A Case Study of Travis County's Precipitation Events Inspired by a "Hyperlocal" Approach from NWS and CoCoRaHS Data

by

William W. Plunket

A directed research report submitted to the Geography Department of

Texas State University in partial fulfillment

of the requirements for the degree of

Master of Applied Geography

with a specialization in Resource and Environmental Studies

December 2020

Committee Members:

Dr. Richard Earl - Chair

Dr. Yanan Li – Committee Member

Table of Contents

Introduction

The Balcones Escarpment in Central Texas is often referred to as "Flash Flood Alley" in part due to the topography, intense precipitation events, and other geophysical characteristics which lead to rapid runoff, high stream discharges, and fast on-setting flood events. Travis County's seat, the City of Austin, Texas is one of many municipalities located along U.S. Interstate 35 which roughly follows the Balcones Escarpment through Central Texas. Along the Interstate 35 corridor, municipal development is outpacing water resource planning, one component of which is flood mitigation and management.

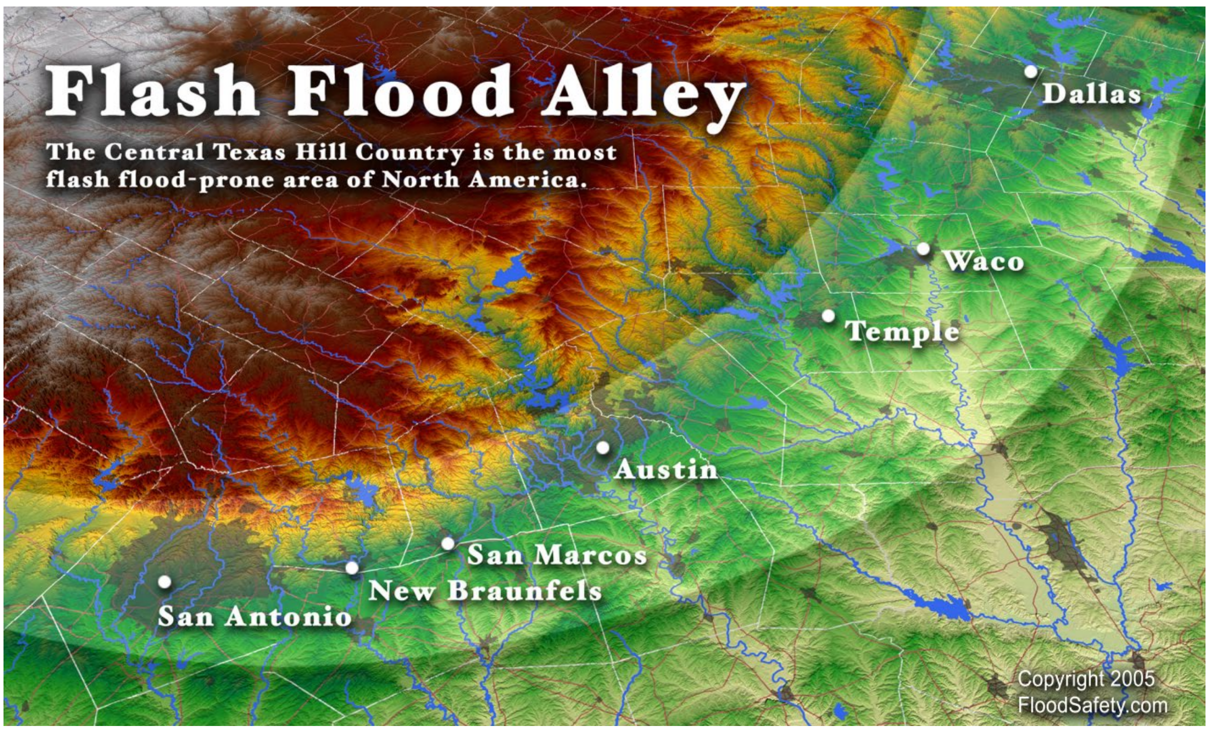


Figure 1: Illustration of the topographic relief contributing to Flash Flood Alley (Frech 2018)

Historically, calculations that yield estimates of precipitation events assume that any given precipitation event is point-based in nature. In practice, precipitation event estimates have been formed from data collected at stations dispersed across the U.S. using long periods of record. These estimates act as design values allowing water resource managers to plan for the future. Combining climatological records and calculated precipitation estimates allows for the creation of recurrence intervals which can help estimate the frequency of occurrence of precipitation events of certain magnitudes. Improvements in understanding the complexity of precipitation events allow for even more beneficial application of these design values.

As many geophysical and meteorological variables including air pressure, humidity, temperature, topographic relief, proximity of major water bodies, ground cover, etc. contribute to how precipitation events manifest across the landscape, further exploration and research of microscale and mesoscale precipitation event characteristics are becoming more prevalent. Floodplain management in Texas is typically conducted at the county level, therefore assessing precipitation estimates across the county is rational. A recent research project through a case study of Fort Collins and Boulder, Colorado has developed a new “hyperlocal” technique to calculate potentially more accurate precipitation estimates that better account for microscale and mesoscale variability. The “hyperlocal” technique developed essentially suggests that precipitation events are heterogeneous across the landscape and that estimates should be derived from numerous data sources rather than individual point sources. The proliferation of volunteered geographic data has increased its reliability and has been proven to aid in data acquisition.

Problem Statement

Basing estimates on the traditional assumption that precipitation events are point-based (i.e. at weather station or gauge scale) fails to account for variability of a precipitation event across the landscape. Failure to account for the areal (relating to the area of interest) diversity of the events can lead to misrepresentations of precipitation event frequencies and subsequently underestimated design values. Decisions made using deficient design values can cause numerous problems such as implementing policies like dam flood operation levels, road closures, etc. or engineering projects that do not adequately mitigate flooding. The treatment of precipitation events as areal phenomena rather than point-based phenomena and subsequent incorporation of estimates derived from the areal perspective should be considered for use as design values to improve the decision-making power for infrastructure development, policy planning, and general water resource management practices in the future.

Purpose Statement

The purpose of this research is to demonstrate *why* volunteered geographic information and a "hyperlocal design value method" like the one developed in Mattingly, Seymour, and Miller (2017), hereafter MSM, should be considered as necessary components for future water resource management plans due to the heterogeneic expression of precipitation events across a region. Volunteered geographic data was obtained from the *Community Collaborative Rain Hail & Snow Network* (CoCoRaHS) stations as well as the two primary Global Historical Climatology Network (GHCN) stations in Travis County. The data was used to find and analyze correlation coefficients between CoCoRaHS and GHCN station observations as they correspond to four historic flood events between 2010 and 2019. Additionally, to emphasize the results of the analysis and contribute to the discussion, visualizations were created in ArcMap that incorporated volunteered geographic information.

Research Question

In a case study of Travis County, Texas, using the largest flood events within the past 10 years on the three largest tributaries of the Colorado River within the county, how did 24hr and 48hr precipitation values recorded by GHCN's Camp Mabry and Austin-Bergstrom International Airport stations correlate to values recorded at CoCoRaHS stations within the county, and what spatial conclusions can be drawn from analysis of these correlations?

Background

Following the devastating floods on the Guadalupe River in 1998 and 2002, PBS raised awareness of the natural flooding hazards of Central Texas through its NOVA documentary, "Flood Alley," released in 2005 (Frech 2010). Since the release of the film there have been several more major flood events that have caused damage far surpassing what many Central Texas residents ever thought possible and all within time frames not previously predicted. Many of these recent floods were products of storms that produced record levels of precipitation within 24hr to 48hr time frames. Four of these major events have been chosen for examination in this work. Three of these events were chosen due to their impact on the three largest streams within Travis County that are tributaries of the Lower Colorado River and their geospatial displacement across the county.

In chronological order, the first event was the peak flow of record within the past 10 years on Bull Creek, recorded September 8, 2010. This event was produced by Tropical Storm Hermine. The second event is the peak flow of record within the past 10 years on Onion Creek, recorded October 31, 2013. The third event was chosen as it represents the peak 24hr precipitation event within the last 10 years for the long term GHCN station at Austin Bergstrom International

Airport (ABIA), one of the two GHCN stations under examination in this work. The ABIA station's 24hr precipitation maximum was recorded October 30, 2015. For the other GHCN station, Camp Mabry, the peak 24hr precipitation event within the last 10 years coincides with Bull Creek's peak flow in 2010. The fourth and last event to be included is the peak flow of record within the past 10 years on Barton Creek, recorded May 4, 2019.

Table 1 shows United States Geological Survey (USGS) stream gage information and observation data for each of the stations top five peak stream flows within the past 10 years. The date of the highest peak flow for each station in this table was used as the date of assessment in this research. Table 2 shows the highest 10 observations of 24hr precipitation for the two GHCN stations within the past 10 years. Table 3 shows 24hr precipitation observations from the two GHCN stations for the days surrounding each of the four events. Figure 2 shows the study area of Travis County, Texas, as well as the location of the relevant stations and gages to demonstrate the spatial distribution. While the inactive CoCoRaHS stations shown in Figure 2 no longer make new observations, their past reports are still useful to study.

