded one month at a time and was modified to select data for the KABX radar station. The final script can be found in Appendix B. The Python script uses the Weather and Climate Toolkit to convert the data from its original binary format to polygon shapefiles. The call function in the script uses two files from the WCT, wct-export.bat, which is a Windows Batch file written in Java, and wctBatchConfig.xml, which is an xml file containing user settings. The XML file was modified to download only polygons with a value of 40 dBZ or greater (see Haberlie et al. 2015). One critical step was found to be that the memory space for the Java virtual machine had to be lowered in the wct-export.bat file. The script uses Python 3 and was run using Spyder, via Anaconda3.

All available timestamps for May through September, from 2001 through 2020, were downloaded as shapefiles. Each shapefile typically contained dozens of polygons, one for each gate that returned data with an intensity value of 40 dBZ or higher. Each polygon contained an intensity value, timestamp, radial angle, beginning and ending distance from the radar antenna, height above the ground, and height above the antenna.

These shapefiles were reprojected into NAD 1983 UTM Zone 13N (EPSG: 26913), clipped to the study area, and merged into monthly aggregates, which retained intensity values and timestamps for each polygon within the shapefiles. Centroid points were then

created for each polygon in the monthly shapefiles. The centroid point shapefiles were then used for further processing, as described in section 4.1. All GIS preprocessing was done in ArcMap 10.6 and ArcGIS Pro using Python scripts and the arcpy site package.

#### 3.2.1.1 NEXRAD Weather Radar Data Limitations and Considerations

As previously mentioned, weather radar is limited by line-of-sight, and as such cannot return data that are obscured by high terrain. There are other limitations to consider as well. By using only one tilt angle, only one "slice" of the atmosphere was sampled. This "slice" rises at a 0.5° angle, meaning that data collected closest to the radar station is 11m (approximately 36 feet) above the radar antenna, while the pulse of energy sampling the atmosphere over Santa Fe, 95km (59 miles) away, is 1,400m (4,593 feet) above the elevation of the radar antenna, and continues rising. Additionally, this "slice" expands as it moves away from the radar station, therefore sampling larger portions of the atmosphere as the pulse moves away from the antenna.

NEXRAD data is collected at regular intervals called gates, and along radials, which are lines extending out from the radar antenna at 0.003 degree intervals. Through the year 2007, the gates were 1,000m apart, and starting in 2008, they were 250m apart (Figure 8). So, while actual thunderstorm activity often has fuzzy boundaries, the data are collected at regularly spaced intervals. These locations are called bins, and they are "the most granular level of digital radar output" (Vasquez, 2015, p. 27).

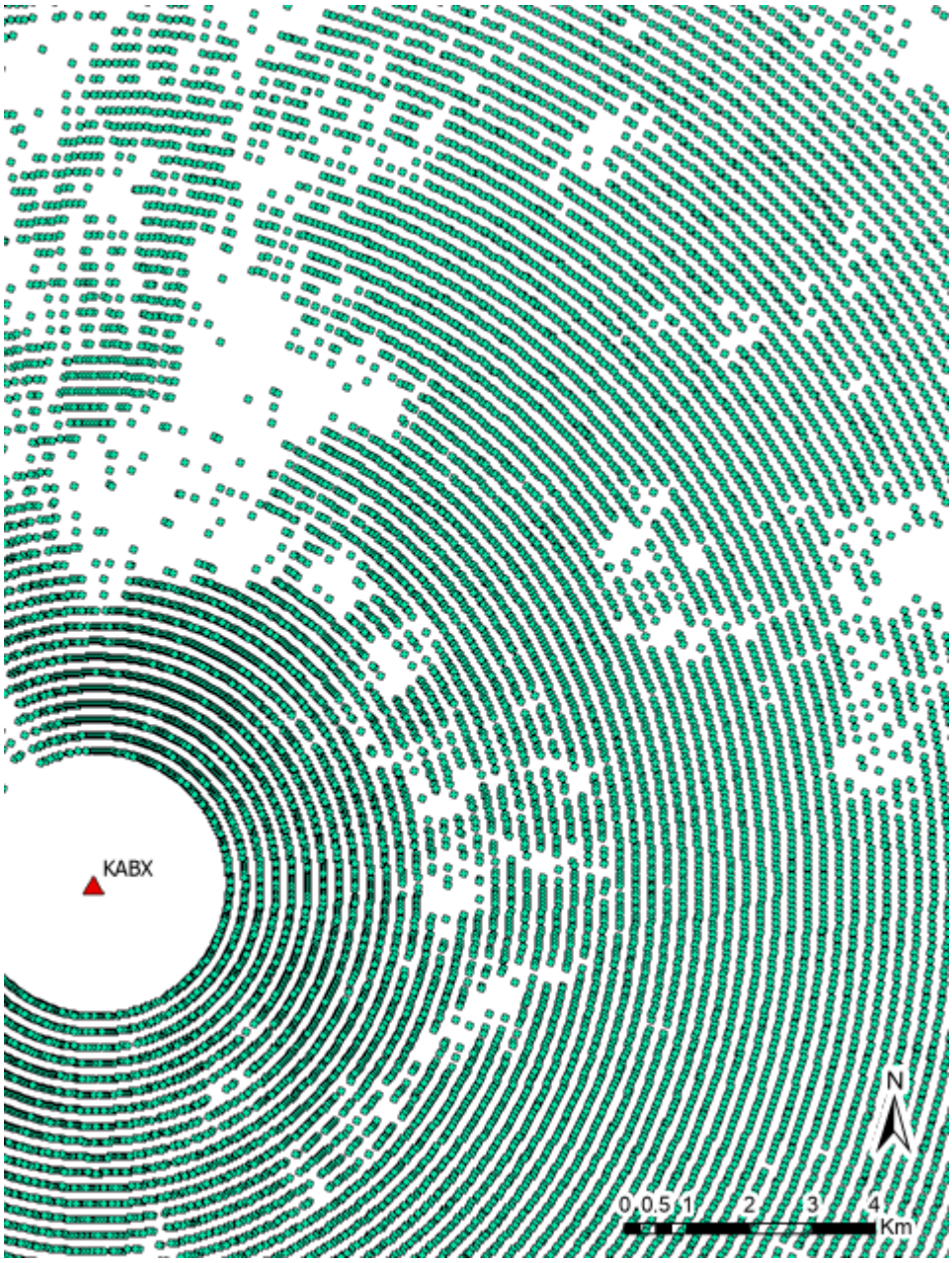



Figure 8. June 2018 points. Detailed look at regularly aligned points (polygon centroids) corresponding with gates and radials. Each gate is 250m apart and each radial is 0.003 degrees apart. Before 2008, the gates were 1,000m apart.

### 3.2.2 National Land Cover Database

The Multi-Resolution Land Characteristics (MRLC) Consortium releases the National Land Cover Database (NLCD) raster dataset every two to three years. NLCD land cover data are derived from Landsat Analysis Ready Data (ARD), as well as "ingesting partner data from 10 different Government partners and Federal agencies that are used to improve accuracy and ensure a cohesive product across these partners" (Dewitz, 2020). Land cover is classified using the Anderson Level II classification system, and the datasets have a spatial resolution of 30m. Descriptions of each class are given in Table 5. The most recent release, from 2019, includes raster datasets for the years 2001, 2004, 2006, 2008, 2011, 2013, 2016, and 2019.

Most changes from one raster dataset to the following one are small. However, land cover data for the years in between datasets cannot be interpolated because significant, localized changes can occur. For example, within a few days or months wildfires caused significant land cover changes across parts of the study area, primarily in the mountains north and southeast of Albuquerque.

As shown in Table 5, NLCD land cover classifications divide developed land cover into four classes based on the percent of impervious surface. The NLCD define "urban impervious surfaces as a percentage of developed surface over every 30-meter pixel" (MRLC, 2021). Additionally, Table 5 describes the most common types of built structures and surfaces found in each developed land cover class. In arid and semi-arid

climates, Developed Open Space land cover, which consists of less than 20% impervious surfaces, and is mainly comprised of irrigated vegetation, such as lawns, parks, and golf courses, may provide a cooling effect by adding more moisture to the atmosphere that would not have been present without human influence (Lazzarini et al., 2015). This moisture may also potentially contribute to the moisture needed to form clouds and thunderstorms.

Table 5. National Land Cover Database legend and description of classes.

| Class | Value | Classification Description |
|---|---|---|
| Water | | |
| | 11 | **Open Water** – areas of open water, generally with less than 25% cover of vegetation or soil. |
| Developed | | |
| | 21 | **Developed, Open Space** – areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20% of the total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes. |
| | 22 | **Developed, Low Intensity** – areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20% to 49% of total cover. These areas most commonly include single-family housing units. |
| | 23 | **Developed, Medium Intensity** – areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50% to 79% of total cover. These areas most commonly include single-family housing units. |
| | 24 | **Developed, High Intensity** – highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses, and commercial / industrial. Impervious surfaces account for 80% to 100% of the total cover. |
| Barren | | |
| | 31 | **Barren Land (Rock / Sand / Clay)** – areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits, and other accumulations of earthen material. Generally, vegetation accounts for less than 15% of total cover. |
| Forest | | |
| | 41 | **Deciduous Forest** – areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75% of the tree species shed foliage simultaneously in response to seasonal change. |
| | 42 | **Evergreen Forest** – areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75% of the tree species maintain their leaves all year. Canopy is never without green foliage. |
| | 43 | **Mixed Forest** – areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. Neither deciduous nor evergreen species are greater than 75% of total tree cover. |

Table 5 (cont.)

| Class | Value | Classification Description |
|---|---|---|
| Shrubland | | |
| | 52 | **Shrub / Scrub** – areas dominated by shrubs; less than 5 meters tall with shrub canopy typically greater than 20% of total vegetation. This class includes true shrubs, young trees in an early successional stage or trees stunted from environmental conditions. |
| Herbaceous | | |
| | 71 | **Grassland / Herbaceous** – areas dominated by gramanoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing. |
| Planted / Cultivated | | |
| | 81 | **Pasture / Hay** – areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture / hay vegetation accounts for greater than 20% of total vegetation. |
| | 82 | **Cultivated Crops** – areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20% of total vegetation. This class also includes all land being actively tilled. |
| Wetlands | | |
| | 90 | **Woody Wetlands** – areas where forest or shrubland vegetation accounts for greater than 20% of vegetative cover and the soil or substrate is periodically saturated with or covered with water. |
| | 95 | **Emergent Herbaceous Wetlands** – areas where perennial herbaceous vegetation accounts for greater than 80% of vegetative cover and the soil or substrate is periodically saturated with or covered with water. |

Source: https://www.mrlc.gov/data/legends/national-land-cover-database-class-legend-and-description

### 3.2.3 PRISM Temperature Data

Maximum, mean, and minimum temperature rasters with a spatial resolution of 800m were downloaded from PRISM (PRISM Climate Group, Oregon State University, https://prism.oregonstate.edu). The temperature rasters were resampled to a spatial resolution of 60m using the "nearest-neighbor" resampling technique. This is the same spatial resolution as the digital elevation model (DEM).

While near-surface temperature data does not give much indication of whether a conditionally unstable lapse rate is present, it does give some indication of the temperature within the lowest portion of the atmosphere. It was deemed impractical to include data showing the atmospheric lapse rate, as those data are collected at only one point within the study area (near the airport that serves the Albuquerque metropolitan area).

The minimum and maximum temperatures were selected to characterize diurnal temperature fluctuations. The minimum temperature was specifically selected because the urban heat island effect is often more pronounced at night (Debbage & Shepherd, 2015). A relatively high minimum temperature, especially in urban areas, would indicate that the land surface is not cooling as much as non-urban areas, likely due to the urban heat island effect.

### 3.2.4 Digital Elevation Model

A digital elevation model (DEM) with a 60m spatial resolution was acquired from RGIS (https://rgis.unm.edu/). The DEM was used to generate elevation, slope, and aspect data. These variables provide an indication of the surface terrain, which can influence the initiation and life cycle of thunderstorms. Airflow over uneven terrain can lift moisture to an altitude where it can condense and form clouds. Rising terrain can also provide the lift necessary to move an air parcel above the level of free convection (LFC), allowing for thunderstorm growth.

## 4.0 Sampling Design

Weather radar data was downloaded as polygons via the Weather and Climate Toolkit, clipped to the study area, and preprocessed into centroid points. Further processing was used to cluster the points by time and location to create thunderstorm event polygons. Underlying data were gathered from raster datasets and added to the event polygon attributes. See Table 6 for a table of the most relevant variables used in this study and see Table 14 in Appendix C for the complete list of variable names and descriptions derived from the radar data.

### 4.1 Derivation of Thunderstorm Events from Weather Radar Data

As previously mentioned, radar data for each timestamp were downloaded as polygons, and each polygon included a timestamp and intensity value. After the polygons were clipped to the study area, they were converted into centroid points which retained the same attributes. The points were merged into monthly shapefiles (Figure 9). Since all of the preprocessing was done in ArcGIS Desktop and ArcGIS Pro, the time field was converted to an ESRI date field. This key step allowed for much of the further processing to be accomplished. The monthly shapefiles were converted to feature classes in a geodatabase in ArcGIS Pro.

Using June 2018 as an example, Figure 9 depicts the centroid points within the study area. Months started at 12:00:00 UTC (6:00:00am local time) on the first day of the month and ended at 11:59:59 UTC (5:59:00am) on the first day of the following month,

except for May and September, which started at 00:00:00 UTC on the first day of the month and ended at 11:59:59 UTC on the last day of the month, respectively. 12:00:00 UTC was chosen as the cut-off time between months because it was expected that thunderstorm activity would be at a minimum at this time, thereby minimizing the number of ongoing thunderstorms that would be arbitrarily divided into two events, one in each month.

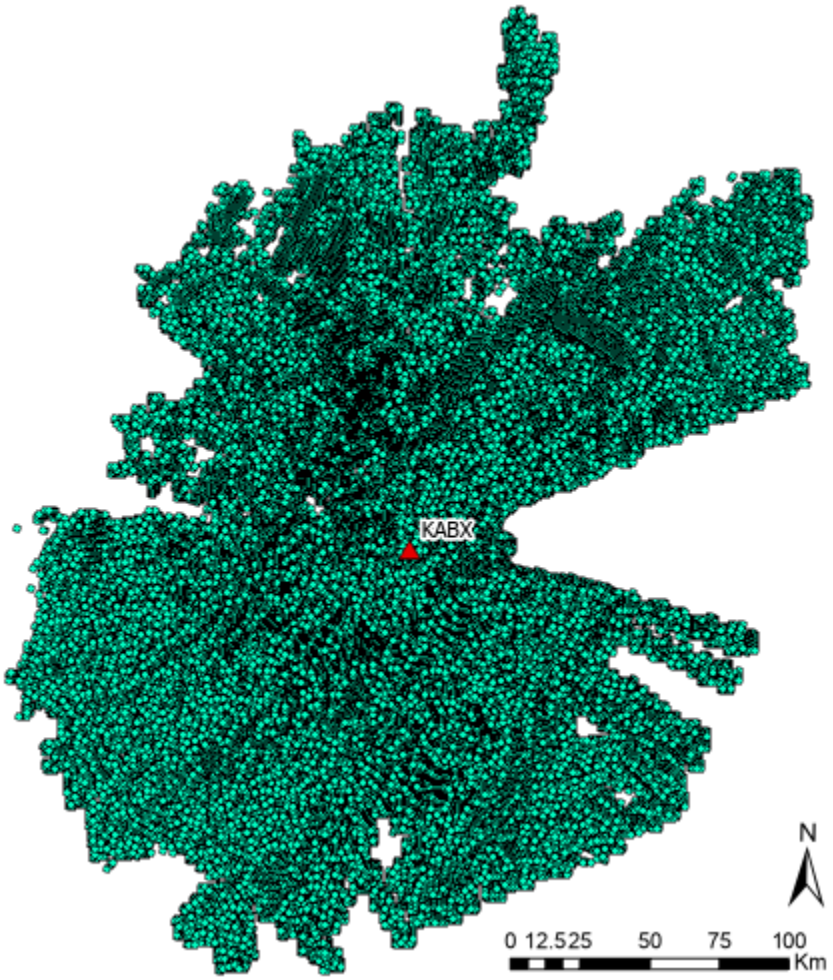



Figure 9. Centroid points derived from polygons for June 2018. There are 1,656,306 points in this month.

The *Find Point Clusters* tool in ArcGIS Pro was used to cluster monthly points by time and distance. The clustering algorithm used the Defined Distance (DBSCAN) method. Cluster parameters were set to contain a minimum of 20 points. Points were grouped into the same cluster if they were within a distance of 1700m and had a start time within 30 minutes of each other (see Haberlie et al., 2015).

The parameter specifying a minimum of 20 points was determined through an iterative process. Clusters with varying numbers of minimum points were compared with screenshots of actual radar weather as displayed on the website aviationweather.gov (Figure 10). The distance parameter of 1700m was selected because points farthest from the radar station, where points are the farthest apart, were found to be approximately 1600m apart, for most years, so a parameter of 1700m would allow points to be grouped together that were likely part of the same thunderstorm activity. The time parameter of 30 minutes was chosen because airmass thunderstorms in New Mexico are typically considered to last approximately 30 minutes or less (Sullivan et al., 2018), thus a parameter of at least 30 minutes would capture thunderstorm activity that signified a weather event that could present hazards to people on the ground.

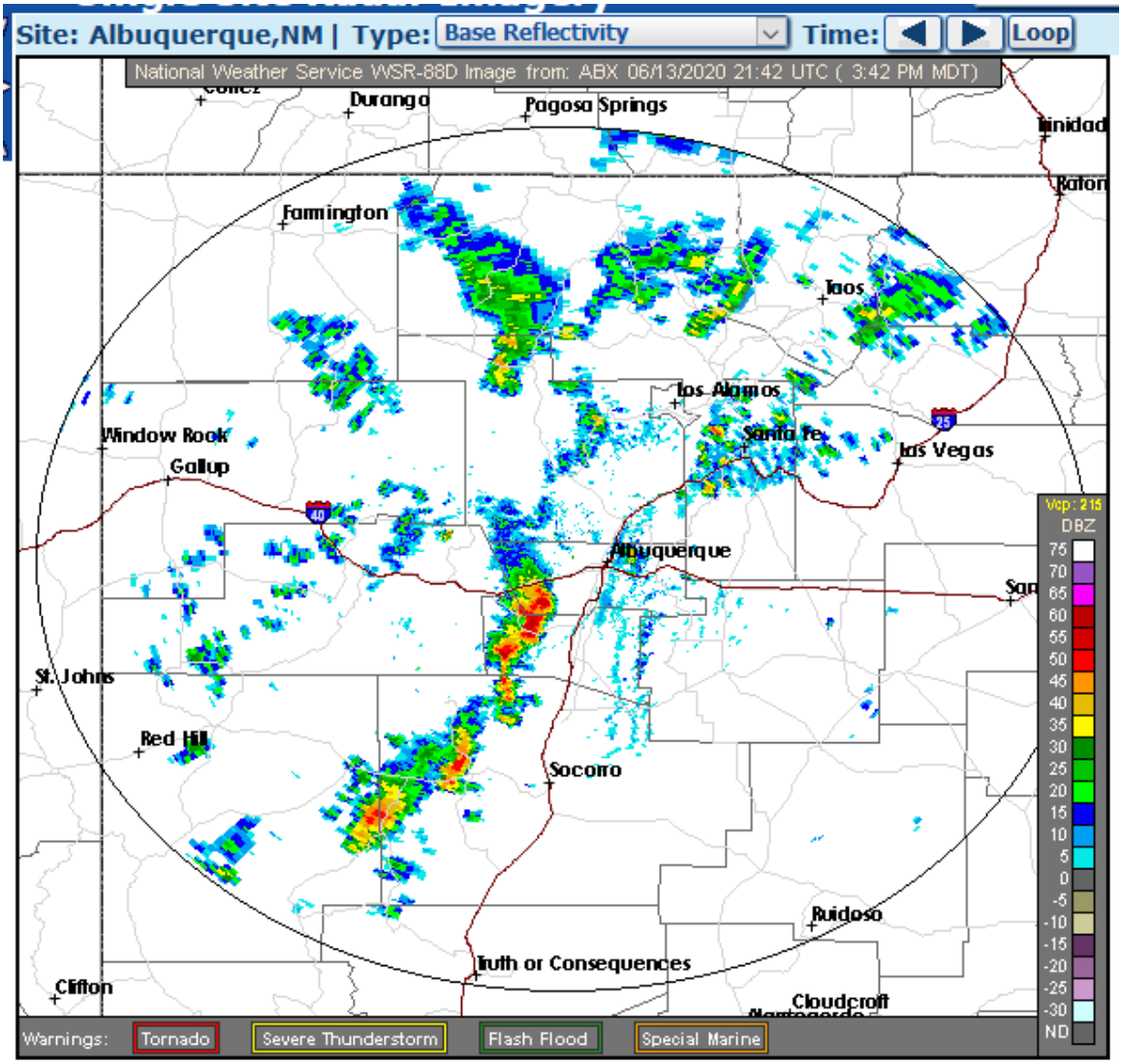



Figure 10. This screenshot from aviationweather.gov depicts the base reflective radar image from June 13, 2020, at 21:42 UTC. The intensity level of 40 dBZ of reflectivity is represented by dark yellow in this image.

Each cluster automatically generated a start date and start time, an end date and end time, and a unique Cluster ID, in addition to the previously mentioned attributes (Figure 11a). The start date and time are the same as the earliest timestamp in the cluster, and the end date and time correspond to the last timestamp in the cluster. Points that did not meet one or more of the criteria were not included in any further analysis.

The *Minimum Bounding Geometry* (MBG) tool was then used to create polygons enclosing each cluster, according to its unique Cluster ID. The polygons were created using convex hull geometry (Figure 11b). The MBG polygons were then clipped to the study area. Although each MBG polygon has a definite outline, the thunderstorm event it represents has fuzzy boundaries. The boundary of each MBG polygon should be considered an approximation of the outline of each event. These polygons are hereafter referred to as thunderstorm event polygons, and are the unit of analysis used in this study.

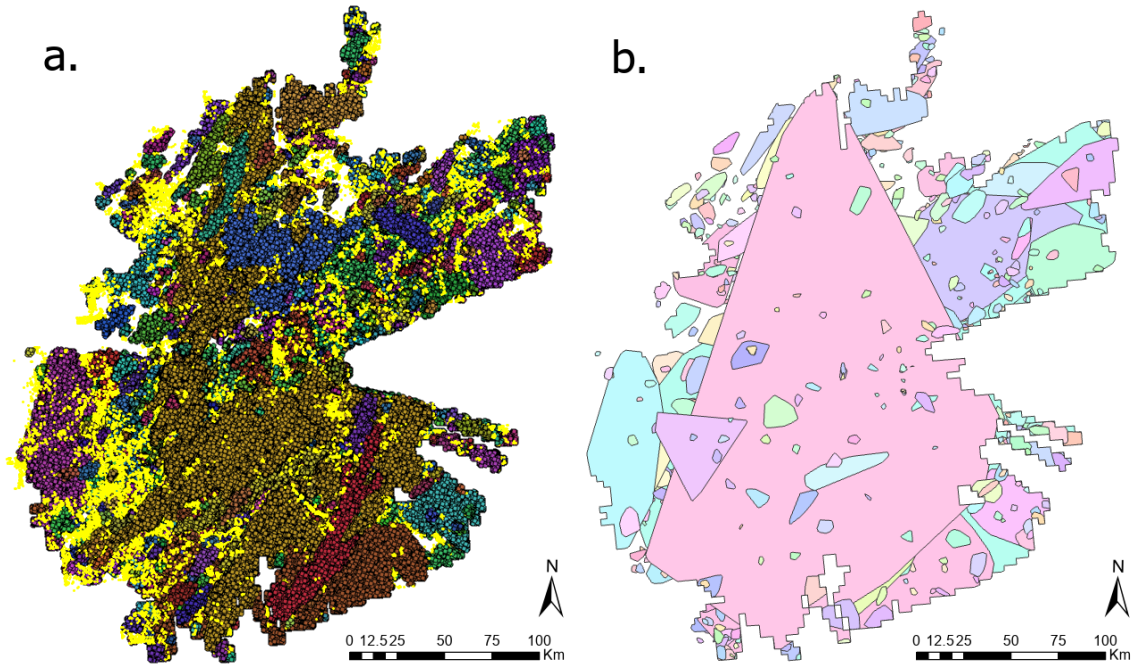



Figure 11. (a.) Points for June 2018, symbolized by Cluster ID. Each color represents a different cluster. There are 1,294 clusters in June 2018. The smaller, yellow points were not grouped into any cluster because they did not meet one or more of the parameters. These unassociated points were disregarded in any further analysis. Not all clusters are visible in the figure. (b.) MBG polygon feature class for June 2018. Not all clusters are visible in the figure.

A second set of polygons was created using the *Minimum Bounding Geometry* tool to group points by both Cluster ID and timestamp, so that each polygon included only one timestamp for each Cluster ID. The convex hull parameter was also used for this set of polygons, and centroid points were created for each polygon. Lines were created to connect the centroid points sequentially. One line was created for each Cluster ID and the vertices of the line were drawn in order by timestamp. This set of lines was used to derive several variables: line bearing in degrees, the distance between the start point and end point of the line, the length of the line, and the ratio of the length of the line to the distance from the start-point to the end-point (line length ratio). The line length ratio was created because it was found that while some thunderstorm events moved fairly linearly for long distances, others stayed in one area. It is beyond the scope of this study, but a

number of observations could be made comparing these variables to the relative size of each thunderstorm event, to terrain and other surface features, or other analyses. Also, some thunderstorms that had a long time duration but stayed in a small geographic area were identified as potentially erroneous storms in the dataset during the data cleaning process.

A number of fields were added to the monthly thunderstorm event polygon feature classes. Each point within a cluster had retained its original intensity value, so statistics on intensity values within each cluster could be derived. Summary tables in ArcGIS Pro were used to generate the maximum intensity and mean intensity values for each cluster from the individual point intensity values, and these values were added to the attribute table of each monthly polygon feature class. The start date and start time and the end date and end time were used to create a time duration field. A full list of the variables is in Table 14 in Appendix C, and Table 6, below, has a summary of the most relevant variables.

### 4.2 Addition of Land Cover, Temperature, and Elevation Attributes to Thunderstorm Events

Additional attributes were added to the thunderstorm event polygons from several raster sources: the NLCD datasets, PRISM temperature data, and the DEM. The *Tabulate Area* tool in ArcGIS Pro was used to add land cover data to the event polygons. The area of each event was used to derive the percent of each type of land cover. The NLCD dataset

is in square pixels, and the minimum bounding geometry polygons surrounding each event are irregular shapes, so in some cases, the percent area was more than 100%. This was addressed during the data cleaning process.

Data derived from the PRISM temperature rasters using the *Zonal Statistics as Table* tool in ArcGIS Pro were added to the thunderstorm event polygon attribute tables. It was found that in order for the radar polygons to sample the temperature data, the center point of at least one pixel must fall within a polygon. Due to the small size of some of the polygons, the temperature rasters were resampled to 60m. Sixty meters was chosen as the new spatial resolution so that it would be sufficiently small and the same as the spatial resolution of the digital elevation model (DEM). The *Zonal Statistics as Table* tool was also used to derive elevation, slope, and aspect variables from the DEM. See Table 6 for a table of the most relevant variables used in the study and see Table 14 in Appendix C for the complete list of variable names and descriptions.

Table 6. A portion of the variables created for this study. See Appendix C for the full table.

| Variable name | Description |
|---|---|
| *Cluster_Mean_Intensity_Value* | Mean intensity |
| *Cluster_Max_Intensity_Value* | Maximum intensity |
| *Time_Duration_minutes* | Duration, in minutes |
| *Percent_All_Developed* | Percentage of the sum of all four Developed areas (Open Space, Low Intensity, Medium Intensity, and High Intensity) |
| *Percent_Developed_High* | Percentage of Developed, High Intensity area |
| *Percent_Develop_Med_Low_OpnSp* | Percentage of the sum of Medium Intensity, Low Intensity, and Open Space Developed areas |
| *Aspect_median* | Median aspect value, in degrees |
| *Elevation_range_meters* | Range of elevation values, in meters |
| *Mean_MaxTemp* | Mean of the maximum temperature, in degrees Celsius |
| *Mean_MinTemp* | Mean of the minimum temperature, in degrees Celsius |
| *Percent_All_Forest* | Percentage of the sum of all three Forest areas (Deciduous, Evergreen, and Mixed) |
| *Percent_PastureHay* | Percentage of Pasture / Hay |
| *Percent_CultivatedCrops* | Percentage of Cultivated Crops |
| *Percent_WoodyWetlands* | Percentage of Woody Wetlands |
| *Percent_EmHerbWetlands* | Percentage of Emergent Herbaceous Wetlands |
| *Percent_OpenWater* | Percentage of Open Water area |
| *Percent_Barren* | Percentage of Barren Land (Rock / Sand / Clay) |

### 4.3 Data Cleaning

After all of the attributes were added to the monthly thunderstorm event polygons, the polygons were combined into yearly feature classes, then into one feature class containing all twenty years of thunderstorm events. Once the dataset had been consolidated, several iterations of data cleaning were applied in an effort to remove the effects of ground clutter and wind turbines as much as possible.

First, a shapefile was found that depicted the locations of wind turbines in New Mexico (https://eerscmap.usgs.gov/uswtdb/data/). The most obvious set of wind turbines in the study area is located on the northeast flank of Mount Taylor. A minimum bounding geometry polygon with convex hull geometry was created around those wind turbines, and thunderstorm polygons that were located completely within that polygon were removed from the dataset. Additionally, a circular polygon with a radius of 2 kilometers was created around the KABX antenna and any polygons that were completely within that circle were removed.

In the next iteration of cleaning, thunderstorm event polygons that met at least one of several criteria were removed. Based on a personal communication with the Albuquerque National Weather Service office, it was determined that events exceeding a time duration of greater than 600 minutes (10 hours) should be removed. Thunderstorm events with an area less than 100 square kilometers and time duration greater than 300 minutes (5 hours) or an area less than 10 square kilometers and time duration greater than 180 minutes (3

hours) were removed. These parameters were based on observations of actual thunderstorm events in central New Mexico.

It was found that a small number of event polygons had land cover percentages exceeding 200%. Those events were removed because it was deemed unreasonable to have percent land cover equal to twice the size of the polygon. This occurred due to the fact that the land cover raster contains square pixels, while the event polygons are not regular shapes and the *Tabulate Area* tool accounted for the entire pixel, rather than only the part that overlapped the polygon.

A number of thunderstorm event polygons that appear to be potentially erroneous remain in the dataset (Figure 12). These polygons are distinct because they are grouped together more so than most of the other polygons, and one or more or their vertices meet at the same locations, instead of being more evenly distributed. However, no objective method of confirming their validity was found. They are primarily located between the Albuquerque metropolitan area and the nearby Sandia Mountains. This is an area where it is reasonable to find numerous thunderstorm events, but it is also apparent that the vertices that are collocated are often found close to large roads in Albuquerque, where antennas and other ground clutter are likely to be found. It is possible that ground clutter returned erroneous data to KABX which the clutter filter did not remove.

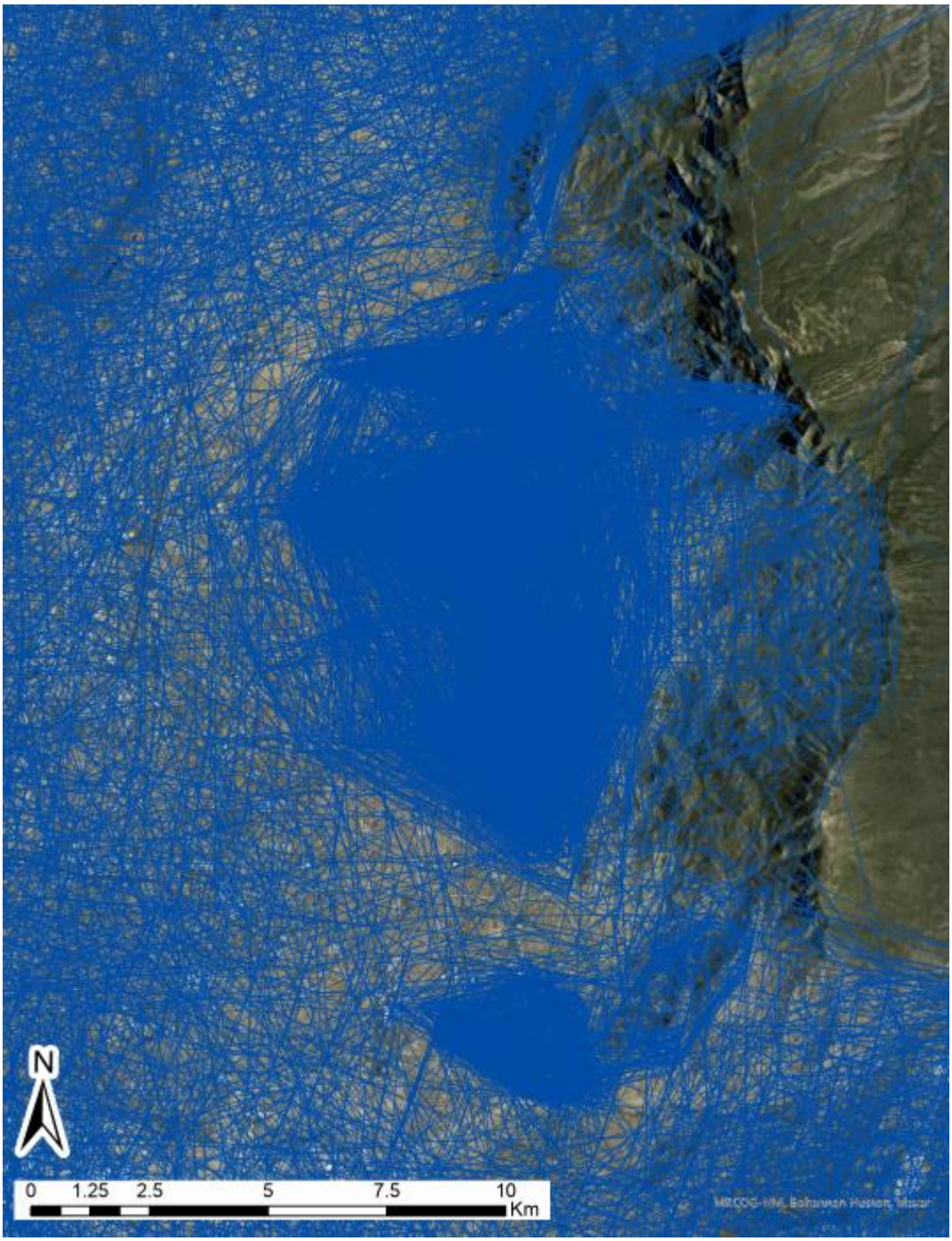



Figure 12. The outlines of polygons are shown in blue, and overlaid on a satellite base map. The base map shows the eastern portion of the Albuquerque metropolitan area, and the Sandia Mountains can be seen in a north-south alignment in the right half of the image. There are two groups of potentially erroneous polygons in this image: a large one in the center, and a smaller one south of it. The western vertices of the smaller polygon are concentrated near Interstate 25, where it is likely that a number of towers, antennas, and other ground clutter are present.

### 4.4 Analysis

A smaller subset of the variables in the MBG polygon attribute tables was needed for the analysis to avoid redundancy in the regression models. A large number of variables were created for this study, as seen in Table 14 in Appendix C, and many of them are likely correlated with each other. It is expected that changes in elevation will produce similar changes in temperature and slope, and potentially corresponding changes in vegetation as well. A subset of variables was selected for their relevance to thunderstorm formation, based on the three "ingredients" needed to form and sustain thunderstorm activity: a lifting action, instability, and moisture (see Table 7), and tested for correlation.

Table 7. Variables most relevant to thunderstorm formation.

| Variable name and thunderstorm "ingredient" | Relevance to thunderstorm activity |
|---|---|
| **Lifting action** | |
| *Aspect_median* | Interaction between predominant wind direction and rising terrain; heating from the sun in the afternoon; uneven heating and presence or lack of moisture as indicated by dry side vs wet side of mountains |
| *Slope_range* | Indicative of rising terrain for wind to flow over |
| *Elevation_range_meters* | High terrain may indicate rising slopes for wind to flow over |
| *Range_MaxTemp* | Variations in temperature imply uneven heating of the surface |
| **Instability** | |
| *Mean_MaxTemp* | How hot is the lowest portion of the atmosphere? Higher temperatures near the surface can indicate less stable atmosphere |
| *Mean_MinTemp* | Is it cooling off at night? Especially over urban areas? |
| **Moisture** | |
| *Percent_All_Forest* | May indicate presence of moisture needed to maintain forest biome; evapotranspiration |
| *Percent_PastureHay, Percent_CultivatedCrops* | Irrigated crops may indicate more moisture available for evaporation, evapotranspiration |
| *Percent_WoodyWetlands, Percent_EmHerbWetlands* | Areas of vegetation where the "soil or substrate is periodically saturated with or covered with water" (NLCD) |
| *Percent_OpenWater* | Can add moisture to the air through evaporation |
| *Percent_Barren* | May indicate a lack of moisture on the surface and lowest portion of the atmosphere |

Correlation matrices were run in RStudio in order to further narrow down the list of variables to those that were not strongly correlated with each other (see Appendix D). A threshold of $r < 0.6$ was used to exclude correlated variables. It was found that *Elevation_range_meters* was strongly correlated with *Slope_range* ($r = 0.81$) and *Range_MaxTemp* ($r = 0.88$). *Mean_MaxTemp* and *Mean_MinTemp* were also strongly correlated ($r = 0.87$), as expected. *Slope_range* and *Range_MaxTemp* were also moderately correlated ($r = 0.65$) with each other, and so *Slope_range* and *Range_MaxTemp* were excluded from analysis. *Elevation_range_meters*, along with the other variables that were not strongly correlated, were retained for regression analysis (Table 8).

*Mean_MaxTemp* and *Mean_MinTemp* were both kept for analysis, but run separately (i.e., *Mean_MaxTemp* would be run with all of the other variables, then *Mean_MinTemp* would be substituted for *Mean_MaxTemp* and run with all of the other variables). As previously mentioned, the minimum temperature can provide a useful metric for measuring the influence of surface features on atmospheric temperature, particularly in urban climate regimes where temperatures may stay elevated after sunset.

The final set of selected variables (Table 8), along with variables representing developed land cover, constituted the independent variables in multi-variate regression to determine the effect of developed land cover on thunderstorm intensity and duration. Separate

correlation matrices were created for each set of the independent variables to be used in each model (see Appendix D).

Table 8. Variables that were not strongly correlated to each other, and relevant to thunderstorm activity, along with developed land cover variables, were used as the independent variables in regression analysis.

| **Independent variables** | | |
|---|---|---|
| **Developed land cover variables** | | **Supporting independent variables** |
| *Percent_All_Developed* | | *Aspect_median* |
| *Percent_Developed_High* | | *Elevation_range_meters* |
| *Percent_Develop_Med_Low_OpnSp* | | *Mean_MaxTemp* (alternatively, *Mean_MinTemp*) |
| | | *Percent_All_Forest* |
| | | *Percent_PastureHay* |
| | | *Percent_CultivatedCrops* |
| | | *Percent_WoodyWetlands* |
| | | *Percent_EmHerbWetlands* |
| | | *Percent_OpenWater* |
| | | *Percent_Barren* |

### 4.5 Modeling

Simple and multivariate Ordinary Least Squares (OLS) regression models were used to determine changes in thunderstorm activity and what effect developed land cover had on thunderstorm intensity and duration. The models were run using the linear model (lm()) function in RStudio.