TABLE 1 – Major floods on Austin Streams, 2010-2019 (USGS 2020)

USGS Gage	Description	Latitude	Longitude	Drainage Area (sq mi)	Distance from ABIA (mi)	Distance from Camp Mabry (mi)	Date	Gage Height (feet)	Peak Stream Flow (cfs)
08155200	Barton Creek at SH71 near Oak Hill, TX	30.2961	-97.9253	89.7	17	10	**5/4/2019**	**24.24**	**24400**
							10/30/2015	23.93	23000
							5/24/2015	17.83	10300
							9/8/2010	15.77	7560
							10/31/2013	15.6	7380

USGS Gage	Description	Latitude	Longitude	Drainage Area (sq mi)	Distance from ABIA (mi)	Distance from Camp Mabry (mi)	Date	Gage Height (feet)	Peak Stream Flow (cfs)
08154700	Bull Creek at Loop 360 near Austin, TX	30.3719	-97.7844	22.3	14	4	**9/8/2010**	**14.97**	**16900**
							7/3/2016	12.87	11300
							10/31/2013	11.52	8340
							5/25/2016	10.68	6750
							8/15/2012	8.35	3280

USGS Gage	Description	Latitude	Longitude	Drainage Area (sq mi)	Distance from ABIA (mi)	Distance from Camp Mabry (mi)	Date	Gage Height (feet)	Peak Stream Flow (cfs)
08158827	Onion Creek at Twin Creeks Rd near Manchaca, TX	30.1261	-97.8208	181	9	14	**10/31/2013**	**36.88**	**60100**
							10/30/2015	32.36	43200
							5/4/2019	25.49	23100
							5/24/2015	22.02	16700
							9/8/2010	15.33	5480

USGS Gage	Description	Latitude	Longitude	Drainage Area (sq mi)	Distance from ABIA (mi)	Distance from Camp Mabry (mi)	Date	Gage Height (feet)	Peak Stream Flow (cfs)
08159000	Onion Creek at US Hwy 183, Austin, TX	30.1778	-97.6883	321	1	11	**10/31/2013**	**40.13**	**135000**
							10/30/2015	39.33	122000
							5/4/2019	27.11	27000
							5/24/2015	22.93	15100
							9/8/2010	20.37	9540

TABLE 2 – Largest recorded precipitation amounts for two selected Austin NWS/NCEI Stations, 2010-2019 (USNCEI 2020)

ABIA	DATE	24 hour Precipitation (in)
72254013904	10/30/2015	12.49
	5/26/2016	8.79
	8/26/2017	7.47
	3/28/2018	5.99
	1/25/2012	5.66
	10/24/2015	5.16
	5/3/2019	4.77
	7/27/2016	4.18
	9/7/2010	3.54
	12/7/2018	3.49

Camp Mabry	DATE	24 hour Precipitation (in)
72254413958	9/7/2010	7.04
	5/25/2015	5.2
	8/26/2017	5.08
	10/30/2015	4.93
	10/24/2015	4.79
	10/13/2013	4.22
	5/5/2015	3.84
	5/4/2018	3.67
	9/18/2014	3.66
	9/2/2010	3.55

TABLE 3 – Recorded precipitation amounts for two selected Austin NWS/NCEI stations for Sept. 2010, May 2015, Oct. 2015, and Apr. 2019 Austin, TX floods (USNCEI 2020)

	ABIA	72254013904	Camp Mabry	72254413958
	Latitude	Longitude	Latitude	Longitude
	30.1831	-97.6799	30.3208	-97.7604
	Date	24hr Precip (in)	Date	24hr Precip (in)
Bull Creek Peak Flow	9/7/2010	3.54	9/7/2010	7.04
	9/8/2010	**0.78**	**9/8/2010**	**0.51**
	9/9/2010	0	9/9/2010	0
Onion Creek Peak Flow	10/30/2013	0.64	10/30/2013	3.24
	10/31/2013	**2.49**	**10/31/2013**	**1.43**
	11/1/2013	0	11/1/2013	0
ABIA 24hr Precipitation Record	10/29/2015	0	10/29/2015	0
	10/30/2015	**12.49**	**10/30/2015**	**4.93**
	10/31/2015	1.11	10/31/2015	0.89
Barton Creek Peak Flow	5/3/2019	4.77	5/3/2019	3.26
	5/4/2019	**0**	**5/4/2019**	**0**
	5/5/2019	0	5/5/2019	0

**Distance between these two stations is approximately 11 miles (Commerce)*

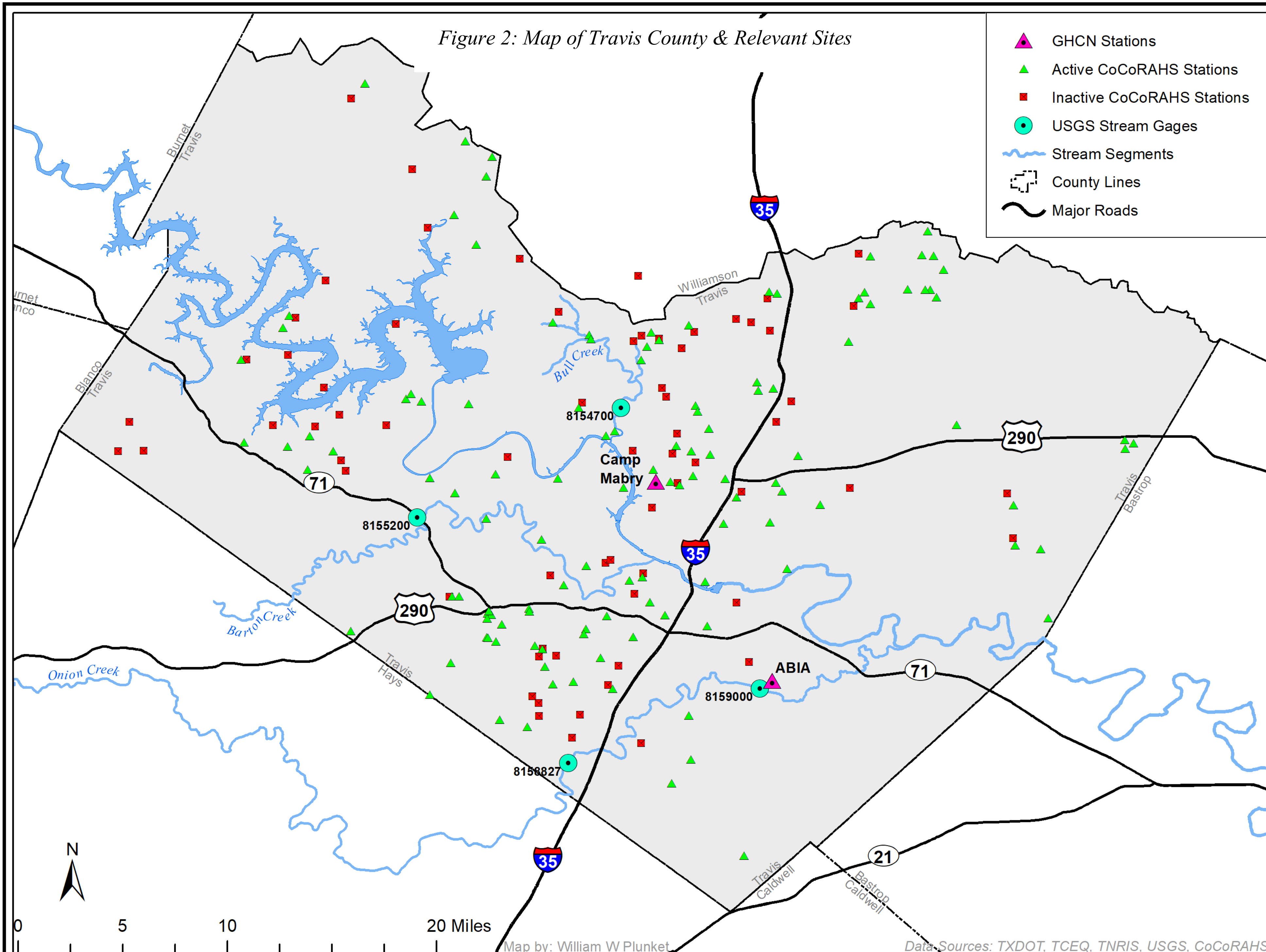


Figure 2: Map of Travis County & Relevant Sites

Literature Review

Importance of Research

Flooding hazards are one of the largest contributors to human casualties and monetary losses related to natural disasters and require ongoing research to better understand the causes and characteristics (Gourley et al. 2012; Brody 2014). Additionally, financial costs related to flood damages have been increasing over the past century (Pielke and Downton 2000; Ashley and Ashley 2008; Changnon 2008; Kundzewicz et al. 2014) further illustrating the need for research related to significant precipitation events and their heterogeneity across a given landscape.

Traditional Methods

Traditional methods used for the establishment of recurrence intervals or "return periods" utilize point-based calculations (Salas et al. 2012). The return periods are used in a wide range of policy decisions as design values for planning including flood mitigation. Gaps in data and general uncertainties are problematic with return periods created from point-based data (Klemeš 2000a, 2000b). Historical statistical techniques such as establishing return periods have been applied for compensation with varying degrees of accuracy (Katz, Parlange, and Naveau 2002; Koutsoyiannis 2004). Extrapolation and interpolation through statistical analysis have traditionally been used as well to account for gaps in the data and to fill in for uncertainties.

Transitional Methods

Increasingly complex mathematical analysis combined with data from regionally homogenous climatic environments or station models themselves have also been used to fill voids in data, however, the most intense precipitation observations are lost through the damping process of regionalization (Hosking and Wallis 2005; Perica et al. 2013). There is increasing support from research that point-based methods underestimate magnitude and recurrence intervals of precipitation events due to assumptions made by traditional methods that precipitation is point-based and uniform across the landscape (Blumenfeld and Skaggs 2011) rather than heterogeneous. Through fine resolution spatial scale studies, research suggests that precipitation events can vary dramatically across the landscape and that point-based records likely fail to capture the most influential, or highest intensity, areas of precipitation (Changnon and Vogel 1981; Winkler 1988; Huff 1994; Villarini, Smith, and Vecchi 2013; Dzotsi et al. 2014; Zolina et al. 2014; Yang et al. 2016).

New Perspectives

Following Blumenfeld and Skaggs (2011), the treatment of precipitation events as non-discrete phenomena has been explored in recent research to assess precipitation events as spatially diverse for the purpose of design value calculations and improve the practicality for real-world decision making (Mattingly, Seymour, and Miller 2017). Research further suggests that increasing spatial density of data collection stations within a shorter period of record derives more accurate precipitation estimates compared to a longer period of record from single stations (Willmott, Robeson, and Janis 1996; Blumenfeld and Skaggs 2011). Spatially expansive estimates are more useful to decision makers and managers as their responsibilities are often regional. The potential use of CoCoRaHS data has been examined and argued to be beneficial for use in design value creation (Keighton et al. 2009; Kelly et al. 2012). Additionally, the use of publicly available data has been shown to be consistent with professionally operated or automated gages (Cifelli et al. 2005; Moon et al. 2009).

While government agency operated data-stations are assumed to follow semi-rigid and formulaic structure and patterns relating to the timing of observations and records, volunteered data of any variety will inherently provide problems and potentially skewed data due to variances in timings of reports. For research that relies on observations within certain time frames, delays in reports are likely to skew data or cause false-positives or negatives in analysis.

Data Background & Coverage

Following MSM (2017), daily observations of precipitation data obtained from the GHCN are best used to assess the traditional precipitation event estimation methods as GHCN stations are rigorously inspected and maintained for quality control purposes (Menne, Durre, Vose, et al. 2012). While the GHCN and CoCoRaHS datasets are publicly accessible, the majority of high-density datasets are not publicly available or are precluded from use due to extreme financial costs required to obtain them (X. Wang et al. 2008; Habib, Larson, and Graschel 2009; Dzotsi et al. 2014; Rafieeinasab et al. 2015). Studies show that CoCoRaHS has recently taken the lead as the most expansive, publicly available, daily precipitation observation system in the United States (Muller et al. 2015).