OLS assumes that data exhibits complete spatial randomness (CSR). CSR means that a phenomenon has an equal probability of existing in any location and that the location of any phenomenon is independent of the location of any other occurrence (O'Sullivan &

Unwin, 2010). The locations of thunderstorms are influenced by their environment, such as the underlying terrain and atmospheric temperature profile (Bowen, 1996; Wallace & Hobbs, 2006), and so it is unlikely that thunderstorms have an equal probability of forming and existing in any location. Also, the outflow from one thunderstorm may help perpetuate or regenerate another thunderstorm (Wallace & Hobbs, 2006) and thus the frequency or intensity of one thunderstorm may affect nearby occurrences. Based on these concepts, it is likely that the dataset does not exhibit CSR, and is spatially autocorrelated. Spatial autocorrelation is the concept that "spatial data from near locations are more likely to be similar than data from distant locations" (O'Sullivan & Unwin, 2010, page 199). Spatial autocorrelation of the dependent variables (*Cluster_Max_Intensity_Value*, *Cluster_Mean_Intensity_Value*, *Time_Duration_minutes*) was tested using the Global Moran's *I* in ArcGIS Pro and all three variables were found to be statistically significant, meaning that they are spatially autocorrelated (Table 9).

Table 9. The three dependent variables were found to be statistically significant.

| Variable | Results |
|---|---|
| *Cluster_Max_Intensity_Value* | p = 0 |
| | z-score = 245 |
| *Cluster_Mean_Intensity_Value* | p = 0 |
| | z-score = 548 |
| *Time_Duration_minutes* | p = 0 |
| | z-score = 402 |

However, in addition to the dataset containing polygons with spatial relationships to each other that may violate CSR, the polygons also exist over a long span of time. The

polygons relate to neighboring polygons and underlying processes in time as well as space, and so it may be assumed that the data also has temporal autocorrelation. Because the data may be affected by both spatial and temporal autocorrelation, one of two approaches is required for analysis: either run the dataset in a spatial-temporal model or use a simpler, non-spatial model such as OLS. As such, OLS was used for both single and multivariate models (Table 10). Debbage and Shepherd (2015) used OLS in their study of spatial contiguity and temperature intensity of the fifty most populous cities in the United States. They acknowledge that the assumptions of OLS can be difficult to meet when researching spatial phenomena, and in their research omitted outliers as a corrective action (Debbage & Shepherd, 2015, p. 185).

Thunderstorm maximum intensity, mean intensity, and duration were tested in the following models (Table 10). Also, a subset of thunderstorms with a maximum intensity in the 95$^{th}$ percentile were tested. This subset allowed for an analysis of the most intense thunderstorms that occurred within the study area. While the full dataset had 220,304 thunderstorm event polygons, the subset consisting of thunderstorms with a maximum intensity in the 95$^{th}$ percentile had 11,704 events. Models 2, 8, 9, and 10 used this smaller subset to test for changes over time and the relationship between land cover and thunderstorm maximum intensity for the most intense thunderstorms.

Table 10. Variables and number of thunderstorm event polygons for each model used.

| Model | Thunderstorm Events (MBG polygons) | Dependent variable | Independent variable(s) |
|---|---|---|---|
| Control set consists of the following variables: *Aspect_median*, *Elevation_range_meters*, *Mean_MaxTemp*, *Percent_All_Forest*, *Percent_PastureHay*, *Percent_CultivatedCrops*, *Percent_WoodyWetlands*, *Percent_EmHerbWetlands*, *Percent_OpenWater*, *Percent_Barren* | | | |
| 1 | 220,304 | *Cluster_Max_Intensity_Value* | *START_DATE* |
| 2 | 11,704 | *Cluster_Max_Intensity_Value* | *START_DATE* |
| 3 | 220,304 | *Cluster_Mean_Intensity_Value* | *START_DATE* |
| 4 | 220,304 | *Time_Duration_minutes* | *START_DATE* |
| 5 | 220,304 | *Cluster_Max_Intensity_Value* | *Percent_All_Developed*, Control set |
| 6 | 220,304 | *Cluster_Max_Intensity_Value* | *Percent_Developed_High*, Control set |
| 7 | 220,304 | *Cluster_Max_Intensity_Value* | *Percent_Develop_Med_Low_OpnSp*, Control set |
| 8 | 11,704 | *Cluster_Max_Intensity_Value* | *Percent_All_Developed*, Control set |
| 9 | 11,704 | *Cluster_Max_Intensity_Value* | *Percent_Developed_High*, Control set |
| 10 | 11,704 | *Cluster_Max_Intensity_Value* | *Percent_Develop_Med_Low_OpnSp*, Control set |
| 11 | 220,304 | *Cluster_Mean_Intensity_Value* | *Percent_All_Developed*, Control set |
| 12 | 220,304 | *Cluster_Mean_Intensity_Value* | *Percent_Developed_High*, Control set |
| 13 | 220,304 | *Cluster_Mean_Intensity_Value* | *Percent_Develop_Med_Low_OpnSp*, Control set |

Table 10 (cont.)

| Model | Thunderstorm Events (MBG polygons) | Dependent variable | Independent variable(s) |
|---|---|---|---|
| 14 | 220,304 | *Time_Duration_minutes* | *Percent_All_Developed*, Control set |
| 15 | 220,304 | *Time_Duration_minutes* | *Percent_Developed_High*, Control set |
| 16 | 220,304 | *Time_Duration_minutes* | *Percent_Develop_Med_Low_OpnSp*, Control set |

## 5.0 Model Interpretation

This study made several hypotheses about the relationship between thunderstorm activity and the surface environment in central New Mexico. The first four hypotheses relate to changes in thunderstorm intensity and duration over time. The subsequent hypotheses relate to the effects of developed land cover on thunderstorm characteristics (Table 11). The hypothesis number in Table 11 corresponds to the model number in Table 10. Hypotheses 2, 8, 9, and 10 used a smaller subset of the data, selecting only those polygons that had a maximum intensity value that was in the $95^{th}$ percentile.

Table 11. Null hypothesis and test for each hypothesis. The Model number corresponds with the Hypothesis number.

| Hypothesis number | Hypothesis | Null | Test |
|---|---|---|---|
| Change over time | | | |
| 1 | *Cluster_Max_Intensity_Value* has increased between 2001 and 2020 | *Cluster_Max_Intensity_Value* has not increased between 2001 and 2020 | Positive slope at $p \leq 0.05$ |
| 2 | *Cluster_Max_Intensity_Value* ($95^{th}$ percentile subset) has increased between 2001 and 2020 | *Cluster_Max_Intensity_Value* ($95^{th}$ percentile subset) has not increased between 2001 and 2020 | Positive slope at $p \leq 0.05$ |
| 3 | *Cluster_Mean_Intensity_Value* has increased between 2001 and 2020 | *Cluster_Mean_Intensity_Value* has not increased between 2001 and 2020 | Positive slope at $p \leq 0.05$ |
| 4 | *Time_Duration_minutes* has increased between 2001 and 2020 | *Time_Duration_minutes* has not increased between 2001 and 2020 | Positive slope at $p \leq 0.05$ |
| Maximum intensity | | | |
| 5 | *Percent_All_Developed* within each thunderstorm event is positively correlated with *Cluster_Max_Intensity_Value* | *Percent_All_Developed* within each thunderstorm event is not statistically significant or positively correlated with *Cluster_Max_Intensity_Value* | *Percent_All_Developed* has a positive slope at 95% |
| 6 | *Percent_Developed_High* within each thunderstorm event is positively correlated with *Cluster_Max_Intensity_Value* | *Percent_Developed_High* within each thunderstorm event is not statistically significant or positively correlated with *Cluster_Max_Intensity_Value* | *Percent_Developed_High* has a positive slope at 95% |
| 7 | *Percent_Develop_Med_Low_OpnSp* within each thunderstorm event is positively correlated with *Cluster_Max_Intensity_Value* | *Percent_Develop_Med_Low_OpnSp* within each thunderstorm event is not statistically significant or positively correlated with *Cluster_Max_Intensity_Value* | *Percent_Develop_Med_Low_OpnSp* has a positive slope at 95% |

Table 11 (cont.)

| Hypothesis number | Hypothesis | Null | Test |
|---|---|---|---|
| Maximum intensity (95th percentile subset) | | | |
| 8 | *Percent_All_Developed* within each thunderstorm event is positively correlated with *Cluster_Max_Intensity_Value* (95th percentile subset) | *Percent_All_Developed* within each thunderstorm event is not statistically significant or positively correlated with *Cluster_Max_Intensity_Value* (95th percentile subset) | *Percent_All_Developed* has a positive slope at 95% |
| 9 | *Percent_Developed_High* within each thunderstorm event is positively correlated with *Cluster_Max_Intensity_Value* (95th percentile subset) | *Percent_Developed_High* within each thunderstorm event is not statistically significant or positively correlated with *Cluster_Max_Intensity_Value* (95th percentile subset) | *Percent_Developed_High* has a positive slope at 95% |
| 10 | *Percent_Develop_Med_Low_OpnSp* within each thunderstorm event is positively correlated with *Cluster_Max_Intensity_Value* (95th percentile subset) | *Percent_Develop_Med_Low_OpnSp* within each thunderstorm event is not statistically significant or positively correlated with *Cluster_Max_Intensity_Value* (95th percentile subset) | *Percent_Develop_Med_Low_OpnSp* has a positive slope at 95% |
| Mean intensity | | | |
| 11 | *Percent_All_Developed* within each thunderstorm event is positively correlated with *Cluster_Mean_Intensity_Value* | *Percent_All_Developed* within each thunderstorm event is not statistically significant or positively correlated with *Cluster_Mean_Intensity_Value* | *Percent_All_Developed* has a positive slope at 95% |
| 12 | *Percent_Developed_High* within each thunderstorm event is positively correlated with *Cluster_Mean_Intensity_Value* | *Percent_Developed_High* within each thunderstorm event is not statistically significant or positively correlated with *Cluster_Mean_Intensity_Value* | *Percent_Developed_High* has a positive slope at 95% |

Table 11 (cont.)

| Hypothesis number | Hypothesis | Null | Test |
|---|---|---|---|
| Mean intensity (cont.) | | | |
| 13 | *Percent_Develop_Med_Low_OpnSp* within each thunderstorm event is positively correlated with *Cluster_Mean_Intensity_Value* | *Percent_Develop_Med_Low_OpnSp* within each thunderstorm event is not statistically significant or positively correlated with *Cluster_Mean_Intensity_Value* | *Percent_Develop_Med_Low_OpnSp* has a positive slope at 95% |
| Duration | | | |
| 14 | *Percent_All_Developed* within each thunderstorm event is positively correlated with *Time_Duration_minutes* | *Percent_All_Developed* within each thunderstorm event is not statistically significant or positively correlated with *Time_Duration_minutes* | *Percent_All_Developed* has a positive slope at 95% |
| 15 | *Percent_Developed_High* within each thunderstorm event is positively correlated with *Time_Duration_minutes* | *Percent_Developed_High* within each thunderstorm event is not statistically significant or positively correlated with *Time_Duration_minutes* | *Percent_Developed_High* has a positive slope at 95% |
| 16 | *Percent_Develop_Med_Low_OpnSp* within each thunderstorm event is positively correlated with *Time_Duration_minutes* | *Percent_Develop_Med_Low_OpnSp* within each thunderstorm event is not statistically significant or positively correlated with *Time_Duration_minutes* | *Percent_Develop_Med_Low_OpnSp* has a positive slope at 95% |

## 6.0 Results

Single and multivariate OLS regression models were used to determine changes in thunderstorm activity and if those changes can be explained by changes in urban land cover. Changes in the maximum intensity, mean intensity, and duration in the study area are summarized in Table 12. The effects of developed land cover variables on thunderstorm characteristics are given in Table 13.

### 6.1 Changes in Thunderstorm Activity

Models 1, 2, 3, and 4 sought to establish general characteristics about thunderstorm activity in the study area. Thunderstorm maximum intensity (H1) did not increase in the study area between 2001 and 2020 ($p < 2.2e\text{-}16$, at a 95% confidence level), and so the null hypothesis could not be rejected. The model adjusted $R^2$ ($R^2 = 0.01074$) showed that time explains only 1% of the change in maximum intensity. Thunderstorm maximum intensity in the $95^{th}$ percentile (H2) showed a slight increase over time at a 95% confidence level (p-value = 0.00454), so the null hypothesis was rejected. Model 2 also had a low model adjusted $R^2$ ($R^2 = 0.0006027$). This model used only 11,704 thunderstorm events, while Models 1, 3, and 4 used the full dataset of 220,304 events.

Thunderstorm mean intensity (H3) did not increase in the study area between 2001 and 2020 ($p < 2.2e\text{-}16$), and so the null hypothesis could not be rejected. The model adjusted $R^2$ ($R^2 = 0.06114$) showed little relationship between the mean intensity and time, at least in a linear model. As seen in Appendix F, a non-linear model may have been better suited

to the fluctuating nature of the data. Thunderstorm duration (H4) also did not increase in the study area between 2001 and 2020 ($p < 2.2e-16$), so the null hypothesis could not be rejected. The model adjusted $R^2$ ($R^2 = 0.01958$) was similar to Model 1.

Table 12. Summary of results for Models 1, 2, 3, and 4.

| Hypothesis | Model adjusted R-squared | Model p-value | Slope | Null Rejected |
|---|---|---|---|---|
| 1 | 0.01074 | <2.2e-16 | -3.640e-09 | No |
| 2 | 0.0006027 | 0.00454 | 0.016235 | Yes |
| 3 | 0.06114 | <2.2e-16 | -2.970e-09 | No |
| 4 | 0.01958 | <2.2e-16 | -4.145e-08 | No |

### 6.2 The Impact of Development on Thunderstorm Maximum Intensity

Models 5, 6, and 7 queried the relationship between developed land cover and thunderstorm maximum intensity. All three models were significant at a 95% confidence level, and all had a model adjusted $R^2$ of ~0.2. The percent of all classes of developed land cover (*Percent_All_Developed*) within each thunderstorm event polygon (H5) was found to be positively correlated with thunderstorm maximum intensity (*Cluster_Max_Intensity_Value*) and significant at a 95% confidence level ($p < 2e-16$), so the null hypothesis was rejected. The model fit ($R^2 = 0.2007$) showed that the independent variables selected for this model explained 20% of the variance in the maximum intensity of thunderstorms in the study area between 2001 and 2020. Additionally, the percent high developed land cover (*Percent_Developed_High*) (H6) was found to be positively correlated with maximum intensity (*Cluster_Max_Intensity_Value*) ($p < 2e-16$), so the null hypothesis was rejected. Further,

the percent of medium, low, and open space developed land cover (*Percent_Develop_Med_Low_OpnSp*) (H7) was found to be positively correlated with maximum intensity (*Cluster_Max_Intensity_Value*) ($p < 2e\text{-}16$), so the null hypothesis was rejected.

### 6.3 The Impact of Development on Thunderstorm Maximum Intensity in the 95th Percentile

Models 8, 9, and 10 looked at the relationship between developed land cover and the subset of thunderstorms with a maximum intensity in the 95th percentile. All three models were significant at a 95% confidence level, but all had a model adjusted $R^2$ of ~0.03, showing that only 3% of the variance in maximum intensity values in the 95th percentile could be explained by these models. The percentage of developed land cover (*Percent_All_Developed*) (H8) was found to be positively correlated with maximum intensity values in the 95th percentile ($p = 7.17e\text{-}07$), so the null hypothesis was rejected. The percent high developed land cover (*Percent_Developed_High*) (H9) was found to be positively correlated with maximum intensity values in the 95th percentile ($p = 0.003904$), so the null hypothesis was rejected. The percent of medium, low, and open space developed land cover (*Percent_Develop_Med_Low_OpnSp*) (H10) was found to be positively correlated with maximum intensity values in the 95th percentile ($p = 4.61e\text{-}07$), so the null hypothesis was rejected.

### 6.4 The Impact of Development on Thunderstorm Mean Intensity

Models 11, 12, and 13 tested the relationship between developed land cover and thunderstorm mean intensity. All three models were significant at a 95% confidence level, but all had a model $R^2$ of ~0.03, showing that only 3% of the variance in mean intensity could be explained by these models. The percent of developed land cover (*Percent_All_Developed*) within each thunderstorm event polygon (H11) was found to be positively correlated with thunderstorm mean intensity and significant at a 95% confidence level ($p < 2e\text{-}16$), so the null hypothesis was rejected. The percent high developed land cover (*Percent_Developed_High*) (H12) was also found to be positively correlated with mean intensity ($p < 2e\text{-}16$) and the null hypothesis was rejected. The percent of medium, low, and open space developed land cover (*Percent_Develop_Med_Low_OpnSp*) (H13) was found to be positively correlated with mean intensity ($p < 2e\text{-}16$), so the null hypothesis was rejected.

### 6.5 The Impact of Development on Thunderstorm Duration

Models 14, 15, and 16 queried the relationship between developed land cover and thunderstorm duration. The three models had model adjusted $R^2$ values of 22% to 25%, showing that 25% of the variance in duration can be explained by the independent variables used in this model. The percent developed land cover (*Percent_All_Developed*) (H14) was found to be positively correlated with thunderstorm duration ($p < 2e\text{-}16$), so the null hypothesis was rejected. The percent high developed land cover (*Percent_Developed_High*) (H15) was found to be positively correlated with duration ($p < 2e\text{-}16$), so the null hypothesis was rejected. The percent of medium, low, and open

space developed land cover (*Percent_Develop_Med_Low_OpnSp*) (H16) was found to be positively correlated with duration ($p < 2e\text{-}16$), so the null hypothesis was rejected.

Table 13. Summary of results for Models 5 through 16.

| Hypothesis | Model adjusted R-squared | Model p-value | Variable | Variable slope | Variable p-value | Null Rejected |
|---|---|---|---|---|---|---|
| Maximum intensity | | | | | | |
| 5 | 0.2007 | < 2.2e-16 | Percent_All_Developed | 3.792e-02 | < 2e-16 | Yes |
| 6 | 0.1937 | < 2.2e-16 | Percent_Developed_High | 2.149e-01 | < 2e-16 | Yes |
| 7 | 0.2009 | < 2.2e-16 | Percent_Develop_Med_Low_OpnSp | 4.151e-02 | < 2e-16 | Yes |
| Maximum intensity (95$^{th}$ percentile subset) | | | | | | |
| 8 | 0.03356 | < 2.2e-16 | Percent_All_Developed | 9.493e-03 | 7.17e-07 | Yes |
| 9 | 0.03221 | < 2.2e-16 | Percent_Developed_High | 5.329e-02 | 0.003904 | Yes |
| 10 | 0.03363 | < 2.2e-16 | Percent_Develop_Med_Low_OpnSp | 1.045e-02 | 4.61e-07 | Yes |
| Mean intensity | | | | | | |
| 11 | 0.03217 | < 2.2e-16 | Percent_All_Developed | 1.035e-02 | < 2e-16 | Yes |
| 12 | 0.02742 | < 2.2e-16 | Percent_Developed_High | 5.482e-02 | < 2e-16 | Yes |
| 13 | 0.03239 | < 2.2e-16 | Percent_Develop_Med_Low_OpnSp | 1.138e-02 | < 2e-16 | Yes |
| Duration | | | | | | |
| 14 | 0.2488 | < 2.2e-16 | Percent_All_Developed | 0.6504379 | < 2e-16 | Yes |
| 15 | 0.2151 | < 2.2e-16 | Percent_Developed_High | 3.1915033 | < 2e-16 | Yes |
| 16 | 0.2508 | < 2.2e-16 | Percent_Develop_Med_Low_OpnSp | 0.7189678 | < 2e-16 | Yes |

## 7.0 Discussion

It was found that thunderstorm maximum intensity, mean intensity, and duration did not increase overall between 2001 and 2020. However, the maximum intensity values of thunderstorms with a maximum intensity in the 95$^{th}$ percentile did increase during that time, showing that the intensity of the most intense thunderstorms increased between 2001 and 2020. As seen in Appendix F, the mean of the maximum intensity values, mean intensity values, and duration fluctuated throughout the study period, but did not result in a net increase.