The GHCN utilizes stations from around the globe that are professionally monitored and calibrated to collect precipitation data on a daily basis as well as many other meteorological variables. The first of two, long standing, Travis County stations that provide data for traditional methods for precipitation estimates is station ID – GHCND:US00013958, “Austin Camp Mabry, TX” located at 30.3208°N, 97.7604°W. This station has been in place and recording daily data since July 1, 1948 with approximately 100% data coverage according to the National Oceanic and Atmospheric Administration's (NOAA), National Centers for Environmental Information (NCEI). The other primary station used by the GHCN is GHCND:US00013904, "Austin Bergstrom International Airport, TX" located at 30.1831°N, 97.6799°W. This station has been in place and recording daily data since January 1, 1949 with approximately 83% data coverage according to NOAA. While the period of record for these stations is substantial, these two stations are roughly 11 miles apart where numerous landscape characteristics that contribute to the onset and intensity of precipitation events vary drastically from topographic relief, to land cover and other physical, geophysical, and environmental variables.

The CoCoRaHS network, an organization that collects daily average, 24-hour, precipitation data voluntarily from the public users associated with the organization, began collecting data in 1998 through the Colorado Climate Center at Colorado State University. Prior significant flash flooding in Fort Collins, Colorado prompted the initiative to establish an enhanced dataset to better understand precipitation characteristics. The state of Texas began contributing data to the network in 2005 and the network now receives data entries from all 50 states of the U.S. Although there are now thousands of collaborators across the network, not every participant records an entry on a daily basis and many regions are slower to develop a quantity of quality collaborators. *Figure 2* illustrates the density of CoCoRaHS users (both active and inactive) in the Travis County as well as indicates where the USGS Gages, GHCN Stations, and case study stream segments are located.

Limitations of Data

As previously acknowledged, the quality of volunteered data can already rival that of professionally operated networks (Keighton et al. 2009; Kelly et al. 2012) and are only increasing in spatial density, in spatial expanse, and in period of record. However, to ensure that the quality of the volunteered data is maintained, research suggests that 80 percent of time recording thresholds are established and that these thresholds are consistent with other climatological studies (Dulière, Zhang, and Salathé 2013). One limitation regarding the data used in this work is related to this threshold. Strictly adhering to this 80 percent threshold would have greatly reduced the number of CoCoRaHS station candidates for inclusion in this study due to the infancy of the CoCoRaHS program in Texas. The effect on the correlations between CoCoRaHS and GHCN stations due to missing observations beyond the aforementioned threshold was identified and included as one of the independent variables used in the regression analysis. Another limitation regarding the methods applied is that the sample size chosen for correlation analysis between two data sources is n=4, the number of major events.

Assumptions of Data

Three considerations were made in this work regarding data due to potential issues related to the timing of and actual reporting of volunteered geographic data. As volunteered geographic data is inherently amateur in nature, guidelines regarding the timing of and reporting of observations are less strict than those of professionally operated stations like the GHCN. The FAQ section on CoCoRaHS's website acknowledges this with the following and demonstrates why considerations must be made:

> **Do I have to check my rain gauge at 7am?** No, but we would prefer it if you did. If you check your gauge at other times, your data may not be directly comparable to other data. If you check your gauge at night, your data will be in our reports but won't show up on our maps. We only map data that is collected within two hours of 7am. (CoCoRaHS)

The three considerations regarding the volunteered geographic data due to the actual date of observation reporting are, A) lag-time between precipitation observation dates and their effect on stream flow gages might exist, B) lag-time between CoCoRaHS observations and GHCN observations might exist, and C) there might exist discrepancies between a 24hr time-frame observation from any given CoCoRaHS station and a 24hr time-frame observation from the GHCN stations.

Research Methods

Methods Background & Outline

This case study of Travis County, Texas was largely inspired by the work conducted in MSM (2017), and its case studies of Fort Collins and Boulder, Colorado. The "hyperlocal" method established in MSM (2017) is a quantitative approach to systematically identify precipitation observations and enhance the creation of estimations, from many spatially dense observations using the CoCoRaHS network within a given representative precipitation region (RPR).

This work's methods differed from those in MSM (2017) due to the difference of purposes between the research. While the methods used in MSM (2017) supported the exploration of developing new precipitation event estimation criteria, the following methods and data used in this case study were intended to examine why the MSM (2017) approach might be necessary. Specifically, the methods used in this research are outlined as follows and explained in more detail in the *Application of Methods* section: 1) determination of rough Pearson correlations between CoCoRaHS and GHCN station observations, 2) identification of a preliminary set of independent variables that potentially impact the correlations, 3) computation of multivariate linear regression analyses in order to assess the independent variables' impact on the correlations between CoCoRaHS and GHCN station observations, and 4) incorporation of CoCoRaHS station observations with GHCN station observations in creating visual representations of 48hr precipitation estimates for each of the four major events.

In MSM (2017), qualified CoCoRaHS station observations "compete" against each other in order to estimate an individual monthly maximum precipitation amount for each month of the period of record. The identified monthly maximums represent the potential for an event of the magnitude to occur somewhere within the RPR. The MSM (2017) "hyperlocal" method uses a circular RPR with a 6 km (3.7 mi) radius and an area of roughly 113 km^2 (43.6 mi^2), similar to that of the 100 km^2 (38.6 mi^2) grid supported by Blumenfeld and Skaggs (2011). In this case study, however, the entirety of Travis County was used as the RPR because it corresponds to the region of responsibility for the local government's flood management. Using the entirety of Travis County yielded and RPR size of approximately 2,650 km^2 (1008 mi^2). Although this RPR is nearly 25 times larger, it was chosen for use as it represents the actual area managed for floods that historically used only two locations of historical data for precipitation estimates within the entire region, Camp Mabry and Austin-Bergstrom International Airport.

Limitations & Assumptions of Methods

Considerable assumptions are made in this work related to the size of the region of study or the RPR, and as mentioned earlier, the reliability of volunteered geographic information. While the literature uses much smaller RPR's compared to this study, it is assumed in this work that the RPR is more significant as it relates to real-world application due to how political boundaries affect decision making. Additionally, for future research, RPR's restricted to certain areal sizes would likely neglect how watershed boundaries actually manifest themselves within a perfect circle and likely lead to the exclusion of observation sites that potentially have real impact on the desired study region.

Reflecting back on the reliability of volunteered geographic information an assumption was made that using statistical interpolation across the entire region, similarly to MSM (2017), would not dampen outlier observations when creating visualizations to demonstrate the heterogeneity of precipitation events across the region. MSM (2017) distinguished between their developed technique and the NWS techniques by describing how the NWS techniques utilize regionalization through interpolation. Interpolation lessens the significance of the most extreme precipitation events at any particular station within a region so as to maintain spatial homogeneity in recurrence interval estimations (Mattingly, Seymour, Miller 2017). While the "hyperlocal" technique also uses regionalization for statistical interpolation, it is applied throughout the entire region of sourced data and does not dampen extreme observations as traditional methods do. The techniques also differ in the use of monthly maximums instead of annual maximums as the period of record for CoCoRaHS would limit the number of available annual observations. MSM (2017) also determined that maximum likelihood estimation was preferred in the field of mathematical statistical inference as opposed to the use of *L*-moments to fit the distributions citing Wackerly, Mendenhall, and Scheaffer (2008) and Casella and Berger (2002).

One limitation regarding the methods used is the acknowledgement that more rigorous and suitable regressions likely exist which would more accurately depict the impact of the independent variables on the correlations. From this limitation, several assumptions were made regarding the use of multivariate linear regression analysis in this work. The first assumption was that a linear relationship exists between the dependent and independent variables. The next assumption made was that there was little to no multicollinearity between the independent variables or rather that the independent variables are not highly correlated with each other. The next assumption made was that there was no autocorrelation. The last assumption made related to the use of multivariate linear regressions was that of homoscedasticity, or rather that there should be roughly equal variance among the predictions made by the produced regression model.

Application of Methods

Following the acquisition of data, the methods previously outlined were applied as detailed below. Table 5A, Table 5B, Table 6A, and Table 6B were created in the first two method steps and used by the third step for analyses.

First step, Pearson correlations between GHCN and CoCoRaHS station observations were determined using greatest 24hr and total 48hr observation values calculated from Table 4. After acquiring the 24hr and 48hr precipitation values for each station and the associated, Pearson correlation coefficients were determined in Excel between each GHCN and CoCoRaHS station.

The resulting correlation coefficients are represented in the dependent variable columns for ABIA in Table 5A and Table 5B, and for Camp Mabry in Table 6A and Table 6B.

Second step, preliminary independent variables were identified for use in regression analysis. It was acknowledged that numerous variables exist that potentially have more significant impact on station correlations, but the following independent variables were used as a seemingly decent starting point, the first independent variable chosen was the distance between each GHCN station and the CoCoRaHS stations. The Euclidean distances between the stations were calculated using ArcMap and their results added as the first independent variable column in Table 5A, Table 5B, Table 6A, and Table 6B. The second independent variable chosen was the elevation of each station. The station elevations were included with original data acquisition and their values entered as the second independent variable column in Table 5A, Table 5B, Table 6A, and Table 6B. The third and last independent variable chosen for regression analysis in this work was the *completeness of reports*, or rather the percentage of days reported out of the 16 possible days of which data was collected. The *null* observations present in Table 4 were used to determine the ratios of *completeness of reports* and are added as the third independent variable column in Table 5A, Table 5B, Table 6A, and Table 6B.

Third step, multivariate linear regressions between the GHCN correlations and the independent variables were run using Excel. The interpretation of the regressions summaries is explored in the following *Results* section.

For ABIA, the multivariate linear regression with three independent variables was performed for the 24hr and 48hr observations and their resulting summaries are shown in Table 7A and Table 7B respectively. As seen in Table 7A and Table 7B, the p-values for each independent variable indicate that none were statistically significant and thus no further regressions were performed.

For Camp Mabry, the multivariate linear regression with three independent variables was performed for the 24hr and 48hr observations and their resulting summaries are shown in Table 8A and Table 8B respectively. As seen in both Table 8A and Table 8B, the p-values for *Report Completeness* were not statistically significant while *Distance to Mabry* and *Station Elevations* were statistically significant. Following these result, in attempts to improve the ability to explain the variation in the correlations, another multivariate linear regression with the two significant independent variables was performed for each of the 24hr and 48hr observations and their resulting summaries are shown in Table 9A and Table 9B respectively. For both new regressions, as both p-values associated with the two independent variables indicated some level of statistical significance, no further regressions were performed.

Last, using ArcMap's Spatial Analyst Tools (SAT), visualizations of precipitation heterogeneity were created from CoCoRaHS and GHCN observations for each of the four major events. The ArcMap SAT Kriging tool used to create the visualizations is a geostatistical interpolation process which attempts to predict missing values across space utilizing given values and variances. Whereas only 18 CoCoRaHS stations were used in the correlations' determination and regressions' analyses, additional CoCoRaHS stations were included to create these visualizations. As a visualization was desired for each of the major events, CoCoRaHS stations were included if their observations met the same stipulations for a 48hr window but did not need to have reported observations for each of the events. The rationale in this decision was that lack of reporting for the missing events likely mean that the station had not been established prior to the missing event or became defunct following the missing event. Additionally, by maintaining the stipulations to qualify for a 48hr event, some data integrity was maintained while allowing many more CoCoRaHS station observations to contribute to and likely improve the interpolation process used to create the visualizations of each event. The resulting visualizations are included and discussed in the following *Results* section.

Results & Analysis

From these considerations of the uncertainty of the time of observation for CoCoRaHS stations, it was decided to use both the greatest 24hr value and the sum of two consecutive 24hr observations (for a total of 48hrs), where the later of the two 24hr observations was the highest reported value within each window around the event. This decision was based on an assumption that using both the 24hr and 48hr observation with the mentioned stipulation would alleviate discrepancies between CoCoRaHS and GHCN station observations due to the differences in observation times. An assumption was also made that due to the professionality associated with the GHCN stations, using the two 24hr reported observations for the date of the event and the date prior would suffice for use as their 48hr observation values.

Data Acquisition

Using ESRI's ArcMap GIS software, the RPR was drawn and used to determine which CoCoRaHS stations fell within the desired area. Station data and geographical information were obtained through the online NOAA data portals for the GHCN, and the online CoCoRaHS database. For inclusion in the correlation determination and regression analyses, CoCoRaHS stations that reported 48hr observations (as stipulated previously in the assumptions section) for each of the four events were included and their precipitation observations are shown in Table 4.

TABLE 4 - GHCN & CoCoRAHS - 24hr Precipitation Observations by Event (inches)																
	Bull Creek Peak Flow Event				Onion Creek Peak Flow Event				ABIA 24hr Precip Max Event				Barton Creek Peak Flow Event			
	9/6/2010	9/7/2010	**9/8/2010**	9/9/2010	10/29/2013	10/30/2013	**10/31/2013**	11/1/2013	10/28/2015	10/29/2015	**10/30/2015**	10/31/2015	5/2/2019	5/3/2019	**5/4/2019**	5/5/2019
ABIA	0.02	3.54	0.78	0.00	0.00	0.64	2.49	0.00	0.00	0.00	12.49	1.11	0.02	4.77	0.00	0.00
Mabry	0.02	7.04	0.51	0.00	0.00	3.24	1.43	0.00	0.00	0.00	4.93	0.89	0.15	3.26	0.00	0.00
TX_TV_1	0.00	0.65	10.14	0.58	0.00	0.10	8.78	0.00	0.00	0.00	0.46	4.80	0.15	0.34	1.07	0.00
TX_TV_2	*null*	1.03	6.68	0.06	*null*	0.15	5.29	0.00	*null*	*null*	0.98	4.43	0.11	0.54	2.06	*null*
TX_TV_9	0.00	0.50	10.58	0.61	0.00	0.10	8.22	0.01	0.00	0.00	0.89	5.51	0.14	0.28	3.76	0.00
TX_TV_10	0.00	1.00	6.69	0.06	*null*	0.23	2.76	*null*	0.00	0.00	0.84	3.67	0.07	0.41	2.68	*null*
TX_TV_27	0.00	0.14	11.01	0.82	0.00	0.01	1.22	0.00	0.00	0.00	0.70	4.45	0.10	1.00	1.20	0.00
TX_TV_30	0.06	0.80	10.80	0.30	0.00	0.11	7.74	0.00	0.00	0.00	1.92	4.82	0.10	0.37	1.18	0.00
TX_TV_34	0.00	0.37	7.35	0.08	0.00	0.04	5.00	0.00	0.00	0.00	0.66	8.13	0.01	0.25	4.15	0.00
TX_TV_44	0.00	0.34	6.54	0.05	0.00	0.13	4.04	0.02	0.00	0.00	0.17	5.39	0.09	0.48	2.93	0.00
TX_TV_52	0.00	0.34	7.52	0.49	0.00	0.14	5.96	0.01	0.00	0.00	0.92	4.40	0.08	0.20	5.42	0.00
TX_TV_87	0.00	0.63	8.39	0.08	0.00	0.17	4.61	0.00	0.00	0.00	0.42	5.67	0.03	0.62	1.75	0.00
TX_TV_96	0.00	0.36	7.36	0.05	0.00	0.00	5.70	0.00	0.00	0.00	1.22	6.88	0.06	0.42	3.25	0.00
TX_TV_99	0.00	0.56	7.60	0.87	0.00	0.00	6.45	0.00	0.00	0.00	2.66	1.57	0.11	0.52	3.44	*null*
TX_TV_113	0.00	0.30	7.16	0.16	*null*	0.08	6.18	*null*	*null*	*null*	0.60	6.26	0.06	0.25	3.80	0.00
TX_TV_114	0.00	0.33	8.06	0.61	0.00	0.01	6.26	*null*	*null*	*null*	0.65	5.27	*null*	0.24	5.75	*null*
TX_TV_117	*null*	0.16	11.24	0.45	0.00	0.08	8.98	0.00	0.00	0.00	1.62	4.45	0.17	0.35	1.86	0.00
TX_TV_118	0.00	0.57	4.50	0.08	0.00	0.16	3.18	0.00	0.00	0.00	0.53	5.23	0.08	0.72	5.08	*null*
TX_TV_122	0.00	0.70	9.72	0.69	*null*	0.12	6.31	0.00	0.00	0.00	0.83	5.10	0.04	0.25	4.48	0.00
TX_TV_123	*null*	0.53	9.58	0.69	0.00	0.13	9.39	0.01	*null*	*null*	0.58	5.14	0.19	0.43	0.99	*null*

Results

Following the methods previously set forth, the correlation coefficients are included as follows where table references with A's are for 24hr precipitation periods and table references with B's are for 48hr precipitation periods. Table 5A, Table 5B, Table 6A, and Table 6B while the regression summaries are included for ABIA in Table 7A and Table 7B, and for Camp Mabry in Table 8A, Table 8B, Table 9A and Table 9B. The visualizations created using 48hr observations for the Bull Creek event, Onion Creek event, ABIA event, and Barton Creek event, are represented by Figure 3, Figure 4, Figure 5, and Figure 6 respectively.

Table 5A and Table 5B represent the data used for the multivariate regression analysis conducted between the Pearson correlation coefficients (the dependent variable), and the three independent variables chosen for this study, the distance between each station and ABIA, the elevation of each station, and the completeness of reports where *n* is out of 16 possible days to report. The calculated Pearson coefficients between ABIA and the other stations interestingly resulted with nearly half of the correlations as negative with the majority of them being relatively weak.

Table 6A and Table 6B represent the data used for the multivariate regression analysis conducted between the Pearson correlation coefficients (the dependent variable), and the three independent variables chosen for this study, the distance between each station and Camp Mabry, the elevation of each station, and the completeness of reports where *n* is out of 16 possible days to report. The calculated Pearson coefficients between Camp Mabry and the other stations were all positive and the majority of them relatively strong.

TABLE 5A - ABIA -24hr - Multivariate Linear Regression Components				
Station	Dependent Variable	Independent Variables		
	Correlation to ABIA	Distance to ABIA (miles)	Station Elevation (ft)	Completeness of Reports (n=16)
ABIA	1.00	0.000	480	1.0
Mabry	0.105	11.0	670	1.0
TX_TV_1	-0.385	17.1	855	1.0
TX_TV_2	-0.208	12.1	689	0.55
TX_TV_9	-0.459	14.1	782	1.0
TX_TV_10	-0.110	11.6	721	0.77
TX_TV_27	-0.013	29.7	1066	1.0
TX_TV_30	-0.349	18.7	848	1.0
TX_TV_34	0.638	9.25	668	1.0
TX_TV_44	0.208	10.7	646	1.0
TX_TV_52	-0.755	14.0	959	1.0
TX_TV_87	0.039	12.4	752	1.0
TX_TV_96	0.266	7.63	701	1.0
TX_TV_99	-0.768	19.0	805	0.93
TX_TV_113	0.041	11.5	752	0.67
TX_TV_114	-0.612	13.8	932	0.55
TX_TV_117	-0.481	16.0	714	0.93
TX_TV_118	0.681	9.48	548	0.93
TX_TV_122	-0.437	12.9	790	0.93
TX_TV_123	-0.359	17.2	865	0.67

TABLE 5B - ABIA -48hr - Multivariate Linear Regression Components				
Station	Dependent Variable	Independent Variables		
	Correlation to ABIA	Distance to ABIA (miles)	Station Elevation (ft)	Completeness of Reports (n=16)
ABIA	1.00	0.000	480	1.0
Mabry	0.164	11.0	670	1.0
TX_TV_1	-0.292	17.1	855	1.0
TX_TV_2	-0.011	12.1	689	0.55
TX_TV_9	-0.291	14.1	782	1.0
TX_TV_10	0.020	11.6	721	0.77
TX_TV_27	0.053	29.7	1066	1.0
TX_TV_30	-0.091	18.7	848	1.0
TX_TV_34	0.728	9.25	668	1.0
TX_TV_44	0.244	10.7	646	1.0
TX_TV_52	-0.508	14.0	959	1.0
TX_TV_87	0.108	12.4	752	1.0
TX_TV_96	0.549	7.63	701	1.0
TX_TV_99	-0.525	19.0	805	0.93
TX_TV_113	0.252	11.5	752	0.67
TX_TV_114	-0.378	13.8	932	0.55
TX_TV_117	-0.536	16.0	714	0.93
TX_TV_118	0.570	9.48	548	0.93
TX_TV_122	-0.260	12.9	790	0.93
TX_TV_123	-0.259	17.2	865	0.67