It was found that the percent of developed land cover within a thunderstorm event polygon was statistically significant and positively correlated with thunderstorm maximum intensity, meaning that an increase in percent developed land cover resulted in an increase in maximum intensity. All three levels of percent developed land cover that were tested resulted in similar positive correlations: percent developed land cover (*Percent_All_Developed*), percent high developed land cover (*Percent_Developed_High*), and percent of medium, low, and open space developed land cover (*Percent_Develop_Med_Low_OpnSp*). Percent high developed land cover (*Percent_Developed_High*) resulted in the highest slope of these three models (slope = 2.149e-01), even though it is only 0.12% of the land cover in the study area and 5% of developed land cover in the study area (as of 2019). Percent high developed land cover (*Percent_Developed_High*) can be expected to have the greatest influence on local atmospheric conditions because it is likely to include surfaces and activities that will contribute to the urban heat island effect (“due to the urban canyon geometry altering

heat storage release" (Debbage & Shepherd, 2015, p. 182)), surface roughness, and the release of air pollution and moisture associated with industrial areas and high density housing (Oke et al., 2017). Thunderstorms that occur near larger and rapidly expanding cities with a greater urban footprint can also be expected to exhibit increasing maximum intensity values and longer duration times.

Most of the urban development in the study area consists of medium, low, and open space developed land cover, which consist of a mixture of impervious surfaces and vegetation. Impervious surfaces account for 20% to 80% of the areas covered in those three land cover classes. The vegetation in these areas can be irrigated surfaces such as parks, golf courses, and lawns (Table 5). Although these three classes have fewer impervious surfaces and less intense industrial activities, they comprise 2.29% of the study area and 95% of developed land cover in the study area, as of 2019. As proposed by Lazzarini et al. (2015), the vegetation introduced in areas of medium, low, and open space developed land cover, such as lawns and golf courses, may introduce moisture to the atmosphere that would not normally be present in a semi-arid climate.

Thunderstorm maximum intensity in the $95^{th}$ percentile, thunderstorm mean intensity, and thunderstorm duration were also statistically significant and positively correlated with the percent developed land cover (*Percent_All_Developed*), percent high developed land cover (*Percent_Developed_High*), and percent of medium, low, and open space developed land cover (*Percent_Develop_Med_Low_OpnSp*). However, the model

adjusted $R^2$ values for maximum intensity in the 95$^{th}$ percentile ($R^2$ = 0.03) and mean intensity ($R^2$ = 0.03) were much lower than those of maximum intensity ($R^2$ = 0.19 to $R^2$ = 0.2) and duration ($R^2$ = 0.22 to $R^2$ = 0.25). This suggests that the selected variables were more explanatory of the peak intensity and longevity of all thunderstorms in the study. Models 8, 9, and 10, which used a subset of the data with maximum intensity values in the 95$^{th}$ percentile as the dependent variable, may have had a smaller model adjusted $R^2$ value because they used a smaller n. Detailed results for Models 8, 9, and 10 (see Appendix G) show that typically only about half of the independent variables used in the model were statistically significant, which is fewer than in the other models. As seen in the box plots in Appendix F, the distribution of values for maximum intensity and duration are fairly even across the 20-year study period. However, mean intensity values contain numerous outliers in the year 2009, which could have adversely affected the model adjusted $R^2$ values. When Debbage and Shepherd (2015) used OLS in their study, they recognized and removed outliers before conducting their analysis. However, for this study, further investigation should be conducted into the distribution of mean intensity values before considering removing any data from the analysis. Avenues of investigation should include changes made to the KABX NEXRAD weather radar station hardware or software, actual weather conditions that occurred in 2009, and any other factors that may have affected the count or intensity of thunderstorms during the study period.

The thunderstorm duration models resulted in higher slope values than the other models in the study, showing that an increase in developed land cover leads to a greater increase in thunderstorm duration than other storm characteristics. Maximum intensity (H1) and

duration (H4) did not increase in the study area as a whole during the study period, but a positive correlation between developed land cover and maximum intensity (H5, H6, H7) and duration (H14, H15, H16) was found. Because it has been shown that developed land cover increased during the study time (Table 3), these findings suggest that increases in maximum intensity and duration have occurred over urban land cover more so than over non-urban parts of the study area.

It was found that maximum intensity, mean intensity, and duration for each thunderstorm event polygon were spatially autocorrelated. Using spatial regression models would most likely result in higher $R^2$ values for Models 5 through 16. By not using spatial models, the correlation coefficients may be larger than their true value, and may not be statistically significant (Gong, 2019). Further research applying spatial regression modeling to the dataset could provide a useful comparison to the results presented in this study. Additionally, a similar research methodology could be applied to larger and expanding cities, particularly those in arid or semi-arid climates, to develop a better understanding of the relationship between urban development and thunderstorm formation and life cycles.

This study joins other related studies by creating a multi-year radar climatology to investigate the relationship between urban land cover and thunderstorm activity. Similar to Bentley et al. (2010), Niyogi et al. (2011), Ashley et al. (2012), and Haberlie et al. (2015), this study found a relationship between urban land cover and thunderstorm

activity. Like Debbage and Shepherd (2015), this study used Ordinary Least Squares (OLS) regression modeling. In line with other studies on urban climates (Debbage and Shepherd, 2015; Oke et al., 2017), the urban heat island (UHI) effect was considered in the methodology of this study. The UHI effect was incorporated when regression analyses were run using the mean minimum temperature for each thunderstorm event polygon. These results are detailed in Appendix G and were found to be similar to those that used the mean maximum temperature.

Unlike previous studies, this investigation used software (the *Find Point Clusters* tool in ArcGIS Pro) to derive thunderstorm events from radar data, rather than through human observation of radar data. Many of the previous studies used Atlanta, GA, a large urban area in an area of flat terrain. This study used a much smaller urban center, the Albuquerque metropolitan area, which is centered in a region with complex topography. Also unlike previous work, this study considered a larger geographic area, without specifying upwind and downwind regions, and did not break down the study period into weekdays and weekends or consider the effects of aerosols.

## 8.0 Conclusion

Regression analyses were used to assess the relationship between developed land cover in central New Mexico and thunderstorm characteristics. This area has complex topography and a relatively small urban footprint compared to cities that have been used in other studies, such as Atlanta, GA, or Phoenix, AZ. However, the analyses in this study resulted in several statistically significant findings. Variables that corresponded with the input for thunderstorm initiation and formation, along with variables representing developed land cover, were used to determine the relationship between developed land cover and thunderstorm characteristics. It was established that while thunderstorm maximum intensity, mean intensity, and duration did not increase over the 20-year study period, maximum intensity values of thunderstorms in the 95$^{th}$ percentile did increase. This finding implies that, by the end of the study period, whenever the most intense thunderstorms occurred, they may have presented a greater hazard to people on the ground and could have caused greater property damage.

It was also found that developed land cover was positively correlated with thunderstorm maximum intensity, maximum intensity in the 95$^{th}$ percentile, mean intensity, and duration. If the population of central New Mexico continues to grow, it may add further developed land cover to previously undeveloped areas, as well as increase the level and contiguity of existing development in established urban areas. This development has the potential to change the local atmosphere and increase thunderstorm intensity and duration. Future analyses of radar climatologies in and around urban areas in other arid and semi-arid regions may provide useful platforms on which to test other variables to

see which ones contribute to thunderstorm activity. Additionally, multi-year datasets can be potentially used to separate the effects of El Niño and La Niña (ENSO) and climate change from urban influences on local atmospheric processes.

Further research, including spatial regression modeling, is needed to continue to extract the relationship between thunderstorm characteristics and surface influences. The radar climatology dataset created for this study will be made available for future research.

## 9.0 Limitations

There are a number of polygons (see Figure 12) that appear to be potentially erroneous. These polygons are primarily located where the Albuquerque metropolitan area meets the foothills of the Sandia Mountains. This is a location where thunderstorms would be expected to form; however, these polygons appear to be grouped together more densely than elsewhere in the study area. No objective parameter was found that could be used to isolate and remove them. Due to the existence of these potentially erroneous polygons, no analysis was made using subdivisions of the study area. Analysis using smaller subdivisions of the study area would allow for comparison of thunderstorm events over or near urban areas to those over non-urban areas, as well as comparison of areas upwind and downwind of the Albuquerque metropolitan area, as in Haberlie et al. (2015).

The use of OLS regression models provided an initial look at the dataset used in this study. However, the use of spatial regression models could have provided a much more robust analysis of thunderstorm activity, as it is inherently spatially autocorrelated, due to its relationship to terrain, temperature, land cover, and other nearby thunderstorms. Additionally, thunderstorms are temporally autocorrelated, which means that spatial-temporal models are needed for a complete analysis.

Atmospheric processes are complex phenomena, with a wide range of factors that can affect local weather. These processes can range in scale from the local to the global. Some of the factors that were not taken into account in this study are aerosols and air

pollution, including particulates from wildfire smoke. Additionally, atmospheric moisture, as well as the effects of El Niño and La Niña (ENSO), and climate change, were not accounted for in this study.

## Appendices

**Appendix A: Acronyms and Abbreviations**

BLM – Bureau of Land Management

CAPE – convective available potential energy

dB – decibel

dBZ – radar reflectivity factor; units of intensity used refer to radar data

ENSO – El Niño-Southern Oscillation

FAA – Federal Aviation Administration

GIS – geographic information systems; geographic information science

ICI events – isolated convective initiation events

KABX – identifier for the Albuquerque, NM NEXRAD WSR-88D weather radar station

LFC – level of free convection

MRLC Consortium – Multi-Resolution Land Characteristics Consortium (formerly National Land Cover Database (NLCD))

MSA – metropolitan statistical area

NAM – North American monsoon

NCEI – National Centers for Environmental Information

NEXRAD – Next Generation Radar

NLCD – National Land Cover Database

NOAA – National Oceanic and Atmospheric Administration

NWS – National Weather Service

RDA – radar data acquisition

ROC – Radar Operations Center

RPG – radar product generator

SALR – saturated adiabatic lapse rate

SUHI – surface urban heat island

UHI – urban heat island

USDA – United States Department of Agriculture

VCP – volume coverage pattern

WCT – Weather and Climate Toolkit

WSR-88D – Weather Surveillance Radar – 1988, Doppler

Z – reflectivity

$Z_e$ – equivalent radar reflectivity factor

**Appendix B: Python Script for Downloading and Formatting NEXRAD Data**

This script, modified from Ansari et al. (2018), accesses NEXRAD binary data files from Amazon's S3 bucket (https://s3.amazonaws.com/noaa-nexrad-level2/index.html) and converts them into shapefiles. The script uses the Weather and Climate Toolkit, which includes wct-export.bat and wctBatchConfig.xml. The wctBatchConfig.xml file needs to be configured to the user's specifications.

Filepath names have been obscured for this document. Any future users should insert the applicable filepath name in the indicated lines in the following script.

```
###
from xml.dom import minidom
#from sys import stdin
from urllib.request import urlopen
from subprocess import call
import os
from datetime import datetime

start = datetime.now()
print ("Download started at " + str(start))
```

```
def getText(nodelist):
    rc = []
    for node in nodelist:
        if node.nodeType == node.TEXT_NODE:
            rc.append(node.data)
    return ''.join(rc)


n = 1
while n < 10:
    date = "2018/05/0"+str(n)
    print (date)
    n+=1


    site = "KABX"
    bucketURL = "http://noaa-nexrad-level2.s3.amazonaws.com"
    dirListURL = bucketURL+ "/?prefix=" + date + "/" + site
```

```
    print ("Listing files from %s" % dirListURL)



#xmldoc = minidom.parse(stdin)

    xmldoc = minidom.parse(urlopen(dirListURL))

    itemlist = xmldoc.getElementsByTagName('Key')

    print (len(itemlist) , "keys found...")



# For this test, WCT is downloaded and unzipped directly in the working directory

# The output files are going in 'Z:/__[username]__/Radar/2018'

# http://www.ncdc.noaa.gov/wct/install.php

    for x in itemlist:

        file = getText(x.childNodes)

    #print "Found %s " % file

        print ("Processing %s " % file)

    #Convert to Shapefile

        call(["wct-export.bat", "%s/%s"%(bucketURL,file),
"Z:/__[username]__/Radar/2018", "shp", "wctBatchConfigAJV.xml"], shell = True)

    #Delete temporary files
```

```
        filelist = [f for f in os.listdir("C:/Users/__[username]__/AppData/Local/Temp/wct-cache/data") if f.endswith(".uncompress")]
        for f in filelist:
            os.remove(os.path.join("C:/Users/__[username]__/AppData/Local/Temp/wct-cache/data", f))


while n <= 31:
    date = "2018/05/"+str(n)
    print (date)
    n+=1


    site = "KABX"
    bucketURL = "http://noaa-nexrad-level2.s3.amazonaws.com"
    dirListURL = bucketURL+ "/?prefix=" + date + "/" + site


    print ("Listing files from %s" % dirListURL)


#xmldoc = minidom.parse(stdin)
    xmldoc = minidom.parse(urlopen(dirListURL))
```

```
    itemlist = xmldoc.getElementsByTagName('Key')
    print (len(itemlist) , "keys found...")


# For this test, WCT is downloaded and unzipped directly in the working directory
# The output files are going in 'Z:/__[username]__/Radar/2018'
# http://www.ncdc.noaa.gov/wct/install.php
    for x in itemlist:
        file = getText(x.childNodes)
    #print "Found %s " % file
        print ("Processing %s " % file)
    #Convert to Shapefile
        call(["wct-export.bat", "%s/%s"%(bucketURL,file),
"Z:/__[username]__/Radar/2018", "shp", "wctBatchConfigAJV.xml"], shell = True)
    #Delete temporary files
        filelist = [f for f in os.listdir("C:/Users/__[username]__/AppData/Local/Temp/wct-
cache/data") if f.endswith(".uncompress")]
        for f in filelist:
            os.remove(os.path.join("C:/Users/__[username]__/AppData/Local/Temp/wct-
cache/data", f))
```

```
end = datetime.now()


elapsedtime = end - start


print ("Download started at " + str(start))

print ("Download ended at " + str(end))

print ("Time to download " + str(elapsedtime))
```

## Appendix C: Variable Names

The following table contains the full names and descriptions of the variables created for this dataset.

Table 14. All variables created for this study.

| Variable name | Description |
|---|---|
| **Cluster intensity variables within each MBG polygon (in dBZ of reflectivity)** | |
| *Cluster_Mean_Intensity_Value* | Mean intensity |
| *Cluster_Max_Intensity_Value* | Maximum intensity |
| *Cluster_Range_Intensity_Value* | Range of intensity values |
| *Cluster_StDev_Intensity_Value* | Standard deviation of the intensity values |
| *Cluster_Median_Intensity_Value* | Median intensity |
| *Cluster_Variance_Intensity_Value* | Variance of the intensity values |
| | |
| **Time variables for each MBG polygon** | |
| *START_DATE* | Beginning time and date |
| *END_DATE* | End time and date |
| *Time_Duration_minutes* | Duration, in minutes |
| | |
| **Line variables within each MBG polygon** | |
| *Line_Bearing_degrees* | The bearing from the starting vertex of the line to the ending vertex of the line |
| *Line_StartToEndDistance_meters* | The distance from the starting vertex of the line to the ending vertex of the line |
| *Line_Length_Ratio* | The value found by dividing *Line_StartToEndDistance_meters* by *Line_Length_meters* |
| *Line_Length_meters* | The length of each line, in meters |

Table 14 (cont.)

| Variable name | Description |
|---|---|
| **Land Cover variables within each MBG polygon (area in square meters)** | |
| *OpenWater_11* | Open Water |
| *Developed_OpenSpace_21* | Developed, Open Space |
| *Developed_Low_22* | Developed, Low Intensity |
| *Developed_Medium_23* | Developed, Medium Intensity |
| *Developed_High_24* | Developed, High Intensity |
| *Barren_31* | Barren Land (Rock / Sand / Clay) |
| *DeciduousForest_41* | Deciduous Forest |
| *EvergreenForest_42* | Evergreen Forest |
| *MixedForest_43* | Mixed Forest |
| *ShrubScrub_52* | Shrub / Scrub |
| *GrasslandHerb_71* | Grassland / Herbaceous |
| *PastureHay_81* | Pasture / Hay |
| *CultivatedCrops_82* | Cultivated Crops |
| *WoodyWetlands_90* | Woody Wetlands |
| *EmergHerbWetlands_95* | Emergent Herbaceous Wetlands |
| *All_GrassHerb_ShrubScrub* | Sum of Grassland / Herbaceous and Shrub / Scrub areas |
| *All_Developed* | Sum of all four Developed areas (Open Space, Low Intensity, Medium Intensity, and High Intensity) |
| *All_Developed_Barren* | Sum of all four Developed areas and Barren Land |
| *All_Forest* | Sum of all three Forest areas (Deciduous, Evergreen, and Mixed) |
| *All_Vegetated* | Sum of Deciduous, Evergreen, and Mixed Forest; Shrub / Scrub; Grassland / Herbaceous; Pasture / Hay; Cultivated Crops; Woody Wetlands; and Emergent Herbaceous Wetlands |

Table 14 (cont.)

| Variable name | Description |
|---|---|
| **Land Cover variables within each MBG polygon (area in square meters) (cont.)** | |
| *Develop_Med_Low* | Sum of Medium Intensity and Low Intensity Developed areas |
| *Develop_Med_Low_OpnSp* | Sum of Medium Intensity, Low Intensity, and Open Space Developed areas |
| *Develop_High_Med* | Sum of High Intensity and Medium Intensity Developed areas |
| *Develop_High_Med_Low* | Sum of High Intensity, Medium Intensity, and Low Intensity areas |
| *OW11_PH81_CC82_WW90_EHW95* | Sum of Open Water, Pasture / Hay, Cultivated Crops, Woody Wetlands, and Emergent Herbaceous Wetlands areas |
| *OW11_WW90_EHW95* | Sum of Open Water, Woody Wetlands, and Emergent Herbaceous Wetlands areas |
| *PH81_CC82* | Sum of Pasture / Hay and Cultivated Crops areas |
| *PH81_CC82_WW90_EHW95* | Sum of Pasture / Hay, Cultivated Crops, Woody Wetlands, and Emergent Herbaceous Wetlands areas |
| | |
| **Land Cover variables within each MBG polygon (percentage)** | |
| *Percent_All_Developed* | Percentage of the sum of all four Developed areas (Open Space, Low Intensity, Medium Intensity, and High Intensity) |
| *Percent_All_Developed_Barren* | Percentage of the sum of all four Developed areas and Barren Land |
| *Percent_All_Forest* | Percentage of the sum of all three Forest areas (Deciduous, Evergreen, and Mixed) |
| *Percent_All_Vegetated* | Percentage of the sum of Deciduous, Evergreen, and Mixed Forest; Shrub / Scrub; Grassland / Herbaceous; Pasture / Hay; Cultivated Crops; Woody Wetlands; and Emergent Herbaceous Wetlands |

Table 14 (cont.)