TABLE 6A - Camp Mabry - 24hr Multivariate Linear Regression Components				
Station	Dependent Variable	Independent Variables		
	Correlation to Mabry	Distance to Mabry (miles)	Station Elevation (ft)	Completeness of Reports (n=16)
ABIA	0.105	11.0	480	1.0
Mabry	1.00	0.000	670	1.0
TX_TV_1	0.566	6.49	855	1.0
TX_TV_2	0.715	1.98	689	0.55
TX_TV_9	0.675	4.68	782	1.0
TX_TV_10	0.977	0.579	721	0.77
TX_TV_27	0.993	18.7	1066	1.0
TX_TV_30	0.695	7.71	848	1.0
TX_TV_34	0.768	7.81	668	1.0
TX_TV_44	0.940	0.705	646	1.0
TX_TV_52	0.573	10.2	959	1.0
TX_TV_87	0.903	3.59	752	1.0
TX_TV_96	0.779	10.1	701	1.0
TX_TV_99	0.446	10.8	805	0.93
TX_TV_113	0.726	9.76	752	0.67
TX_TV_114	0.721	10.1	932	0.55
TX_TV_117	0.598	5.10	714	0.93
TX_TV_118	0.250	5.76	548	0.93
TX_TV_122	0.821	6.11	790	0.93
TX_TV_123	0.479	6.79	865	0.67

TABLE 6B - Camp Mabry - 48hr Multivariate Linear Regression Components				
Station	Dependent Variable	Independent Variables		
	Correlation to Mabry	Distance to Mabry (miles)	Station Elevation (ft)	Completeness of Reports (n=16)
ABIA	0.164	11.0	480	1.0
Mabry	1.00	0.000	670	1.0
TX_TV_1	0.782	6.49	855	1.0
TX_TV_2	0.951	1.98	689	0.55
TX_TV_9	0.861	4.68	782	1.0
TX_TV_10	0.932	0.579	721	0.77
TX_TV_27	0.914	18.7	1066	1.0
TX_TV_30	0.913	7.71	848	1.0
TX_TV_34	0.795	7.81	668	1.0
TX_TV_44	0.994	0.705	646	1.0
TX_TV_52	0.746	10.2	959	1.0
TX_TV_87	0.997	3.59	752	1.0
TX_TV_96	0.879	10.1	701	1.0
TX_TV_99	0.741	10.8	805	0.93
TX_TV_113	0.919	9.76	752	0.67
TX_TV_114	0.813	10.1	932	0.55
TX_TV_117	0.732	5.10	714	0.93
TX_TV_118	0.015	5.76	548	0.93
TX_TV_122	0.908	6.11	790	0.93
TX_TV_123	0.733	6.79	865	0.67

Table 7A, Table 8A, and Table 9A show the 24hr precipitation period, multivariate linear regression summaries for ABIA with three independent variables, Camp Mabry with three independent variables, and Camp Mabry with two independent variables, respectively. The values in these summaries were used to determine the impact each of the independent variables chosen had on the correlations between the GHCN stations and the CoCoRaHS stations. The first value examined in each of the tables is the adjusted R square value as it represents the approximate percentage of impact the collective independent variables chosen estimated to have on the correlations. The standard error shown under each adjusted R square value indicates the potential error in the estimation when using the predictions. The next value of import is the significance-F in each of the tables as it indicates the probability of the regression model to be incorrect. The significance-F value for the regression performed for ABIA suggests a significance level over 99%, while the significance-F value for the second regression performed for Camp Mabry suggests a significance level over 95%.

Similar to the significance-F values, the p-values for each independent variable indicate the probability that the calculated coefficients within the regression are incorrect. For ABIA's related regression, while the significance-F suggests the model is fairly reliable, the p-values for each of the independent variables have poor significance levels, the highest being that related to Station Elevation just above 0.05. For Camp Mabry, examining the p-values in Table 8A demonstrate why another multivariate linear regression was performed. While the p-values in Table 8A suggest significance levels smaller than 0.05 for both the Distance to Mabry and the Station Elevation, the Report Completeness significance level was very low and thus removed for the next regression in order to obtain a more accurate adjusted R square value.

Continuing with Table 9A, Camp Mabry's multivariate regression performed with only two independent variables, the new significance-F value suggested that the regression model improved from that of the prior. Conversely, the p-values for the two independent variables degraded slightly yet still suggested significance levels above 95% for both Distance to Mabry and Station Elevation.

Table 7A - ABIA - 24hr Multivariate Linear Regression Summary

Three Independent Variables

Regression Statistics	
Multiple R	0.74
R Square	0.55
Adjusted R Square	0.46
Standard Error	0.35
Observations	20

ANOVA	df	SS	MS	F	Sig. F
Regression	3	2.4	0.81	6.4	0.0046
Residual	16	2.0	0.13		
Total	19	4.4			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	1.25	0.82	1.5	0.15	-0.49	3.0
Distance to ABIA	-0.00569	0.0260	-0.219	0.830	-0.0608	0.0494
Station Elevations	-0.00226	0.00110	-2.06	0.0562	-0.00459	0.0000670
Report Completeness	0.51	0.52	0.98	0.34	-0.59	1.6

TABLE 8A - Camp Mabry - 24hr Multivariate Linear Regression Summary

Three Independent Variables

Regression Statistics	
Multiple R	0.61
R Square	0.37
Adjusted R Square	0.25
Standard Error	0.21
Observations	20

ANOVA	df	SS	MS	F	Sig. F
Regression	3	0.40	0.13	3.1	0.055
Residual	16	0.69	0.043		
Total	19	1.1			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	-0.14	0.42	-0.33	0.75	-1.0	0.75
Distance to Mabry	-0.0313	0.0123	-2.54	0.0219	-0.0575	-0.00518
Station Elevations	0.00112	0.000405	2.77	0.0136	0.000264	0.00198
Report Completeness	0.208	0.30	0.69	0.50	-0.43	0.85

TABLE 9A - Camp Mabry - 24hr Multivariate Linear Regression Summary

Two Independent Variables

Regression Statistics	
Multiple R	0.593
R Square	0.352
Adjusted R Square	0.275
Standard Error	0.204
Observations	20

ANOVA	df	SS	MS	F	Sig. F
Regression	2	0.383	0.192	4.61	0.0252
Residual	17	0.707	0.042		
Total	19	1.09			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0.0793	0.271	0.293	0.773	-0.492	0.651
Distance to Mabry	-0.0298	0.0119	-2.49	0.0232	-0.0550	-0.00460
Station Elevations	0.00107	0.000390	2.73	0.0142	0.000243	0.00189

Table 7B, Table 8B, and Table 9B show the 48hr precipitation period, multivariate linear regression summaries for ABIA with three independent variables, Camp Mabry with three independent variables, and Camp Mabry with two independent variables, respectively. The values in these summaries were used to determine the impact each of the independent variables chosen had on the correlations between the GHCN stations and the CoCoRaHS stations. The first value examined in each of the tables is the adjusted R square value as it represents the approximate percentage of impact the collective independent variables chosen are estimated to have on the correlations. The standard error shown under each adjusted R square value indicates the potential error in the estimation when using the predictions. The next value of import is the significance-F in each of the tables as it indicates the probability of the regression model to be incorrect. Each of the significance-F values for the regressions suggest the regression models are accurate with significance levels equal to or above 99%.

Similar to the significance-F values, the p-values for each independent variable indicate the probability that the calculated coefficients within the regression are incorrect. For ABIA's related regression, while the significance-F suggests the model is fairly reliable, the p-values for each of the independent variables have poor significance levels, the highest being that related to Station Elevation at roughly 80%. For Camp Mabry, examining the p-values in Table 8B demonstrate why another multivariate linear regression was performed. While the p-values in Table 8B suggest significance-For the Distance to Mabry and the Station Elevation, the Report Completeness significance level was very low and thus removed for the next regression in order to obtain a more accurate adjusted R square value.

Continuing with Table 9B, Camp Mabry's multivariate regression performed with only two independent variables, the new significance-F value suggested that the regression model greatly improved from that of the prior. Additionally, the p-values for the two independent variables improved slightly yet still suggested significance levels above 99% for both Distance to Mabry and Station Elevation.

TABLE 7B - ABIA - 48hr Multivariate Linear Regression Summary	
Three Independent Variables	

Regression Statistics	
Multiple R	0.71
R Square	0.50
Adjusted R Square	0.41
Standard Error	0.33
Observations	20

ANOVA	df	SS	MS	F	Sig. F
Regression	3	1.8	0.60	5.3	0.0096
Residual	16	1.8	0.11		
Total	19	3.6			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0.98	0.77	1.3	0.22	-0.66	2.6
Distance to ABIA	-0.0200	0.0246	-0.815	0.427	-0.0721	0.0321
Station Elevations	-0.00139	0.00104	-1.34	0.198	-0.00359	0.000806
Report Completeness	0.42	0.49	0.86	0.40	-0.62	1.5

TABLE 8B - Camp Mabry - 48hr Multivariate Linear Regression Summary	
Three Independent Variables	

Regression Statistics	
Multiple R	0.71
R Square	0.50
Adjusted R Square	0.40
Standard Error	0.20
Observations	20

ANOVA	df	SS	MS	F	Sig. F
Regression	3	0.62	0.21	5.3	0.010
Residual	16	0.63	0.039		
Total	19	1.3			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	-0.13	0.40	-0.32	0.75	-0.98	0.72
Distance to Mabry	-0.0349	0.0118	-2.95	0.00937	-0.0599	-0.00983
Station Elevations	0.00145	0.000388	3.73	0.00181	0.000625	0.00227
Report Completeness	0.064	0.29	0.22	0.83	-0.55	0.68

TABLE 9B - Camp Mabry - 48hr Multivariate Linear Regression Summary	
Two Independent Variables	

Regression Statistics	
Multiple R	0.704
R Square	0.496
Adjusted R Square	0.437
Standard Error	0.193
Observations	20

ANOVA	df	SS	MS	F	Sig. F
Regression	2	0.622	0.311	8.37	0.00295
Residual	17	0.632	0.037		
Total	19	1.25			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	-0.0627	0.256	-0.245	0.809	-0.603	0.478
Distance to Mabry	-0.0344	0.0113	-3.05	0.00728	-0.0583	-0.0106
Station Elevations	0.00143	0.000369	3.87	0.00122	0.000651	0.00221

The visualizations represented in Figure 3, Figure 4, Figure 5, and Figure 6, were created to represent how each of the four events precipitations varied spatially and how the incorporation of CoCoRaHS stations improved the level at which the variation could be conveyed. Whereas the previous correlations and regressions used only 18 CoCoRaHS stations due to stipulations discussed previously, each of the following visualizations used any CoCoRaHS station that provided observations related to the events, also stipulated previously. The interpolations performed in ArcMap provide more accurate predictions with more data points provided, thus the incorporation of more CoCoRaHS stations improved the visualizations. Chronologically, the visualization for Bull Creek's event benefitted from 54 total stations, 52 CoCoRaHS and the two GHCN, the visualization for Onion Creek's event benefitted from 70 total stations, 68 CoCoRaHS and the two GHCN, the visualization for the ABIA record 24hr precipitation benefitted from 70 total stations, 68 CoCoRaHS and the two from GHCN, and lastly the visualization for Barton Creek's event benefitted from 64 total stations, 62 from CoCoRaHS and the two GHCN.

The visualizations show the location of the station observations used to create each visual, the interpolated precipitation across the region represented by the varying color bands, and the proportional peak stream flow for the three USGS gages related to the event, one on Bull Creek, one on Onion Creek, and the last on Barton Creek. While the gage number was left in place for reference, the USGS gage for Bull Creek did not have peak flows ranking in the top ten for the last ten years related to the ABIA and Barton Creek events and thus no proportional representation was shown.

Figure 3: Visualization of Interpolated 48hr Precipitation from Station Reports

Bull Creek Event – 9/8/2010 – [n=54 stations reported]

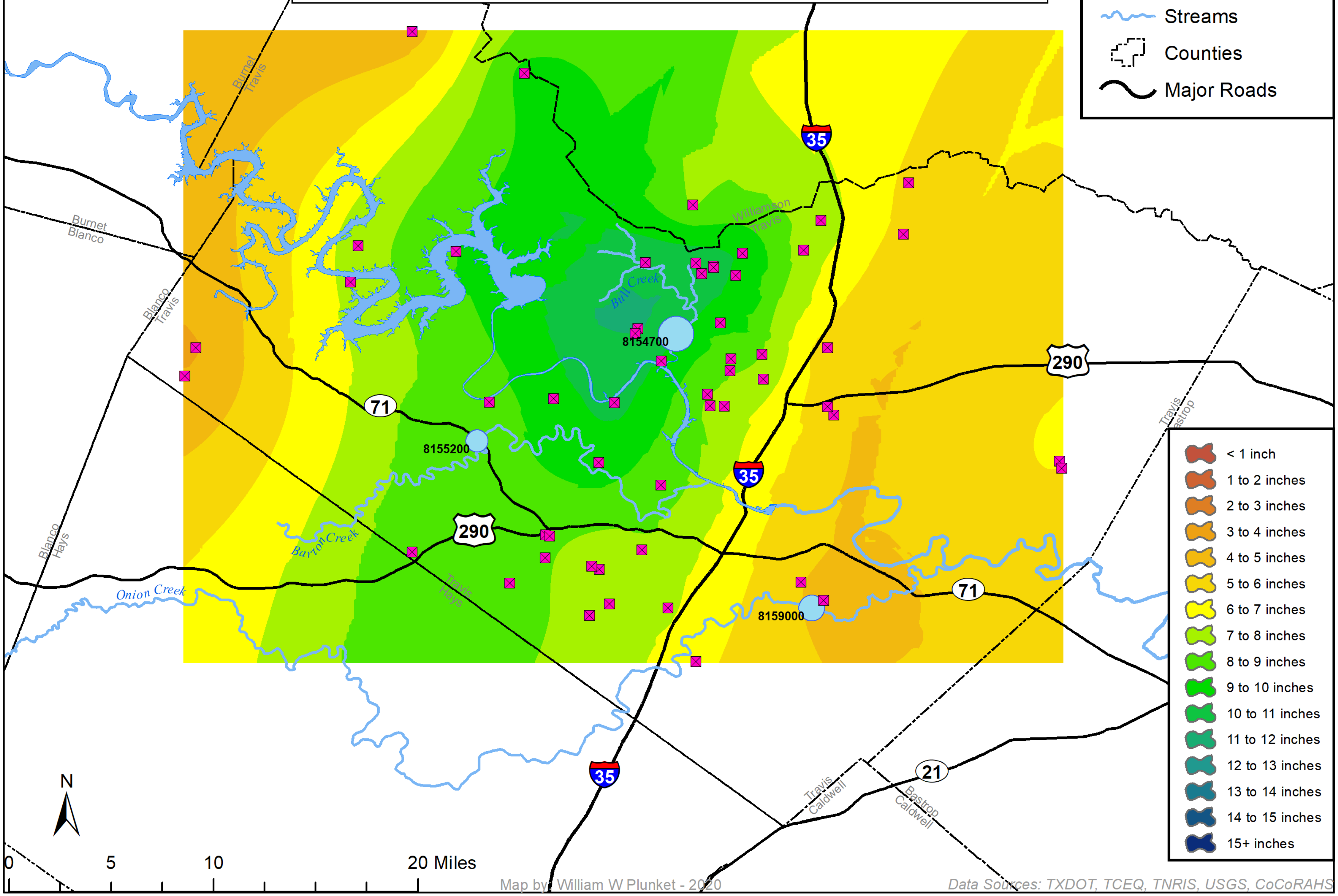
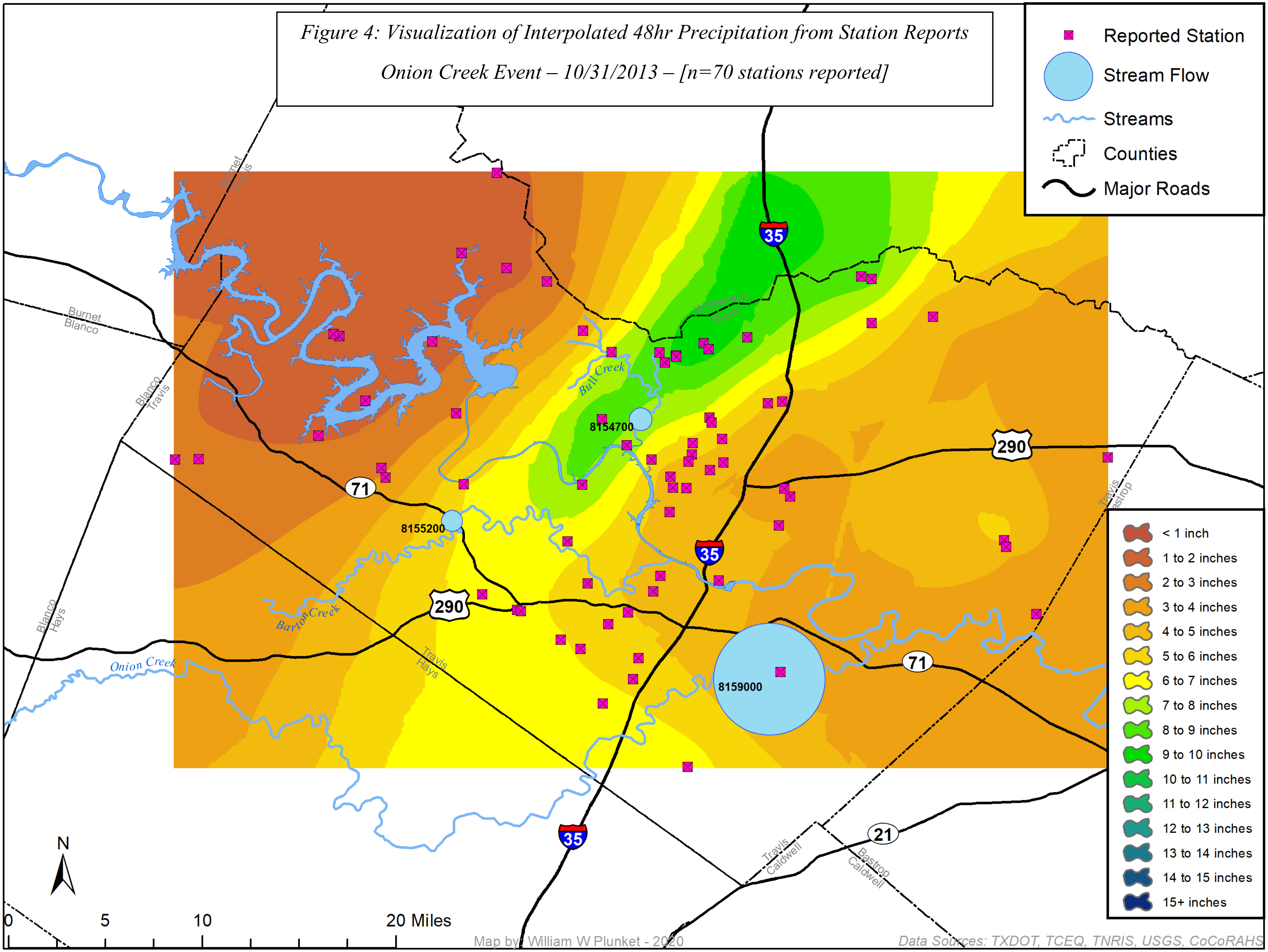


Figure 4: Visualization of Interpolated 48hr Precipitation from Station Reports

Onion Creek Event – 10/31/2013 – [n=70 stations reported]



Map by: William W Plunket - 2020

Data Sources: TXDOT, TCEQ, TNRIS, USGS, CoCoRAHS

Figure 5: Visualization of Interpolated 48hr Precipitation from Station Reports

ABIA 24hr Precip Max Event – 10/30/2015 – [n=70 stations reported]