| Variable name | Description |
|---|---|
| **Land Cover variables within each MBG polygon (percentage) (cont.)** | |
| *Percent_GrassHerb_ShrubScrub* | Percentage of the sum of Grassland / Herbaceous and Shrub / Scrub areas |
| *Percent_OpenWater* | Percentage of Open Water area |
| *Percent_Developed_OpenSpace* | Percentage of Developed, Open Space area |
| *Percent_Developed_Low* | Percentage of Developed, Low Intensity area |
| *Percent_Developed_Medium* | Percentage of Developed, Medium Intensity area |
| *Percent_Developed_High* | Percentage of Developed, High Intensity area |
| *Percent_Barren* | Percentage of Barren Land (Rock / Sand / Clay) |
| *Percent_DeciduousForest* | Percentage of Deciduous Forest |
| *Percent_EvergreenForest* | Percentage of Evergreen Forest |
| *Percent_MixedForest* | Percentage of Mixed Forest |
| *Percent_ShrubScrub* | Percentage of Shrub / Scrub |
| *Percent_GrasslandHerb* | Percentage of Grassland / Herbaceous |
| *Percent_PastureHay* | Percentage of Pasture / Hay |
| *Percent_CultivatedCrops* | Percentage of Cultivated Crops |
| *Percent_WoodyWetlands* | Percentage of Woody Wetlands |
| *Percent_EmHerbWetlands* | Percentage of Emergent Herbaceous Wetlands |
| *Percent_Develop_Med_Low* | Percentage of the sum of Medium Intensity and Low Intensity Developed areas |
| *Percent_Develop_Med_Low_OpnSp* | Percentage of the sum of Medium Intensity, Low Intensity, and Open Space Developed areas |
| *Percent_Develop_High_Med* | Percentage of the sum of High Intensity and Medium Intensity Developed areas |
| *Percent_Develop_High_Med_Low* | Percentage of the sum of High Intensity, Medium Intensity, and Low Intensity areas |

Table 14 (cont.)

| Variable name | Description |
|---|---|
| **Land Cover variables within each MBG polygon (percentage) (cont.)** | |
| *Percent_OW11PH81CC82WW90EHW95* | Percentage of the sum of Open Water, Pasture / Hay, Cultivated Crops, Woody Wetlands, and Emergent Herbaceous Wetlands areas |
| *Percent_OW11_WW90_EHW95* | Percentage of the sum of Open Water, Woody Wetlands, and Emergent Herbaceous Wetlands areas |
| *Percent_PH81_CC82* | Percentage of the sum of Pasture / Hay and Cultivated Crops areas |
| *Percent_PH81_CC82_WW90_EHW95* | Percentage of the sum of Pasture / Hay, Cultivated Crops, Woody Wetlands, and Emergent Herbaceous Wetlands areas |
| | |
| **Terrain variables within each MBG polygon** | |
| *Elevation_min_meters* | Minimum elevation value, in meters |
| *Elevation_max_meters* | Maximum elevation value, in meters |
| *Elevation_range_meters* | Range of elevation values, in meters |
| *Elevation_mean_meters* | Mean elevation value, in meters |
| *Aspect_median* | Median aspect value, in degrees |
| *Slope_min* | Minimum slope value, in degrees |
| *Slope_max* | Maximum slope value, in degrees |
| *Slope_range* | Range of slope values, in degrees |
| *Slope_mean* | Mean slope value, in degrees |
| | |
| **Temperature variables within each MBG polygon (in degrees Celsius)** | |
| *Min_MaxTemp* | Minimum of the maximum temperature, in degrees Celsius |
| *Max_MaxTemp* | Maximum of the maximum temperature, in degrees Celsius |
| *Range_MaxTemp* | Range of maximum temperature values, in degrees Celsius |
| *Mean_MaxTemp* | Mean of the maximum temperature, in degrees Celsius |

Table 14 (cont.)

| Variable name | Description |
|---|---|
| **Temperature variables within each MBG polygon (in degrees Celsius) (cont.)** | |
| *Min_MeanTemp* | Minimum of the mean temperature, in degrees Celsius |
| *Max_MeanTemp* | Maximum of the mean temperature, in degrees Celsius |
| *Range_MeanTemp* | Range of mean temperature values, in degrees Celsius |
| *Mean_MeanTemp* | Mean of the mean temperature, in degrees Celsius |
| *Min_MinTemp* | Minimum of the minimum temperature, in degrees Celsius |
| *Max_MinTemp* | Maximum of the minimum temperature, in degrees Celsius |
| *Range_MinTemp* | Range of minimum temperature values, in degrees Celsius |
| *Mean_MinTemp* | Mean of the minimum temperature, in degrees Celsius |
| | |
| **Area variables within each MBG polygon** | |
| *MBG_Area_Sqm* | Area in square meters |
| *Area_SqKilometers* | Area in square kilometers |
| | |
| **Additional time variables for each MBG polygon** | |
| *Year* | Year of the start date |
| *Month* | Month of the start date |
| *MonthTxt* | Month of the start date, in text data format |
| *Month_and_Year* | Month and year of the start date |
| *Week* | Week of the start date |
| *Year_short* | Year of the start date, in short data format |
| *Month_short* | Month of the start date, in short data format |
| *Week_short* | Week of the start date, in short data format |

## Appendix D: Correlation Matrices

The following correlation matrices were run for relevant variables used in this study.

Table 15. Correlation matrix for variables related to thunderstorm formation.

Correlation Matrix for All Relevant Variables

| | Aspect_median | Slope_range | Elevation_range_meters | Range_MaxTemp | Mean_MaxTemp | Mean_MinTemp | Percent_All_Forest | Percent_PastureHay | Percent_CultivatedCrops | Percent_WoodyWetlands | Percent_EmHerbWetlands | Percent_OpenWater | Percent_Barren |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Aspect_median | 0.017**** | | | | | | | | | | | | |
| Slope_range | 0.031**** | 0.81**** | | | | | | | | | | | |
| Elevation_range_meters | 0.025**** | 0.65**** | 0.88**** | | | | | | | | | | |
| Range_MaxTemp | -0.033**** | -0.34**** | -0.32**** | -0.23**** | | | | | | | | | |
| Mean_MaxTemp | 0.019**** | -0.23**** | -0.19**** | -0.13**** | 0.87**** | | | | | | | | |
| Mean_MinTemp | 0.047**** | 0.54**** | 0.47**** | 0.33**** | -0.56**** | -0.45**** | | | | | | | |
| Percent_All_Forest | -0.061**** | -0.071**** | -0.039**** | -0.02**** | 0.072**** | 0.059**** | -0.049**** | | | | | | |
| Percent_PastureHay | -0.078**** | -0.085**** | -0.057**** | -0.033**** | 0.068**** | 0.019**** | -0.08**** | 0.15**** | | | | | |
| Percent_CultivatedCrops | 0.017**** | -0.012**** | -0.011**** | 0.0041 | 0.034**** | 0.027**** | -0.0031 | 0.11**** | 0.21**** | | | | |
| Percent_WoodyWetlands | -0.11**** | -0.054**** | -0.036**** | -0.025**** | 0.013**** | -7.00E-04 | -0.03**** | 0.32**** | 0.097**** | 0.22**** | | | |
| Percent_EmHerbWetlands | -0.04**** | -0.019**** | -0.022**** | -0.0078*** | 0.057**** | 0.05**** | -0.059**** | 0.17**** | 0.13**** | 0.24**** | 0.26**** | | |
| Percent_OpenWater | -0.037**** | 0.014**** | 0.03**** | -0.0016 | -0.031**** | -0.0015 | -0.025**** | 0.026**** | 0.027**** | 0.018**** | 0.11**** | 0.021**** | |
| Percent_Barren | | | | | | | | | | | | | |

Table 16. Correlation matrix for variables used in Models 5, 8, 11, and 14.

| | Percent_All_Developed | Aspect_median | Elevation_range_meters | Mean_MaxTemp | Percent_All_Forest | Percent_PastureHay | Percent_CultivatedCrops | Percent_WoodyWetlands | Percent_EmHerbWetlands | Percent_OpenWater | Percent_Barren |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Percent_All_Developed | | 0.17**** | -0.18**** | 0.1**** | -0.21**** | 0.12**** | 0.024**** | -0.0086**** | 0.026**** | 0.033**** | -0.0067** |
| Aspect_median | | | 0.031**** | -0.033**** | 0.047**** | -0.061**** | -0.078**** | 0.017**** | -0.11**** | -0.04**** | -0.037**** |
| Elevation_range_meters | | | | -0.32**** | 0.47**** | -0.039**** | -0.057**** | -0.011**** | -0.036**** | -0.022**** | 0.03**** |
| Mean_MaxTemp | | | | | -0.56**** | 0.072**** | 0.068**** | 0.034**** | 0.013**** | 0.057**** | -0.031**** |
| Percent_All_Forest | | | | | | -0.049**** | -0.08**** | -0.0031 | -0.03**** | -0.059**** | -0.025**** |
| Percent_PastureHay | | | | | | | 0.15**** | 0.11**** | 0.32**** | 0.17**** | 0.026**** |
| Percent_CultivatedCrops | | | | | | | | 0.21**** | 0.097**** | 0.13**** | 0.027**** |
| Percent_WoodyWetlands | | | | | | | | | 0.22**** | 0.24**** | 0.018**** |
| Percent_EmHerbWetlands | | | | | | | | | | 0.26**** | 0.11**** |
| Percent_OpenWater | | | | | | | | | | | 0.021**** |
| Percent_Barren | | | | | | | | | | | |

Table 17. Correlation matrix for variables used in Models 6, 9, 12, and 15.

| | Percent_Developed_High | Aspect_median | Elevation_range_meters | Mean_MaxTemp | Percent_All_Forest | Percent_PastureHay | Percent_CultivatedCrops | Percent_WoodyWetlands | Percent_EmHerbWetlands | Percent_OpenWater | Percent_Barren |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Percent_Developed_High | | 0.098**** | -0.11**** | 0.063**** | -0.12**** | 0.038**** | -0.0057** | -0.015**** | 0.0058** | 0.02**** | -0.013**** |
| Aspect_median | | | 0.031**** | -0.033**** | 0.047**** | -0.061**** | -0.078**** | 0.017**** | -0.11**** | -0.04**** | -0.037**** |
| Elevation_range_meters | | | | -0.32**** | 0.47**** | -0.039**** | -0.057**** | -0.011**** | -0.036**** | -0.022**** | 0.03**** |
| Mean_MaxTemp | | | | | -0.56**** | 0.072**** | 0.068**** | 0.034**** | 0.013**** | 0.057**** | -0.031**** |
| Percent_All_Forest | | | | | | -0.049**** | -0.08**** | -0.0031 | -0.03**** | -0.059**** | -0.025**** |
| Percent_PastureHay | | | | | | | 0.15**** | 0.11**** | 0.32**** | 0.17**** | 0.026**** |
| Percent_CultivatedCrops | | | | | | | | 0.21**** | 0.097**** | 0.13**** | 0.027**** |
| Percent_WoodyWetlands | | | | | | | | | 0.22**** | 0.24**** | 0.018**** |
| Percent_EmHerbWetlands | | | | | | | | | | 0.26**** | 0.11**** |
| Percent_OpenWater | | | | | | | | | | | 0.021**** |
| Percent_Barren | | | | | | | | | | | |

Table 18. Correlation matrix for variables used in Models 7, 10, 13, and 16.

| | Percent_Develop_Med_Low_OpnSp | Aspect_median | Elevation_range_meters | Mean_MaxTemp | Percent_All_Forest | Percent_PastureHay | Percent_CultivatedCrops | Percent_WoodyWetlands | Percent_EmHerbWetlands | Percent_OpenWater | Percent_Barren |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Percent_Develop_Med_Low_OpnSp | | 0.17**** | -0.18**** | 0.1**** | -0.21**** | 0.12**** | 0.027**** | -0.0076*** | 0.027**** | 0.033**** | -0.0057** |
| Aspect_median | | | 0.031**** | -0.033**** | 0.047**** | -0.061**** | -0.078**** | 0.017**** | -0.11**** | -0.04**** | -0.037**** |
| Elevation_range_meters | | | | -0.32**** | 0.47**** | -0.039**** | -0.057**** | -0.011**** | -0.036**** | -0.022**** | 0.03**** |
| Mean_MaxTemp | | | | | -0.56**** | 0.072**** | 0.068**** | 0.034**** | 0.013**** | 0.057**** | -0.031**** |
| Percent_All_Forest | | | | | | -0.049**** | -0.08**** | -0.0031 | -0.03**** | -0.059**** | -0.025**** |
| Percent_PastureHay | | | | | | | 0.15**** | 0.11**** | 0.32**** | 0.17**** | 0.026**** |
| Percent_CultivatedCrops | | | | | | | | 0.21**** | 0.097**** | 0.13**** | 0.027**** |
| Percent_WoodyWetlands | | | | | | | | | 0.22**** | 0.24**** | 0.018**** |
| Percent_EmHerbWetlands | | | | | | | | | | 0.26**** | 0.11**** |
| Percent_OpenWater | | | | | | | | | | | 0.021**** |
| Percent_Barren | | | | | | | | | | | |

## Appendix E: Count of Thunderstorm Events

The following graphs show the number of thunderstorm events, as captured by this dataset, by year and by month. Monthly counts show that there is substantial variation in the number of events throughout the summer seasons of the study period, and provide a useful reference for understanding the intensity and duration measures gathered from this dataset.

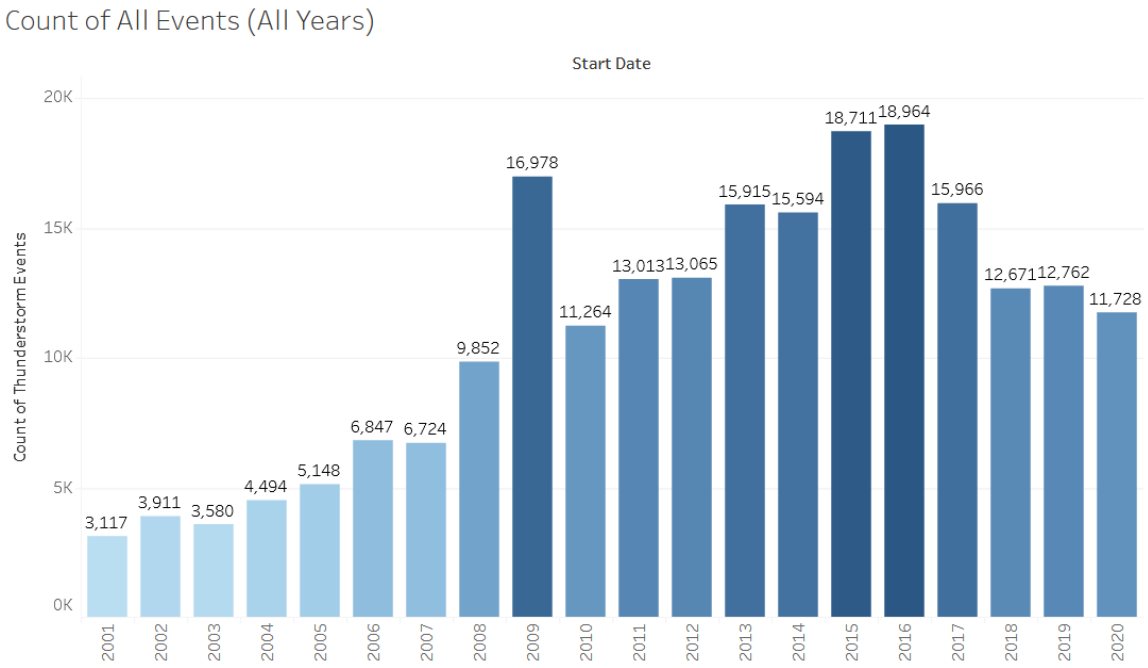



Figure 13. Thunderstorm event count by year. The year 2009 stands out in the first half of the study period as having more thunderstorm events, and it is only surpassed by the years 2015 and 2016. It is unknown why early years in the study have far fewer thunderstorm events captured in this dataset than later years, though it may be due to changes and updates in radar technology. However, the most recent years do not have the highest count, so the numbers did not increase continuously over time. Other than the year 2009, the peak in thunderstorm count occurred in 2016, and then showed a decrease through the end of the study in 2020. Numerous other factors may account for these variations, including El Niño / La Niña (ENSO) and climate change.

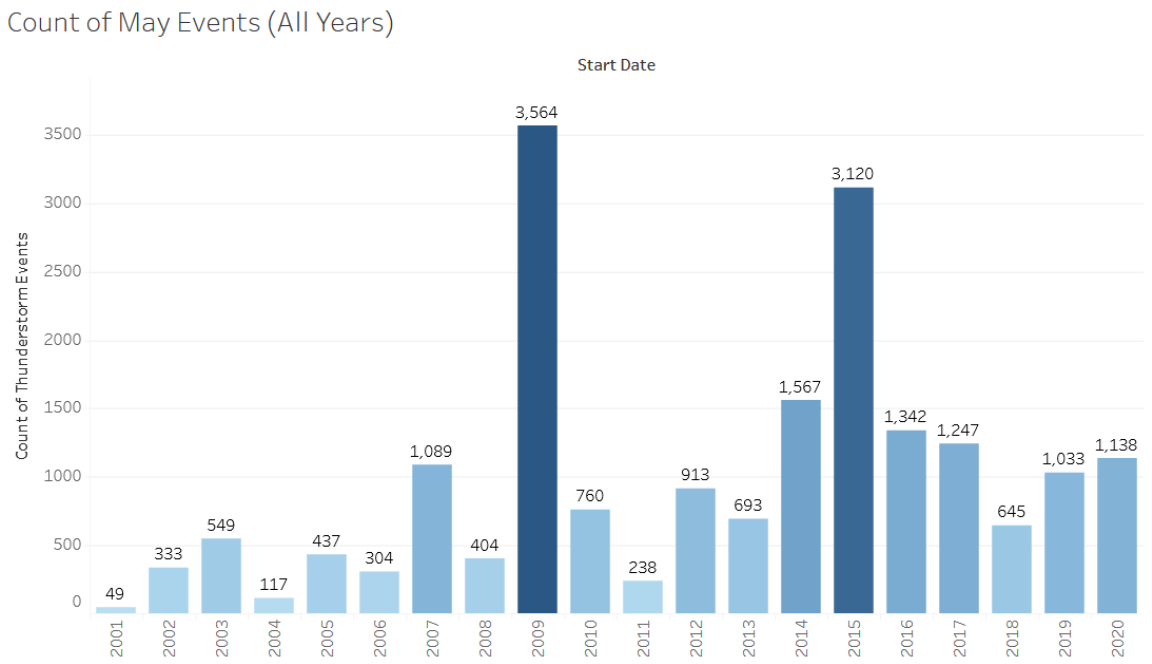



Figure 14. Thunderstorm event count for the month of May for each year. In a monthly breakdown of the count, for the month of May, the year 2009 is especially conspicuous as having the most thunderstorm events. The only other year that similarly stands out is 2015.

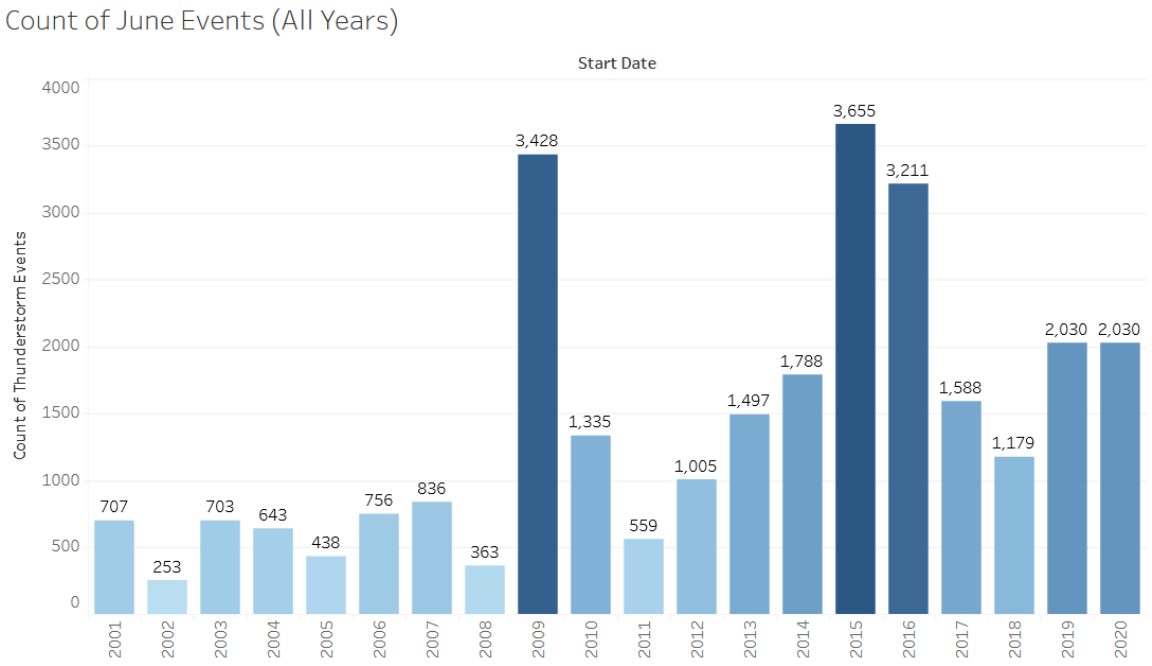



Figure 15. Thunderstorm event count for the month of June for each year. The year 2009 again stands out as having far more thunderstorms than most other years for the month of June. Again, 2015, and also 2016, have similarly high counts.