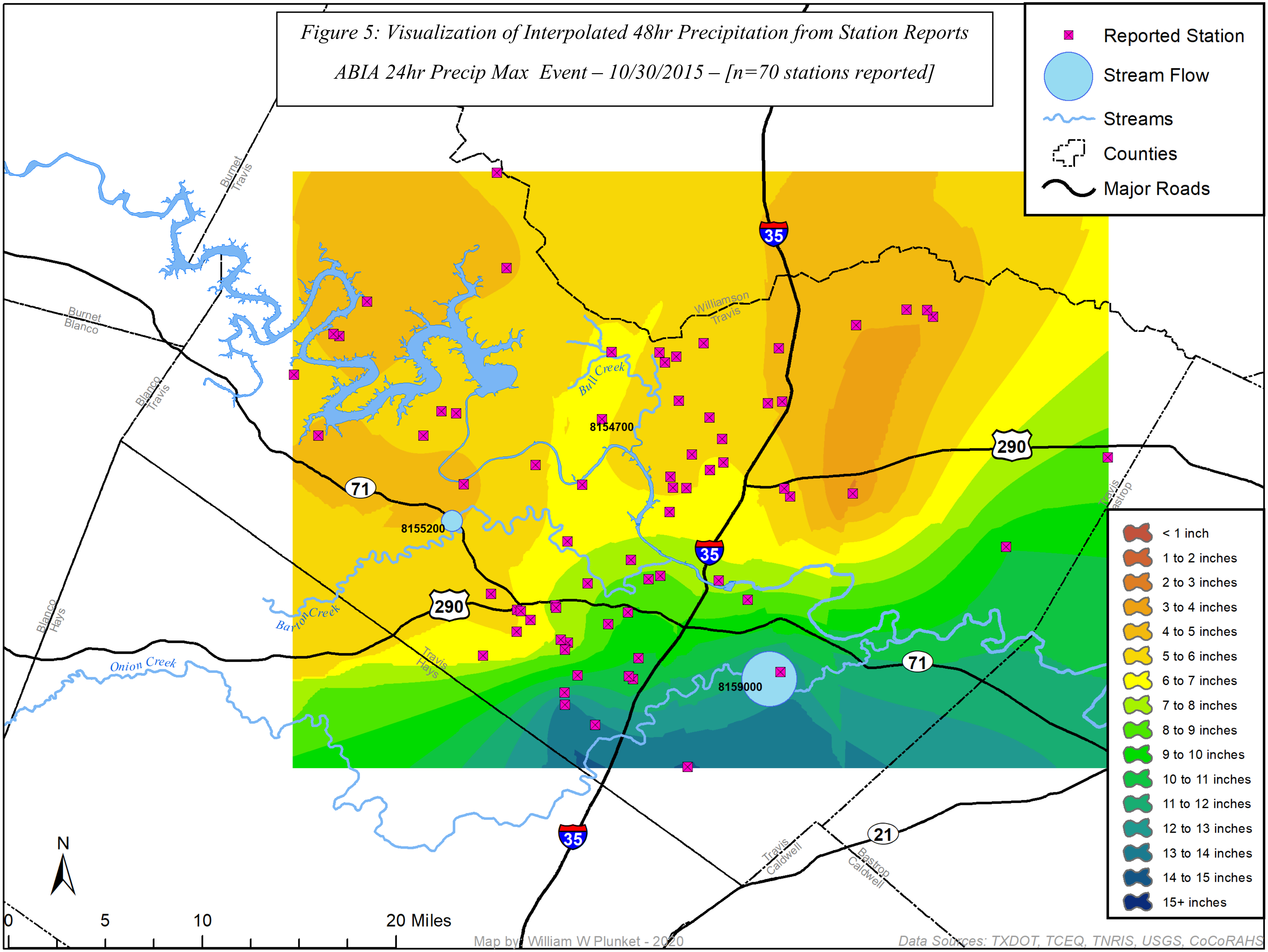
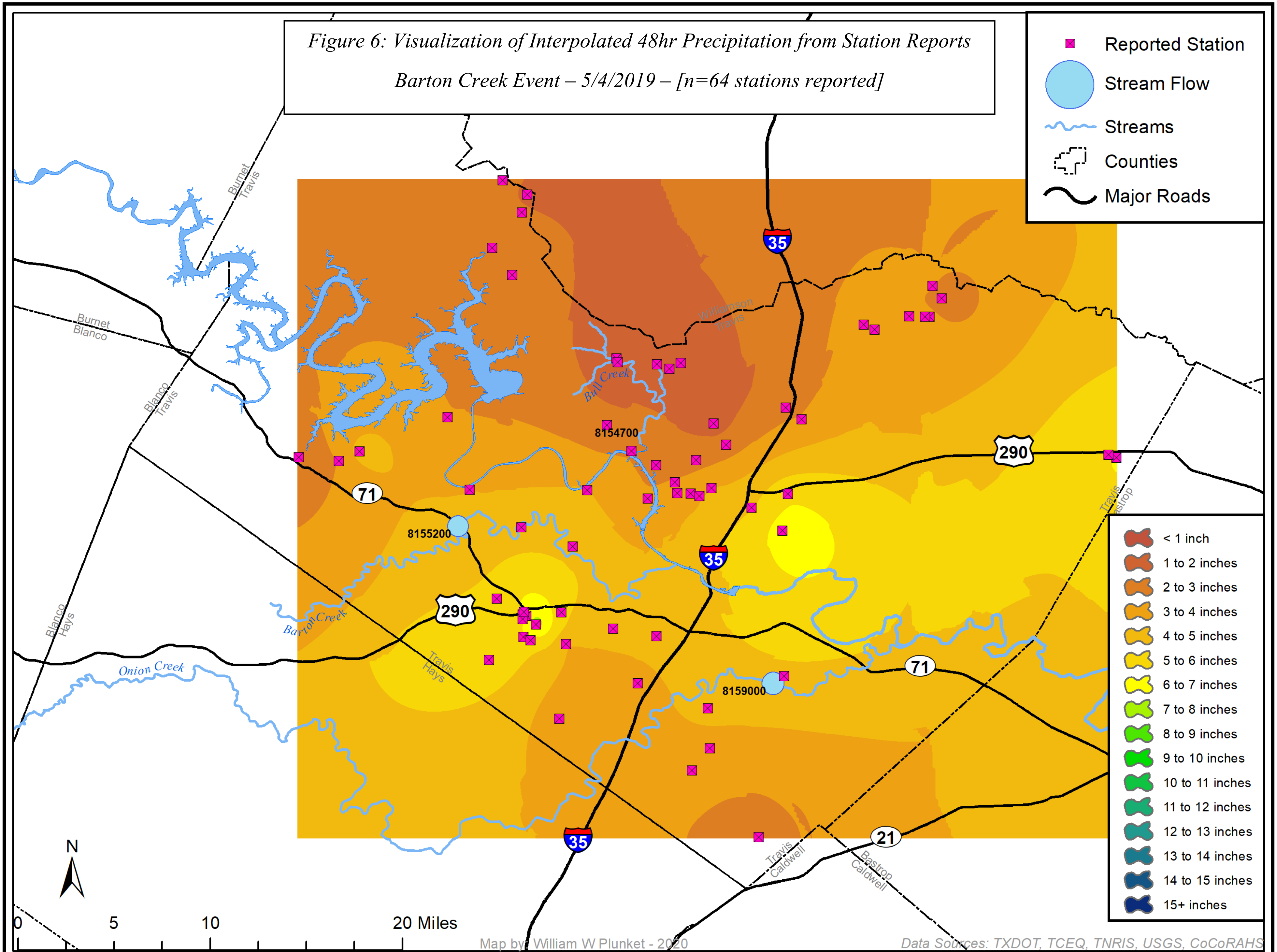


Figure 6: Visualization of Interpolated 48hr Precipitation from Station Reports

Barton Creek Event – 5/4/2019 – [n=64 stations reported]



Discussion

How did the CoCoRaHS observations correlate with the GHCN observations?

Upon initial review of the resulting correlation coefficients between ABIA and the other stations as shown in Table 5A and Table 5B, and before examining the regression analysis results, approximately half of the stations exhibited a negative correlation with ABIA's observations with the majority of the stations having relatively weak correlation magnitudes with ABIA. Unlike ABIA, upon initial review of the resulting correlation coefficients between Camp Mabry and the other stations as shown in Table 6A and Table 6B, and before examining the regression analysis results, the all of the stations exhibited a positive correlation with Camp Mabry's observations with nearly the all of the stations having relatively very strong correlation magnitudes with Camp Mabry.

What do the results of the regression analyses indicate about the correlation coefficients?

The null hypothesis of the performed regressions is that there is no relationship between the independent and dependent variables. Rejection of the null hypothesis in these regressions suggests that there is a relationship between the chosen independent variables and the Pearson coefficients and the stations. Examining both Table 7A and 7B for ABIA, the significance-F values of 0.0046 and 0.0096 respectively suggest rejection of both null hypotheses with a significance level of 0.01. Continuing, the adjusted R-Square value suggests that the independent variables are responsible for roughly 46% of the change in correlation values for ABIA using 24hr observations and roughly 41% of the change in correlation values for ABIA using 48hr observations. The p-values for the independent variable coefficients do not support the coefficients' reliability. With the smallest p-value, Station Elevations has the most support for being the strongest factor affecting ABIA correlations among the independent variables assessed.

Comparing the results for Camp Mabry, the significance-F value in Table 9A is 0.0252 with a significance level of 0.05 while the significance-F value in Table 9B is 0.00295 with a significance of 0.01. Both significance levels suggest rejection of both null hypotheses where the regression model for the 48hr observations has slightly stronger significance. Continuing with the comparison of the adjusted R-square values, the regression model for 24hr observations indicate that the independent variables are responsible roughly 27.5% of the change in correlation values for Camp Mabry, and the regression model for 48hr observations indicate that the independent variables are responsible for roughly 43.7% of the change in correlation values. The p-values for *Distance to Mabry* and *Station Elevation* coefficients both support their respective coefficient's reliability with significance levels each of 0.05 for the 24hr observation regression model compared to the 48hr observation model results where both variables coefficient's p-values support their reliability with significance levels of 0.01. Similar to the ABIA results, *Station Elevation* has the smallest p-values and therefore is marginally likely the strongest factor affecting Camp Mabry correlations among the independent variables assessed.

What do the interpolated visualizations for each event demonstrate?

By simply comparing the four visualizations altogether it should be clear that precipitation events express themselves heterogeneically across the region. Further examination shows that proximity of precipitation observations to USGS stream gages alone likely represents what relationship the two have.

A useful final step in analyzing the four flood events is to compare the greatest 24hr and 48hr precipitation recorded for each of the four floods to the updated values in the new (2018) precipitation atlas for Texas (USNOAA 2018) (Table 10). When considered in light of these precipitation frequency values, it can be seen that the 2010 flood on Bull Creek and the floods on Onion Creek were rare events with recurrence intervals of greater than 25 years whereas the 2019 flood on Barton Creek was not an exceptional event.   For the 2010 and 2013 floods, the precipitation values of the NOAA stations grossly underestimated the recurrence interval for the precipitation that produced the floods.