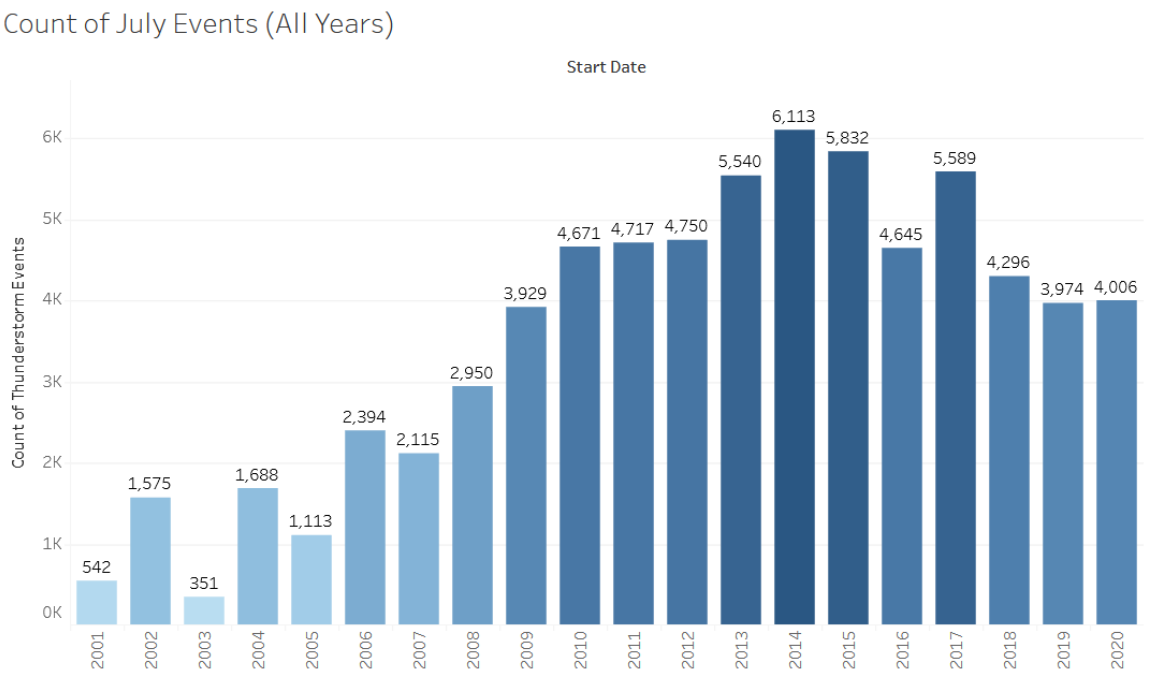



Figure 16. Thunderstorm event count for the month of July for each year. The year 2009 does not stand out as having a high count of thunderstorm events.

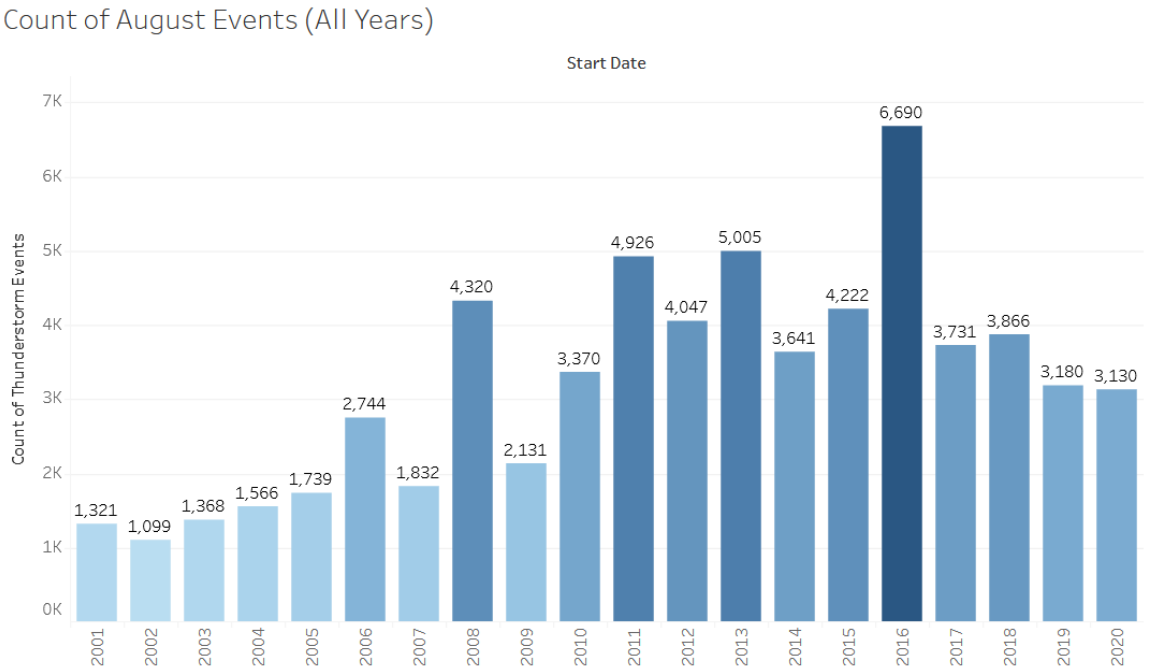



Figure 17. Thunderstorm event count for the month of August for each year. Again, in August, the year 2009 does not stand out as having a high count of thunderstorm events.

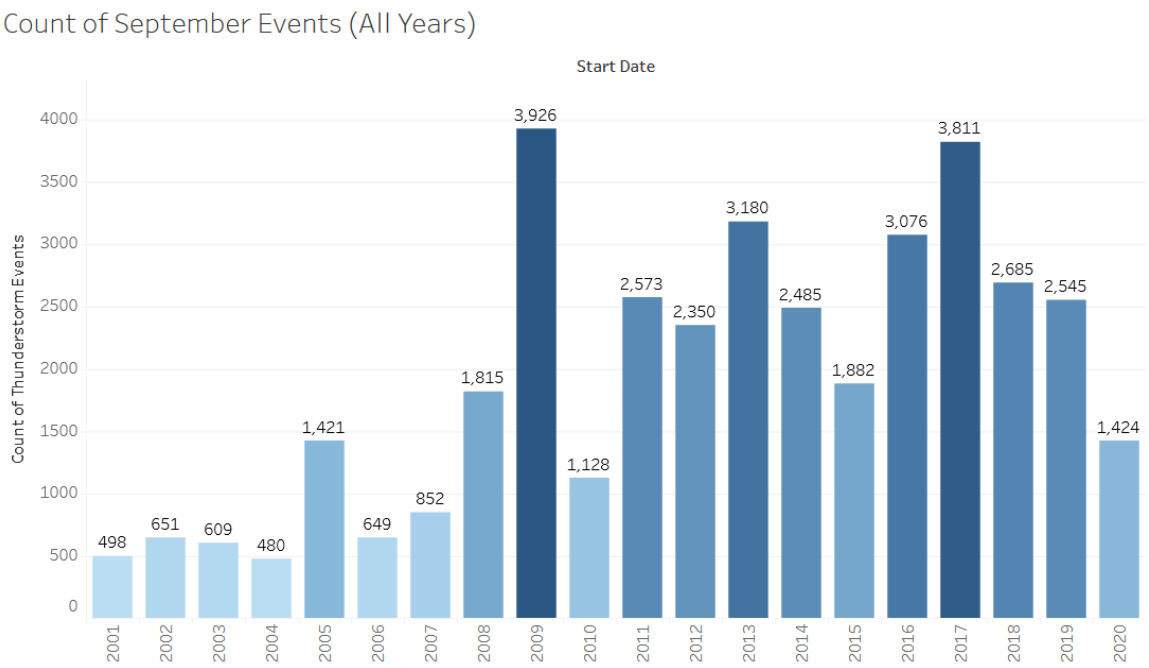



Figure 18. Thunderstorm event count for the month of September for each year. In September, the year 2009 once again stands out as having a high number of thunderstorm events, though there is more variation between years.

## Appendix F: Graphical Results for Models 1, 2, 3, and 4

Scatterplots and box plots for Models 1, 2, 3, and 4 are provided below. As noted in Appendix E, the year 2009 had more thunderstorm events than other years in the first half of the study, especially in May, June, and September.

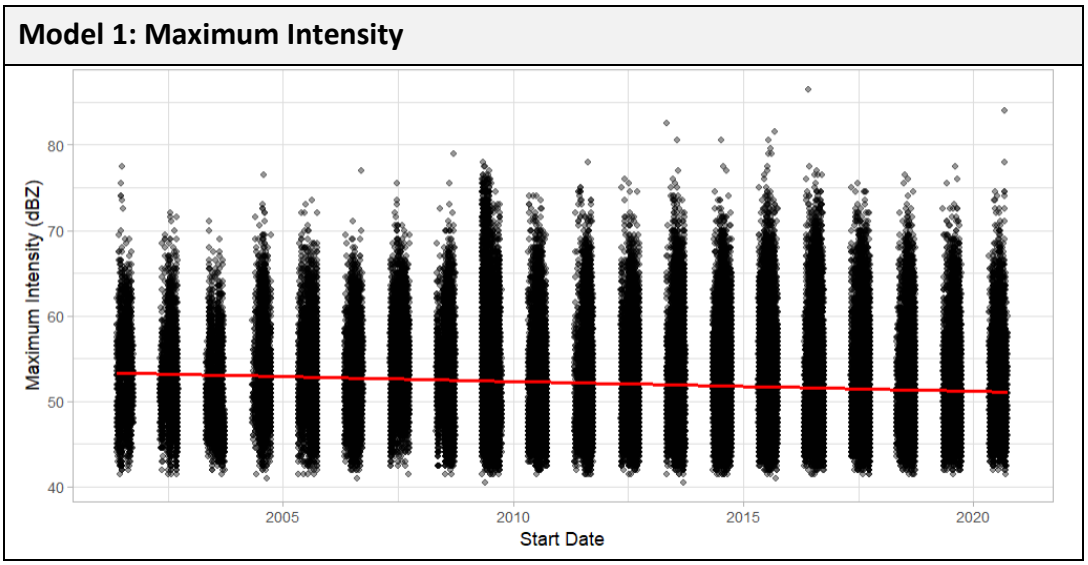



Figure 19. Scatterplot for Model 1. The regression line has a slope of -3.640e-09.

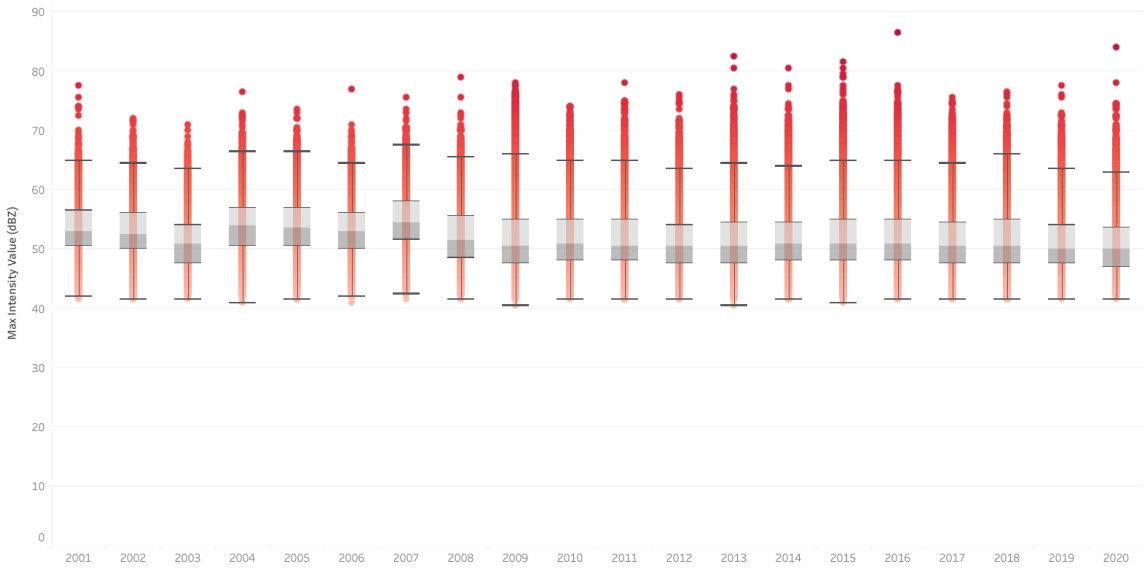



Figure 20. Box plot showing the distribution of maximum intensity values for each year of the study.

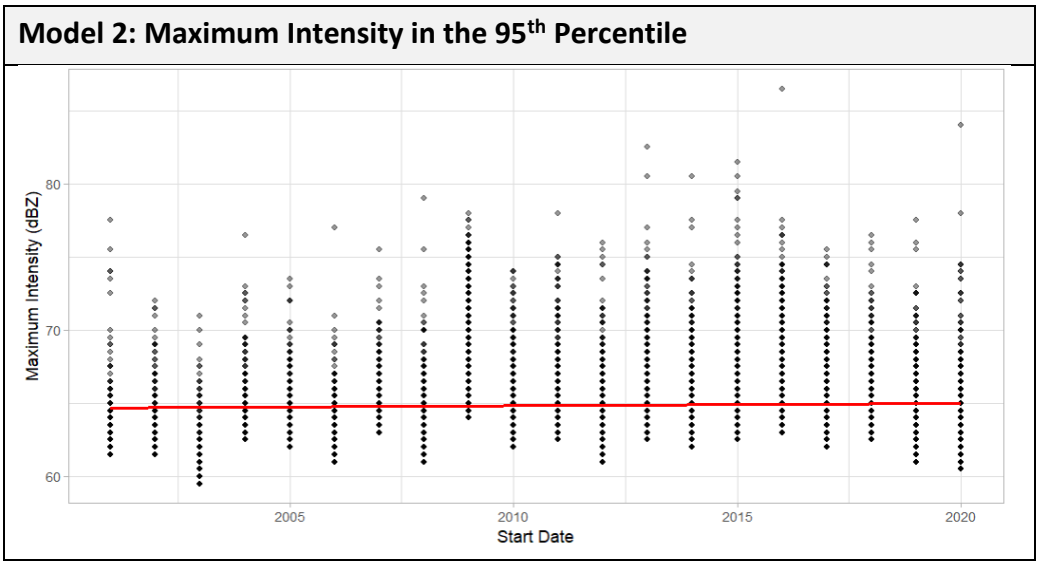



Figure 21. Scatterplot for Model 2. The regression line has a slope of 0.016235.

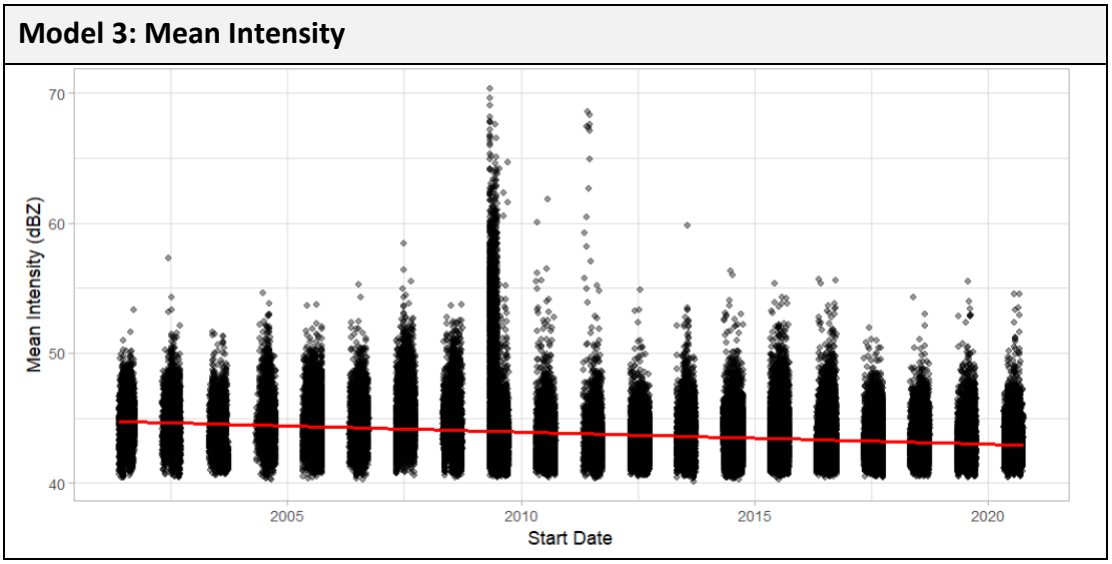



Figure 22. Scatterplot for Model 3. The regression line has a slope of -2.970e-09.

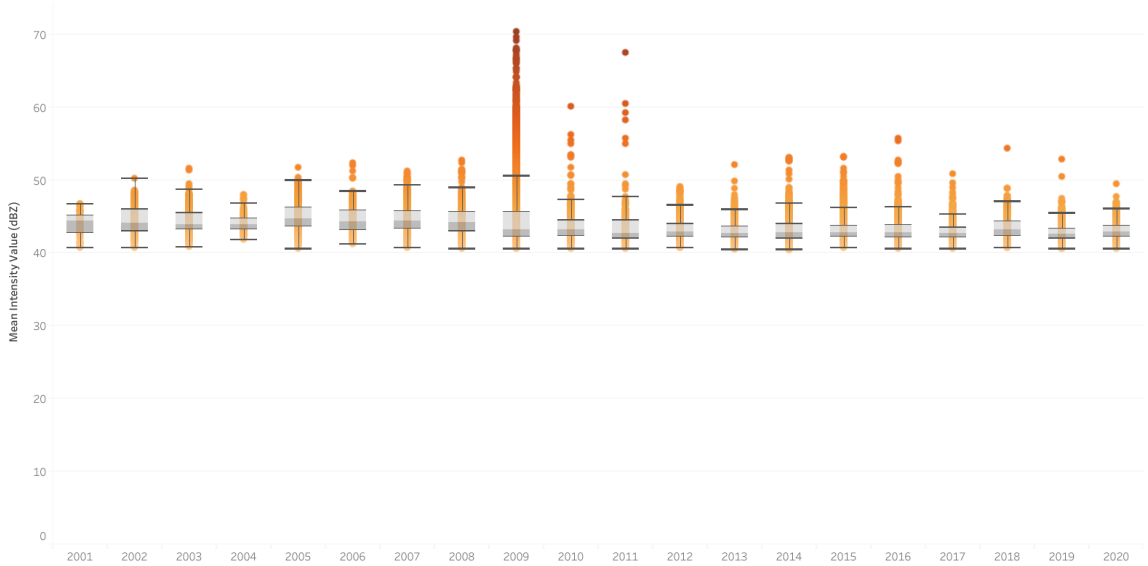



Figure 23. Box plot showing the distribution of mean intensity values for each year of the study.

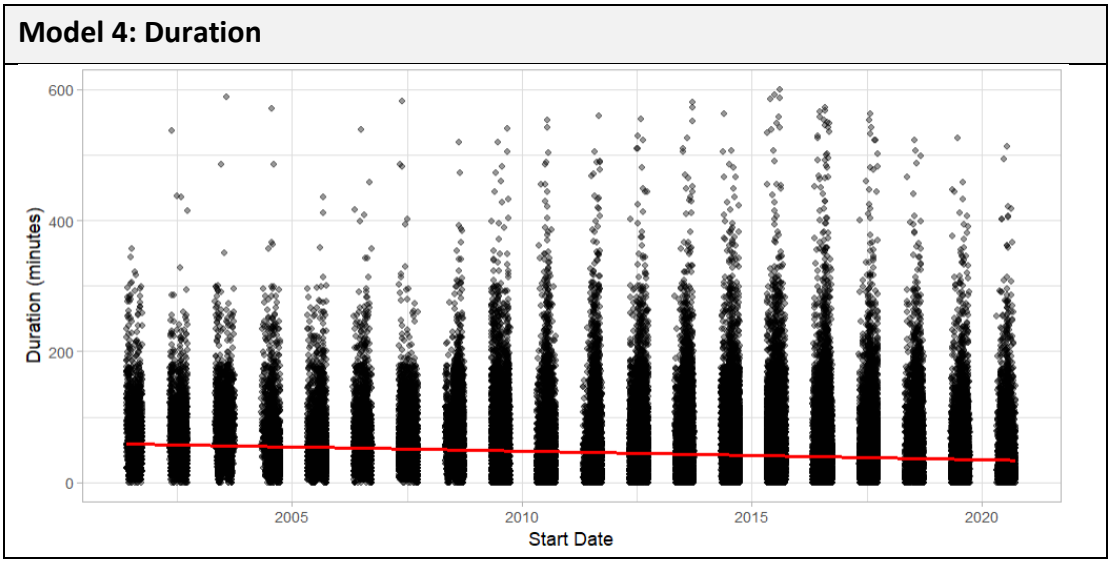



Figure 24. Scatterplot for Model 4. The regression line has a slope of -4.145e-08.

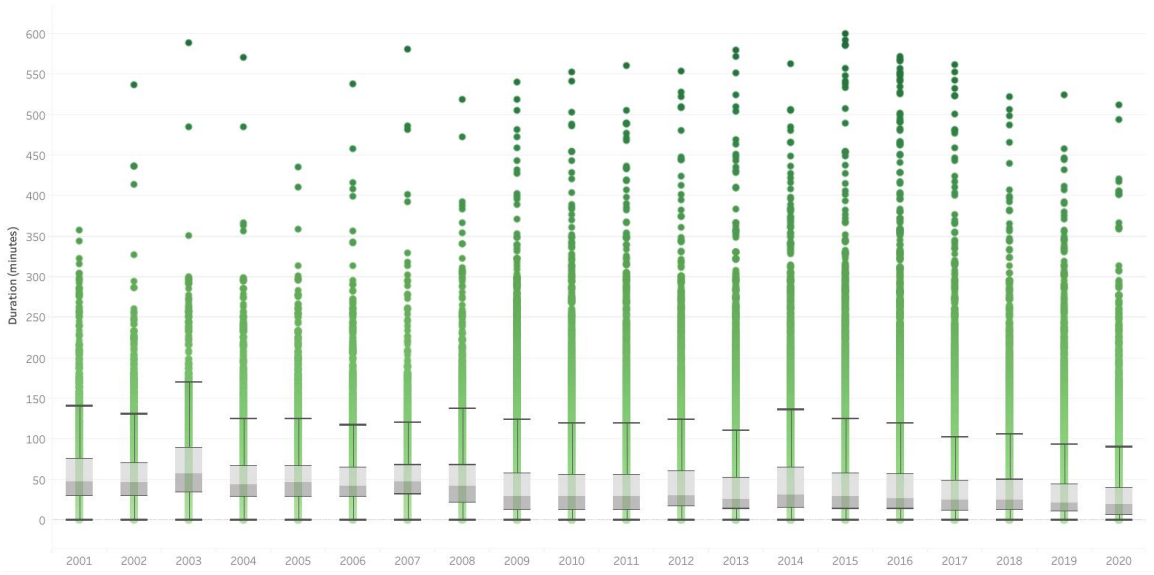



Figure 25. Box plot showing the distribution of duration times for each year of the study.

## Appendix G: Detailed Regression Results for Models 5 through 16

The following results include the slope, p-value, and statistical significance for each variable used in the OLS multivariate regression. Results are also included for models that used the mean of the minimum temperature (*Mean_MinTemp*) instead of the mean of the maximum temperature.

Table 19. Results for Model 5.

| **Model 5** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity<br>Model adjusted $R^2$: 0.2007<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.478e+01 | < 2e-16 | *** |
| *Percent_All_Developed* | 3.792e-02 | < 2e-16 | *** |
| *Aspect_median* | 8.586e-05 | 0.583 | None |
| *Elevation_range_meters* | 8.725e-03 | < 2e-16 | *** |
| *Mean_MaxTemp* | 1.821e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -2.293e-02 | < 2e-16 | *** |
| *Percent_PastureHay* | -4.082e-02 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | 7.763e-03 | 0.167 | None |
| *Percent_WoodyWetlands* | -5.086e-02 | 7.29e-12 | *** |
| *Percent_EmHerbWetlands* | 5.008e-02 | 6.47e-14 | *** |
| *Percent_OpenWater* | -1.071e-01 | 4.59e-13 | *** |
| *Percent_Barren* | -1.642e-01 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 20. Results for Model 5, using the mean minimum temperature.

| **Model 5 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity<br>Model adjusted $R^2$: 0.196<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.847e+01 | < 2e-16 | *** |
| *Percent_All_Developed* | 3.487e-02 | < 2e-16 | *** |
| *Aspect_median* | -1.300e-04 | 0.407941 | None |
| *Elevation_range_meters* | 8.496e-03 | < 2e-16 | *** |
| *Mean_MinTemp* | 1.436e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -2.713e-02 | < 2e-16 | *** |
| *Percent_PastureHay* | -3.647e-02 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | 1.864e-02 | 0.000939 | *** |
| *Percent_WoodyWetlands* | -4.777e-02 | 1.40e-10 | *** |
| *Percent_EmHerbWetlands* | 4.834e-02 | 5.36e-13 | *** |
| *Percent_OpenWater* | -1.056e-01 | 1.15e-12 | *** |
| *Percent_Barren* | -1.905e-01 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 21. Results for Model 6.

| **Model 6** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity<br>Model adjusted $R^2$: 0.1937<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.491e+01 | < 2e-16 | *** |
| *Percent_Developed_High* | 2.149e-01 | < 2e-16 | *** |
| *Aspect_median* | 1.169e-03 | 5.14e-14 | *** |
| *Elevation_range_meters* | 8.603e-03 | < 2e-16 | *** |
| *Mean_MaxTemp* | 1.784e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -2.495e-02 | < 2e-16 | *** |
| *Percent_PastureHay* | -2.315e-02 | 3.38e-11 | *** |
| *Percent_CultivatedCrops* | 1.251e-02 | 0.0266 | * |
| *Percent_WoodyWetlands* | -5.750e-02 | 1.22e-14 | *** |
| *Percent_EmHerbWetlands* | 5.018e-02 | 7.32e-14 | *** |
| *Percent_OpenWater* | -1.032e-01 | 3.86e-12 | *** |
| *Percent_Barren* | -1.634e-01 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 22. Results for Model 6, using the mean minimum temperature.

| **Model 6 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity<br>Model adjusted $R^2$: 0.1901<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.842e+01 | < 2e-16 | *** |
| *Percent_Developed_High* | 1.971e-01 | < 2e-16 | *** |
| *Aspect_median* | 8.577e-04 | 3.62e-08 | *** |
| *Elevation_range_meters* | 8.384e-03 | < 2e-16 | *** |
| *Mean_MinTemp* | 1.493e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -2.851e-02 | < 2e-16 | *** |
| *Percent_PastureHay* | -2.072e-02 | 3.21e-09 | *** |
| *Percent_CultivatedCrops* | 2.320e-02 | 4.09e-05 | *** |
| *Percent_WoodyWetlands* | -5.452e-02 | 2.91e-13 | *** |
| *Percent_EmHerbWetlands* | 4.907e-02 | 2.91e-13 | *** |
| *Percent_OpenWater* | -1.029e-01 | 5.04e-12 | *** |
| *Percent_Barren* | -1.890e-01 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 23. Results for Model 7.

| **Model 7** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity<br>Model adjusted $R^2$: 0.2009<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.477e+01 | < 2e-16 | *** |
| *Percent_Develop_Med_Low_OpnSp* | 4.151e-02 | < 2e-16 | *** |
| *Aspect_median* | 3.908e-05 | 0.803 | None |
| *Elevation_range_meters* | 8.729e-03 | < 2e-16 | *** |
| *Mean_MaxTemp* | 1.823e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -2.284e-02 | < 2e-16 | *** |
| *Percent_PastureHay* | -4.207e-02 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | 7.103e-03 | 0.206 | None |
| *Percent_WoodyWetlands* | -5.080e-02 | 7.68e-12 | *** |
| *Percent_EmHerbWetlands* | 5.012e-02 | 6.12e-14 | *** |
| *Percent_OpenWater* | -1.067e-01 | 5.68e-13 | *** |
| *Percent_Barren* | -1.647e-01 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 24. Results for Model 7, using the mean minimum temperature.

| **Model 7 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity<br>Model adjusted $R^2$: 0.1962<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.847e+01 | < 2e-16 | *** |
| *Percent_Develop_Med_Low_OpnSp* | 3.818e-02 | < 2e-16 | *** |
| *Aspect_median* | -1.730e-04 | 0.27103 | None |
| *Elevation_range_meters* | 8.499e-03 | < 2e-16 | *** |
| *Mean_MinTemp* | 1.435e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -2.706e-02 | < 2e-16 | *** |
| *Percent_PastureHay* | -3.760e-02 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | 1.803e-02 | 0.00137 | ** |
| *Percent_WoodyWetlands* | -4.769e-02 | 1.49e-10 | *** |
| *Percent_EmHerbWetlands* | 4.836e-02 | 5.21e-13 | *** |
| *Percent_OpenWater* | -1.052e-01 | 1.41e-12 | *** |
| *Percent_Barren* | -1.910e-01 | < 2e-16 | *** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 25. Results for Model 8.

| **Model 8** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity ($95^{th}$ percentile subset)<br>Model adjusted $R^2$: 0.03356<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 6.739e+01 | < 2e-16 | *** |
| *Percent_All_Developed* | 9.493e-03 | 7.17e-07 | *** |
| *Aspect_median* | 2.642e-03 | 3.01e-07 | *** |
| *Elevation_range_meters* | 5.849e-04 | < 2e-16 | *** |
| *Mean_MaxTemp* | -1.168e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -2.293e-03 | 0.06753 | . |
| *Percent_PastureHay* | 6.825e-03 | 0.68282 | None |
| *Percent_CultivatedCrops* | 3.299e-02 | 0.33607 | None |
| *Percent_WoodyWetlands* | -1.686e-01 | 0.00108 | ** |
| *Percent_EmHerbWetlands* | 1.918e-02 | 0.65684 | None |
| *Percent_OpenWater* | -7.288e-02 | 0.64271 | None |
| *Percent_Barren* | -1.369e-01 | 0.10105 | None |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 26. Results for Model 8, using the mean minimum temperature.

| **Model 8 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity (95$^{th}$ percentile subset)<br>Model adjusted $R^2$: 0.0321<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 6.523e+01 | < 2e-16 | *** |
| *Percent_All_Developed* | 1.138e-02 | 2.17e-09 | *** |
| *Aspect_median* | 2.948e-03 | 1.43e-08 | *** |
| *Elevation_range_meters* | 6.325e-04 | < 2e-16 | *** |
| *Mean_MinTemp* | -1.102e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | 6.621e-05 | 0.95441 | None |
| *Percent_PastureHay* | 6.549e-03 | 0.69525 | None |
| *Percent_CultivatedCrops* | 1.195e-02 | 0.72832 | None |
| *Percent_WoodyWetlands* | -1.660e-01 | 0.00131 | ** |
| *Percent_EmHerbWetlands* | 3.998e-03 | 0.92627 | None |
| *Percent_OpenWater* | -3.757e-02 | 0.81123 | None |
| *Percent_Barren* | -1.061e-01 | 0.20458 | None |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 27. Results for Model 9.

| **Model 9** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity (95$^{th}$ percentile subset)<br>Model adjusted $R^2$: 0.03221<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 6.749e+01 | < 2e-16 | *** |
| *Percent_Developed_High* | 5.329e-02 | 0.003904 | ** |
| *Aspect_median* | 3.069e-03 | 1.26e-09 | *** |
| *Elevation_range_meters* | 5.604e-04 | < 2e-16 | *** |
| *Mean_MaxTemp* | -1.205e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -3.039e-03 | 0.014395 | * |
| *Percent_PastureHay* | 1.120e-02 | 0.501767 | None |
| *Percent_CultivatedCrops* | 3.443e-02 | 0.315852 | None |
| *Percent_WoodyWetlands* | -1.763e-01 | 0.000632 | *** |
| *Percent_EmHerbWetlands* | 2.088e-02 | 0.628786 | None |
| *Percent_OpenWater* | -5.707e-02 | 0.716514 | None |
| *Percent_Barren* | -1.443e-01 | 0.083935 | . |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 28. Results for Model 9, using the mean minimum temperature.

| **Model 9 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity (95$^{th}$ percentile subset)<br>Model adjusted $R^2$: 0.03013<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 6.523e+01 | < 2e-16 | *** |
| *Percent_Developed_High* | 6.420e-02 | 0.000487 | *** |
| *Aspect_median* | 3.454e-03 | 1.3e-11 | *** |
| *Elevation_range_meters* | 6.034e-04 | < 2e-16 | *** |
| *Mean_MinTemp* | -1.113e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -5.481e-04 | 0.634496 | None |
| *Percent_PastureHay* | 1.154e-02 | 0.489572 | None |
| *Percent_CultivatedCrops* | 1.356e-02 | 0.693805 | None |
| *Percent_WoodyWetlands* | -1.762e-01 | 0.000648 | *** |
| *Percent_EmHerbWetlands* | 5.810e-03 | 0.893140 | None |
| *Percent_OpenWater* | -1.880e-02 | 0.904952 | None |
| *Percent_Barren* | -1.156e-01 | 0.167311 | None |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 29. Results for Model 10.

| **Model 10** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity (95$^{th}$ percentile subset)<br>Model adjusted $R^2$: 0.03363<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 6.739e+01 | < 2e-16 | *** |
| *Percent_Develop_Med_Low_OpnSp* | 1.045e-02 | 4.61e-07 | *** |
| *Aspect_median* | 2.624e-03 | 3.72e-07 | *** |
| *Elevation_range_meters* | 5.854e-04 | < 2e-16 | *** |
| *Mean_MaxTemp* | -1.167e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | -2.267e-03 | 0.0707 | . |
| *Percent_PastureHay* | 6.553e-03 | 0.6949 | None |
| *Percent_CultivatedCrops* | 3.268e-02 | 0.3407 | None |
| *Percent_WoodyWetlands* | -1.683e-01 | 0.0011 | ** |
| *Percent_EmHerbWetlands* | 1.906e-02 | 0.6587 | None |
| *Percent_OpenWater* | -7.294e-02 | 0.6424 | None |
| *Percent_Barren* | -1.367e-01 | 0.1014 | None |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 30. Results for Model 10, using the mean minimum temperature.

| **Model 10 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: maximum intensity (95$^{th}$ percentile subset)<br>Model adjusted $R^2$: 0.03221<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 6.523e+01 | < 2e-16 | *** |
| *Percent_Develop_Med_Low_OpnSp* | 1.253e-02 | 1.14e-09 | *** |
| *Aspect_median* | 2.926e-03 | 1.86e-08 | *** |
| *Elevation_range_meters* | 6.332e-04 | < 2e-16 | *** |
| *Mean_MinTemp* | -1.103e-01 | < 2e-16 | *** |
| *Percent_All_Forest* | 8.395e-05 | 0.94221 | None |
| *Percent_PastureHay* | 6.239e-03 | 0.70905 | None |
| *Percent_CultivatedCrops* | 1.153e-02 | 0.73733 | None |
| *Percent_WoodyWetlands* | -1.656e-01 | 0.00134 | ** |
| *Percent_EmHerbWetlands* | 3.846e-03 | 0.92908 | None |
| *Percent_OpenWater* | -3.755e-02 | 0.81130 | None |
| *Percent_Barren* | -1.058e-01 | 0.20574 | None |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 31. Results for Model 11.

| **Model 11** | | | |
|---|---|---|---|
| Dependent variable: mean intensity<br>Model adjusted $R^2$: 0.03217<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.228e+01 | < 2e-16 | *** |
| *Percent_All_Developed* | 1.035e-02 | < 2e-16 | *** |
| *Aspect_median* | 5.390e-04 | < 2e-16 | *** |
| *Elevation_range_meters* | 1.030e-03 | < 2e-16 | *** |
| *Mean_MaxTemp* | 3.414e-02 | < 2e-16 | *** |
| *Percent_All_Forest* | -1.089e-03 | 7.93e-14 | *** |
| *Percent_PastureHay* | -1.516e-02 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | -4.466e-03 | 0.03441 | * |
| *Percent_WoodyWetlands* | -1.636e-02 | 4.52e-09 | *** |
| *Percent_EmHerbWetlands* | 1.941e-03 | 0.43952 | None |
| *Percent_OpenWater* | -1.619e-02 | 0.00362 | ** |
| *Percent_Barren* | -3.920e-02 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 32. Results for Model 11, using the mean minimum temperature.

| **Model 11 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: mean intensity<br>Model adjusted $R^2$: 0.03175<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.291e+01 | < 2e-16 | *** |
| *Percent_All_Developed* | 9.703e-03 | < 2e-16 | *** |
| *Aspect_median* | 4.909e-04 | < 2e-16 | *** |
| *Elevation_range_meters* | 9.850e-04 | < 2e-16 | *** |
| *Mean_MinTemp* | 3.251e-02 | < 2e-16 | *** |
| *Percent_All_Forest* | -1.617e-03 | < 2e-16 | *** |
| *Percent_PastureHay* | -1.461e-02 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | -2.154e-03 | 0.30790 | None |
| *Percent_WoodyWetlands* | -1.618e-02 | 6.77e-09 | *** |
| *Percent_EmHerbWetlands* | 2.012e-03 | 0.42296 | None |
| *Percent_OpenWater* | -1.649e-02 | 0.00306 | ** |
| *Percent_Barren* | -4.393e-02 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 33. Results for Model 12.

| **Model 12** | | | |
|---|---|---|---|
| Dependent variable: mean intensity<br>Model adjusted $R^2$: 0.02742<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.232e+01 | < 2e-16 | *** |
| *Percent_Developed_High* | 5.482e-02 | < 2e-16 | *** |
| *Aspect_median* | 8.447e-04 | < 2e-16 | *** |
| *Elevation_range_meters* | 9.951e-04 | < 2e-16 | *** |
| *Mean_MaxTemp* | 3.309e-02 | < 2e-16 | *** |
| *Percent_All_Forest* | -1.658e-03 | < 2e-16 | *** |
| *Percent_PastureHay* | -1.027e-02 | 4.80e-15 | *** |
| *Percent_CultivatedCrops* | -3.209e-03 | 0.12946 | None |
| *Percent_WoodyWetlands* | -1.828e-02 | 6.38e-11 | *** |
| *Percent_EmHerbWetlands* | 1.979e-03 | 0.43161 | None |
| *Percent_OpenWater* | -1.497e-02 | 0.00728 | ** |
| *Percent_Barren* | -3.906e-02 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 34. Results for Model 12, using the mean minimum temperature.

| **Model 12 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: mean intensity<br>Model adjusted $R^2$: 0.02755<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.289e+01 | < 2e-16 | *** |
| *Percent_Developed_High* | 5.095e-02 | < 2e-16 | *** |
| *Aspect_median* | 7.759e-04 | < 2e-16 | *** |
| *Elevation_range_meters* | 9.526e-04 | < 2e-16 | *** |
| *Mean_MinTemp* | 3.419e-02 | < 2e-16 | *** |
| *Percent_All_Forest* | -2.015e-03 | < 2e-16 | *** |
| *Percent_PastureHay* | -1.015e-02 | 9.30e-15 | *** |
| *Percent_CultivatedCrops* | -9.203e-04 | 0.66377 | None |
| *Percent_WoodyWetlands* | -1.817e-02 | 8.17e-11 | *** |
| *Percent_EmHerbWetlands* | 2.234e-03 | 0.37479 | None |
| *Percent_OpenWater* | -1.560e-02 | 0.00518 | ** |
| *Percent_Barren* | -4.357e-02 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 35. Results for Model 13.

| **Model 13** | | | |
|---|---|---|---|
| Dependent variable: mean intensity<br>Model adjusted $R^2$: 0.03239<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.228e+01 | < 2e-16 | *** |
| *Percent_Develop_Med_Low_OpnSp* | 1.138e-02 | < 2e-16 | *** |
| *Aspect_median* | 5.241e-04 | < 2e-16 | *** |
| *Elevation_range_meters* | 1.031e-03 | < 2e-16 | *** |
| *Mean_MaxTemp* | 3.421e-02 | < 2e-16 | *** |
| *Percent_All_Forest* | -1.060e-03 | 3.48e-13 | *** |
| *Percent_PastureHay* | -1.553e-02 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | -4.650e-03 | 0.02761 | * |
| *Percent_WoodyWetlands* | -1.633e-02 | 4.83e-09 | *** |
| *Percent_EmHerbWetlands* | 1.951e-03 | 0.43698 | None |
| *Percent_OpenWater* | -1.608e-02 | 0.00385 | ** |
| *Percent_Barren* | -3.933e-02 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 36. Results for Model 13, using the mean minimum temperature.

| **Model 13 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: mean intensity<br>Model adjusted $R^2$: 0.03195<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 4.291e+01 | < 2e-16 | *** |
| *Percent_Develop_Med_Low_OpnSp* | 1.068e-02 | < 2e-16 | *** |
| *Aspect_median* | 4.768e-04 | 5.90e-16 | *** |
| *Elevation_range_meters* | 9.862e-04 | < 2e-16 | *** |
| *Mean_MinTemp* | 3.248e-02 | < 2e-16 | *** |
| *Percent_All_Forest* | -1.595e-03 | < 2e-16 | *** |
| *Percent_PastureHay* | -1.495e-02 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | -2.328e-03 | 0.27047 | None |
| *Percent_WoodyWetlands* | -1.614e-02 | 7.34e-09 | *** |
| *Percent_EmHerbWetlands* | 2.015e-03 | 0.42220 | None |
| *Percent_OpenWater* | -1.638e-02 | 0.00326 | ** |
| *Percent_Barren* | -4.406e-02 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 37. Results for Model 14.

| **Model 14** | | | |
|---|---|---|---|
| Dependent variable: duration<br>Model adjusted $R^2$: 0.2488<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 6.9644006 | < 2e-16 | *** |
| *Percent_All_Developed* | 0.6504379 | < 2e-16 | *** |
| *Aspect_median* | 0.0169918 | < 2e-16 | *** |
| *Elevation_range_meters* | 0.0768347 | < 2e-16 | *** |
| *Mean_MaxTemp* | 0.5111810 | < 2e-16 | *** |
| *Percent_All_Forest* | -0.2096986 | < 2e-16 | *** |
| *Percent_PastureHay* | -0.5715456 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | -0.0636767 | 0.161 | None |
| *Percent_WoodyWetlands* | -0.2817821 | 2.71e-06 | *** |
| *Percent_EmHerbWetlands* | 0.0721773 | 0.182 | None |
| *Percent_OpenWater* | -1.1723964 | < 2e-16 | *** |
| *Percent_Barren* | -0.9115502 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 38. Results for Model 14, using the mean minimum temperature.

| **Model 14 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: duration<br>Model adjusted $R^2$: 0.2482<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 17.6753635 | < 2e-16 | *** |
| *Percent_All_Developed* | 0.6422580 | < 2e-16 | *** |
| *Aspect_median* | 0.0164240 | < 2e-16 | *** |
| *Elevation_range_meters* | 0.0761991 | < 2e-16 | *** |
| *Mean_MinTemp* | 0.3751707 | < 2e-16 | *** |
| *Percent_All_Forest* | -0.2227624 | < 2e-16 | *** |
| *Percent_PastureHay* | -0.5580507 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | -0.0345101 | 0.448 | None |
| *Percent_WoodyWetlands* | -0.2711129 | 6.43e-06 | *** |
| *Percent_EmHerbWetlands* | 0.0653054 | 0.227 | None |
| *Percent_OpenWater* | -1.1651221 | < 2e-16 | *** |
| *Percent_Barren* | -0.9864771 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 39. Results for Model 15.

| **Model 15** | | | |
|---|---|---|---|
| Dependent variable: duration<br>Model adjusted $R^2$: 0.2151<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 9.3107785 | < 2e-16 | *** |
| *Percent_Developed_High* | 3.1915033 | < 2e-16 | *** |
| *Aspect_median* | 0.0368763 | < 2e-16 | *** |
| *Elevation_range_meters* | 0.0745410 | < 2e-16 | *** |
| *Mean_MaxTemp* | 0.4433159 | < 2e-16 | *** |
| *Percent_All_Forest* | -0.2466935 | < 2e-16 | *** |
| *Percent_PastureHay* | -0.2584371 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | 0.0127846 | 0.783 | None |
| *Percent_WoodyWetlands* | -0.4088155 | 2.75e-11 | *** |
| *Percent_EmHerbWetlands* | 0.0752225 | 0.173 | None |
| *Percent_OpenWater* | -1.0860826 | < 2e-16 | *** |
| *Percent_Barren* | -0.9076648 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 40. Results for Model 15, using the mean minimum temperature.

| **Model 15 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: duration<br>Model adjusted $R^2$: 0.2153<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 16.5486945 | < 2e-16 | *** |
| *Percent_Developed_High* | 3.1365904 | < 2e-16 | *** |
| *Aspect_median* | 0.0358968 | < 2e-16 | *** |
| *Elevation_range_meters* | 0.0739607 | < 2e-16 | *** |
| *Mean_MinTemp* | 0.4918309 | < 2e-16 | *** |
| *Percent_All_Forest* | -0.2499082 | < 2e-16 | *** |
| *Percent_PastureHay* | -0.2586941 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | 0.0450374 | 0.332 | None |
| *Percent_WoodyWetlands* | -0.4097202 | 2.46e-11 | *** |
| *Percent_EmHerbWetlands* | 0.0810333 | 0.142 | None |
| *Percent_OpenWater* | -1.0980253 | < 2e-16 | *** |
| *Percent_Barren* | -0.9668699 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 41. Results for Model 16.

| **Model 16** | | | |
|---|---|---|---|
| Dependent variable: duration<br>Model adjusted $R^2$: 0.2508<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 6.8049009 | < 2e-16 | *** |
| *Percent_Develop_Med_Low_OpnSp* | 0.7189678 | < 2e-16 | *** |
| *Aspect_median* | 0.0159069 | < 2e-16 | *** |
| *Elevation_range_meters* | 0.0769303 | < 2e-16 | *** |
| *Mean_MaxTemp* | 0.5157808 | < 2e-16 | *** |
| *Percent_All_Forest* | -0.2076419 | < 2e-16 | *** |
| *Percent_PastureHay* | -0.5967611 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | -0.0755440 | 0.096 | . |
| *Percent_WoodyWetlands* | -0.2786398 | 3.4e-06 | *** |
| *Percent_EmHerbWetlands* | 0.0727905 | 0.177 | None |
| *Percent_OpenWater* | -1.1667353 | < 2e-16 | *** |
| *Percent_Barren* | -0.9193800 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 42. Results for Model 16, using the mean minimum temperature.

| **Model 16 (*Mean_MinTemp*)** | | | |
|---|---|---|---|
| Dependent variable: duration<br>Model adjusted $R^2$: 0.2501<br>Model p-value: < 2.2e-16 | | | |
| **Independent variables** | **Estimate** | **p-value** | **Signif.** |
| (Intercept) | 17.6957052 | < 2e-16 | *** |
| *Percent_Develop_Med_Low_OpnSp* | 0.7100289 | < 2e-16 | *** |
| *Aspect_median* | 0.0153534 | < 2e-16 | *** |
| *Elevation_range_meters* | 0.0762905 | < 2e-16 | *** |
| *Mean_MinTemp* | 0.3717767 | < 2e-16 | *** |
| *Percent_All_Forest* | -0.2211524 | < 2e-16 | *** |
| *Percent_PastureHay* | -0.5825645 | < 2e-16 | *** |
| *Percent_CultivatedCrops* | -0.0463041 | 0.308 | None |
| *Percent_WoodyWetlands* | -0.2674124 | 8.35e-06 | *** |
| *Percent_EmHerbWetlands* | 0.0653691 | 0.226 | None |
| *Percent_OpenWater* | -1.1587850 | < 2e-16 | *** |
| *Percent_Barren* | -0.9951197 | < 2e-16 | *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**References**

Adams, D. K., & Comrie, A. C. (1997). The North American Monsoon. *Bulletin of the American Meteorological Society, 78,* 2197–2214. https://doi.org/10.1175/1520-0477(1997)078<2197:TNAM>2.0.CO;2

Ahrens, C. D. (2007). *Meteorology today* (8th ed.). Thompson Higher Education.

Ansari, S., Del Greco, S., Kearns E., Brown, O., Wilkins, S., Ramamurthy, M., Weber, J., May, R., Sundwall, J., Layton, J., Gold, A., Pasch, A., & Lakshmanan, V. (2018). Unlocking the potential of NEXRAD data through NOAA's Big Data Partnership. *Bulletin of the American Meteorological Society, 99,* 189–204. https://doi.org/10.1175/BAMS-D-16-0021.1

Armstrong, L., Butler, K., Settelmaier, J., Vance, T., & Wilhelmi, O. (Eds.). (2015). *Mapping and modeling weather and climate with GIS.* ESRI Press.

Ashley, W. S., Bentley, M. L., & Stallins, J. A. (2012). Urban-induced thunderstorm modification in the Southeast United States. *Climatic Change*, *113*(2), 481–498. https://doi.org/10.1007/s10584-011-0324-1

Barry, R. G., & Blanken, P. D. (2016). *Microclimate and local climate*. Cambridge University Press.

Benson, M., Llewellyn, D., Morrison, R., & Stone, M. (2014). Water governance challenges in New Mexico's Middle Rio Grande Valley: A resilience assessment. *Idaho Law Review, 51*(195). http://dx.doi.org/10.2139/ssrn.2464387

Bentley, M. L., Ashley, W. S., & Stallins, J. A. (2010). Climatological radar delineation of urban convection for Atlanta, Georgia. *International Journal of Climatology*, *30*(11), 1589–1594. https://doi.org/10.1002/joc.2020

Bentley, M. L., Stallins, J. A., & Ashley, W. S. (2012). Synoptic environments favourable for urban convection in Atlanta, Georgia. *International Journal of Climatology*, *32*(8), 1287–1294. https://doi.org/10.1002/joc.2344

Bowen, B. M. (1996). Rainfall and climate variation over a sloping New Mexico plateau during the North American Monsoon. *Journal of Climate, 9*, 3432–3442. https://doi.org/10.1175/1520-0442(1996)009<3432:RACVOA>2.0.CO;2

Debbage, N., & Shepherd, J. M. (2015). The urban heat island effect and city contiguity. *Computers, Environment and Urban Systems*, *54*, 181–194. https://doi.org/10.1016/j.compenvurbsys.2015.08.002

Dewitz, J. (2020). *National Land Cover Database (NLCD)*. U.S. Department of the Interior, U.S. Geological Survey, Earth Resources Observation and Science (EROS) Center. https://eros.usgs.gov/doi-remote-sensing-activities/2020/usgs/national-land-cover-database-nlcd

FAA (Federal Aviation Administration). (2016, August 23). *Aviation weather* (Advisory Circular 00-6B). U.S. Department of Transportation. https://www.faa.gov/documentlibrary/media/advisory_circular/ac_00-6b.pdf

Fabry, F. (2015). *Radar meteorology: Principles and practice*. Cambridge University Press.

Fosdick, E. K., & Watson, A. I. (1995). Cloud-to-ground lightning patterns in New Mexico during the summer. *National Weather Digest, 19*(4), 17–24.

Gong, X. (2019). *Spatial regression* [PowerPoint slides].

Haberlie, A. M., Ashley, W. S., & Pingel, T. J. (2015). The effect of urbanisation on the climatology of thunderstorm initiation. *Quarterly Journal of the Royal Meteorological Society*, *141*(688), 663–675. https://doi.org/10.1002/qj.2499

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift, *15*(3), 259–263. https://www.weather.gov/media/jetstream/global/Koppen-Geiger.pdf

Lazzarini, M., Molini, A., Marpu, P. R., Ouarda, T. B. M. J., & Ghedira, H. (2015). Urban climate modifications in hot desert cities: The role of land cover, local climate, and seasonality. *Geophysical Research Letters*, *42*(22), 9980–9989. https://doi.org/10.1002/2015GL066534

López, R. E., Holle, R. L., Watson, A. I., & Skindlov, J. (1997). Spatial and temporal distributions of lightning over Arizona from a power utility perspective. *Journal of Applied Meteorology*, *36*(6), 825–831. https://doi.org/10.1175/1520-0450(1997)036<0825:SATDOL>2.0.CO;2

Maddox, R. A., McCollum, D. M., & Howard, K. W. (2002). Large-scale patterns associated with severe summertime thunderstorms over central Arizona. *Weather and Forecasting*, *10*(4), 763–778. https://doi.org/10.1175/1520-0434(1995)010<0763:lspaws>2.0.co;2

MRLC (Multi-Resolution Land Characteristics Consortium). (2021). *Urban imperviousness*. Multi-Resolution Land Characteristics Consortium. https://www.mrlc.gov/data/type/urban-imperviousness

Myers, J. N. (1964). Preliminary radar climatology of central Pennsylvania. *Journal of Applied Meteorology and Climatology*, *3*(4), 421–429. https://doi.org/10.1175/1520-0450(1964)003<0421:PRCOCP>2.0.CO;2

NCEI (National Centers for Environmental Information). (2021, May 17). *NOAA's Weather and Climate Toolkit*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration. https://www.ncdc.noaa.gov/wct/batch.php

Niyogi, D., Pyle, P., Lei, M., Arya, S. P., Kishtawal, C. M., Shepherd, M., Chen, F., & Wolfe, B. (2011). Urban modification of thunderstorms: An observational storm climatology and model case study for the Indianapolis urban region. *Journal of Applied Meteorology and Climatology*, *50*(5), 1129–1144. https://doi.org/10.1175/2010JAMC1836.1

NOAA (National Oceanic and Atmospheric Administration). (2022a, May 19). *Frequently asked questions (FAQs)*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service Radar Operations Center. https://www.roc.noaa.gov/wsr88d/WindFarm/FAQs.aspx?wid=dev

NOAA (National Oceanic and Atmospheric Administration). (2022b, July 22). *NEXRAD technical information*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service Radar Operations Center, Technical Information. https://www.roc.noaa.gov/WSR88D/Engineering/NEXRADTechInfo.aspx

NOAA (National Oceanic and Atmospheric Administration). (2022c, August 30). *NEXRAD and TDWR radar locations*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service Radar Operations Center. https://www.roc.noaa.gov/WSR88D/Maps.aspx

NOAA (National Oceanic and Atmospheric Administration). (2023, August 11). *Radar images: Reflectivity*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration. https://www.noaa.gov/jetstream/reflectivity

Novo, S., & Raga, G. B. (2013). The properties of convective storms in central Mexico: A radar and lightning approach. *Atmosfera*, *26*(4), 461–472. https://doi.org/10.1016/S0187-6236(13)71088-9

NWS (National Weather Service). (n.d.a). *About our WSR 88-D radar*. National Oceanic and Atmospheric Administration, Weather Forecast Office, Northern Indiana. https://www.weather.gov/iwx/wsr_88d

NWS (National Weather Service). (n.d.b). *NWS weather radar*. National Oceanic and Atmospheric Administration, Weather Forecast Office, Melbourne, FL. https://www.weather.gov/mlb/Doppler_Dual_Pol_Weather_Radar

OFCM (Office of the Federal Coordinator for Meteorological Services and Supporting Research). (2006, February). *Doppler Radar Meteorological Observations*. Federal Meteorological Handbook No. 11, Part D, WSR-88D Unit Description and Operational Applications. FCM-H11B-2006. U.S. Department of Commerce, National Oceanic and Atmospheric Administration. https://www.icams-portal.gov/resources/ofcm/fmh/FMH11/FMH11D-2006.pdf

OFCM (Office of the Federal Coordinator for Meteorological Services and Supporting Research). (2021, July). *WSR-88D Meteorological Observations*. Federal Meteorological Handbook No. 11, Part A, System Concepts, Responsibilities, and Procedures. FCM-H11A-2021. U.S. Department of Commerce, National Oceanic and Atmospheric Administration. https://www.icams-portal.gov/resources/ofcm/fmh/FMH11/2021_fmh11_parta.pdf

Oke, T. R., Mills, G., Christen, A., & Voogt, J. A. (2017). *Urban climates*. Cambridge University Press.

O'Sullivan, D., and Unwin, D. J. (2010). *Geographic information analysis* (2nd ed.). Wiley.

PRISM Climate Group. (2001–2020). *PRISM climate data*. PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, data created between 2001 and 2020.

Seto, K. C., & Shepherd, J. M. (2009). Global urban land-use trends and climate impacts. *Current Opinion in Environmental Sustainability, 1*(1), 89–95. https://doi.org/10.1016/j.cosust.2009.07.012

Shem, W., & Shepherd, M. (2009). On the impact of urbanization on summertime thunderstorms in Atlanta: Two numerical model case studies. *Atmospheric Research*, *92*(2), 172–189. https://doi.org/10.1016/j.atmosres.2008.09.013

Shepherd, J. M. (2006). Evidence of urban-induced precipitation variability in arid climate regimes. *Journal of Arid Environments*, *67*(4), 607–628. https://doi.org/10.1016/j.jaridenv.2006.03.022

Sullivan, S., Kann, D., & Mogil, H. M. (2018). The weather and climate of New Mexico: The land of meteorological enchantment. *Weatherwise, 71*(6), 12–19. https://doi.org/10.1080/00431672.2018.1514181

Vasquez, T. (2015). *Weather radar handbook*. Weather Graphics Technologies.

Wagner, G., Fuelberg, H. E., Kann, D., Wynne, R., & Cobb, S. (2006, January 28–February 2). *A GIS-based approach to lightning studies for West Texas and New Mexico* [Conference presentation]. Second Conference on Meteorological Applications of Lightning Data, Atlanta, GA.

Wallace, J. M., & Hobbs, P. V. (2006). *Atmospheric science: An introductory survey* (2nd ed.). Academic Press.

Watson, A. I., López, R. E., & Holle, R. L. (1994). Diurnal cloud-to-ground lightning patterns in Arizona during the Southwest Monsoon. *Monthly Weather Review, 122*(8), 1716–1725. https://doi.org/10.1175/1520-0493(1994)122<1716:DCTGLP>2.0.CO;2

Williams, A. P., Cook, E. R., Smerdon, J. E., Cook, B. I., Abatzoglou, J. T., Bolles, K., Baek, S. H., Badger, A. M., & Livneh, B. (2020). Large contribution from anthropogenic warming to an emerging North American megadrought. *Science*, *368*(6488), 314–318. https://doi.org/10.1126/science.aaz9600