TABLE 10 - Precipitation recurrence interval values for the flood events (USNOAA 2018)

Flood Date	CoCoRaHS Greatest 24hr (in)	CoCoRaHS Greatest 48hr (in)	NOAA Greatest 24hr (in)	NOAA Greatest 48hr (in)
Bull Creek @ Loop 360, 9/8/2010	11.2	11.4	7.04	7.55
T_R *Years*	50 - 100	25 - 50	10 - 25	5 - 10
Onion Creek @ Twin Creeks, 10/31/2013	9.39	9.52	3.24	4.67
T_R *Years*	25 - 50	10 - 25	1 - 2	2 - 5
Onion Creek @ US 183, 10/31/2015	8.13	8.79	12.5	13.6
T_R *Years*	10 - 25	25	50 - 100	100
Barton Creek nr Oak Hill, 5/4/2019	5.75	5.99	4.77	4.77
T_R *Years*	5 - 10	2 - 5	2 - 5	2 - 5

Conclusions

The importance of broadening the scope of inclusion for precipitation observations was made evident in the results of this work. The correlation analysis and regression results suggest that there are more variables that need to be considered other than distance, elevation, and the simple completeness of precipitation records. While the correlations between the volunteered geographic information stations and Camp Mabry were relatively strong and all positive, the resulting regression model indicates that other variables are at play between the stations' correlations or that simple linear models do not suffice to represent their relationships.

The visualizations provide further evidence if nothing else that the use of political boundaries is a poor choice for determining boundaries related to natural disaster management and mitigation. The Onion Creek event's visualization clearly indicates extreme peak flow with relatively little precipitation within Travis County. The majority of its contributing watershed lies outside of Travis County, mostly in Hays County, and thus should also be incorporated not only in future planning, but also future research conducted on the use of volunteered geographic information. Reassessing RPRs should also be examined and likely based on need. RPRs for research likely differ from RPRs of real-world application and need to be considered depending on the goals of the stakeholders.

When the CoCoRaHS precipitation values are considered with their recurrence intervals it demonstrates that the floods at two of the four USGS gages (9/8/2010 and 10/31/2013) in the county were produced by precipitation events with recurrence intervals of greater than 25 years and a third flood (10/31/2015) was produced by a storm with a greater than 10 year probability. The much lower recurrence intervals for precipitation at the NWS stations at Camp Mabry and

ABIA leads to a gross underestimation of the severity of the precipitation that produced those floods.

Further research should include and consider other independent variables that likely influence how precipitation events express themselves related to stream flow gages such as hydrologic efficiency, land cover, watershed boundaries, soil moisture conditions, and research should also consider incorporation of other data sources such as the Lower Colorado River Authority’s Hydromet system (Lower Colorado River Authority).

Summary

This work examined correlations between volunteered geographic information observations and historically accepted professional observations of precipitation in Travis County across four major events related to precipitation. Correlation analysis suggests that the CoCoRaHS stations examined more closely correlate to Camp Mabry and that precipitation amounts at ABIA are often not related to major floods, or lack thereof, in Travis, County, which is somewhat logical as ABIA is located east/downstream of the watersheds that produce floods in Travis. County.

References

Ashley, S. T., and W. S. Ashley. 2008. Flood fatalities in the United States. *Journal of Applied Meteorology and Climatology* 47:805–18.

Blumenfeld, K. A., and R. H. Skaggs. 2011. Using a high-density rain gauge network to estimate extreme rainfall frequencies in Minnesota. *Applied Geography* 31:5–11.

Casella, G., and R. L. Berger. 2002. Statistical inference. Pacific Grove, CA: Duxbury.

Changnon, S. A. 2008. Assessment of flood losses in the United States. *Journal of Contemporary Water Research & Education* 138:38–44.

Changnon, S. A., and J. L. Vogel. 1981. Hydroclimatological characteristics of isolated severe rainstorms. *Water Resources Research* 17:1694–1700.

Cifelli, R., N. Doesken, P. Kennedy, L. D. Carey, S. A. Rutledge, C. Gimmestad, and T. Depue. 2005. The Community Collaborative Rain, Hail, and Snow Network: Informal education for scientists and citizens. *Bulletin of the American Meteorological Society* 86:1069–77.

Dulière, V., Y. Zhang, and E. P. Salathé, Jr. 2013. Changes in twentieth-century extreme temperature and precipitation over the western United States based on observations and regional climate model simulations. *Journal of Climate* 26:8556–75.

Dzotsi, K., C. Matyas, J. Jones, G. Baigorria, and G. Hoogenboom. 2014. Understanding high resolution spacetime variability of rainfall in southwest Georgia, United States. *International Journal of Climatology* 34:3188–3203.

Frech, M. 2018. *Flash Flood Alley*. Flood Statistics. https://www.floodsafety.com/media/maps/texas/index.htm (accessed March 14, 2018)

Gourley, J. J., J. M. Erlingis, Y. Hong, and E. B. Wells. 2012. Evaluation of tools used for monitoring and forecasting flash floods in the United States. *Weather and Forecasting* 27:158–73.

Habib, E., B. F. Larson, and J. Graschel. 2009. Validation of NEXRAD multisensor precipitation estimates using an experimental dense rain gauge network in south Louisiana. *Journal of Hydrology* 373:463–78.

Hosking, J. R. M., and J. R. Wallis. 2005. Regional frequency analysis: An approach based on L-moments. Cambridge, UK: *Cambridge University Press*.

Huff, F. A. 1994. Record-breaking microstorm system supports new rainfall frequency estimates in Illinois. *Bulletin of the American Meteorological Society* 75:1223–26.

Katz, R. W., M. B. Parlange, and P. Naveau. 2002. Statistics of extremes in hydrology. *Advances in Water Resources* 25:1287–1304.

Keighton, S., L. Lee, B. Holloway, D. Hotz, S. Zubrick, J. Hovis, G. Votaw, L. B. Perry, G. Lackmann, and S. E. Yuter. 2009. A collaborative approach to study northwest flow snow in the southern Appalachians. *Bulletin of the American Meteorological Society* 90:979–91.

Kelly, G. M., L. B. Perry, B. F. Taubman, and P. T. Soulé. 2012. Synoptic classification of 2009–2010 precipitation events in the southern Appalachian Mountains, USA. *Climate Research* 55:1–15.

Klemeš, V. 2000a. Tall tales about tails of hydrological distributions I. *Journal of Hydrologic Engineering* 5:227–31.

Klemeš, V. 2000b. Tall tales about tails of hydrological distributions II. *Journal of Hydrologic Engineering* 5:232–39.

Koutsoyiannis, D. 2004. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. *Hydrological Sciences Journal* 49:591–610.

Kundzewicz, Z.W., S. Kanae, S. I. Seneviratne, J. Handmer, N. Nicholls, P. Peduzzi, R. Mechler, L. M. Bouwer, N. Arnell, and K. Mach. 2014. Flood risk and climate change: Global and regional perspectives. *Hydrological Sciences Journal* 59:1–28.

Lower Colorado River Authority (LCRA 2020). 2020. LCRA's Hydromet. https://hydromet.lcra.org/ (accessed September 2, 2020)

Mattingly, K. S., Seymour, L., Miller, P. W. (2017) Estimates of Extreme Precipitation Frequency Derived from Spatially Dense Rain Gauge Observations: A Case Study of Two Urban Areas in the Colorado Front Range Region. *Annals of the American Association of Geographers* 107:6, 1499-1518, DOI: 10.1080/24694452.2017.1309961

Menne, M. J., I. Durre, B. Korzeniewski, S. McNeal, K. Thomas, X. Yin, S. Anthony, et al. 2012. *Global Historical Climatology Network—Daily (GHCN-Daily), 3.20*. Asheville, NC: National Centers for Environmental Information.

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston. 2012. An overview of the Global Historical Climatology Network–Daily database. *Journal of Atmospheric and Oceanic Technology* 29:897–910.

Moon, J. T., P. E. Guinan, D. J. Snider, and A. R. Lupo. 2009. CoCoRaHS in Missouri: Four years later, the importance of observations. *Transactions of the Missouri Academy of Science* 43:1–12.

Muller, C., L. Chapman, S. Johnston, C. Kidd, S. Illingworth, G. Foody, A. Overeem, and R. Leigh. 2015. Crowdsourcing for climate and atmospheric sciences: Current status and future potential. *International Journal of Climatology* 35:3185–3203.

Perica, S., D. Martin, S. Pavlovic, I. Roy, M. St. Laurent, C. Trypaluk, D. Unruh, M. Yekta, and G. M. Bonnin. 2013. NOAA Atlas 14*: Precipitation-frequency atlas of the United States*. Silver Spring, MD: National Oceanic and Atmospheric Administration.

Pielke, R. A., Jr., and M. W. Downton. 2000. Precipitation and damaging floods: Trends in the United States, 1932–97. *Journal of Climate* 13:3625–37.

Rafieeinasab, A., A. Norouzi, D.-J. Seo, and B. Nelson. 2015. Improving high-resolution quantitative precipitation estimation via fusion of multiple radar-based precipitation products. *Journal of Hydrology* 531:320–36.

Salas, J. D., J. H. Heo, D. J. Lee, and P. Burlando. 2012. Quantifying the uncertainty of return period and risk in hydrologic design. *Journal of Hydrologic Engineering* 18:518–26.

Thom, H. C. 1958. A note on the gamma distribution. *Monthly Weather Review* 86:117–22.

U.S. Department of Commerce. National Centers for Environmental Information, National Oceanic and Atmospheric Administration (USNOAA 2018). 2018. NOAA Atlas 14 Point Precipitation Frequency Estimates, TX. https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_cont.html?bkmrk=tx (accessed 6 September 2020)

U.S. Department of Commerce. National Centers for Environmental Information, National Oceanic and Atmospheric Administration. 2020. Latitude/Longitude Distance Calculator https://www.nhc.noaa.gov/gccalc.shtml (accessed March 5, 2020)

Villarini, G., J. A. Smith, and G. A. Vecchi. 2013. Changing frequency of heavy rainfall over the central United States. *Journal of Climate* 26:351–57.

Wackerly, D. D., W. Mendenhall, III, and R. L. Scheaffer. 2008. *Mathematical statistics with applications.* 7th ed. Belmont, CA: Brooks/Cole.

Wang, X., H. Xie, H. Sharif, and J. Zeitler. 2008. Validating NEXRAD MPE and Stage III precipitation products for uniform rainfall on the Upper Guadalupe River Basin of the Texas Hill Country. *Journal of Hydrology* 348:73– 86.

Willmott, C. J., S. M. Robeson, and M. J. Janis. 1996. Comparison of approaches for estimating time-averaged precipitation using data from the USA. *International Journal of Climatology* 16:1103–15.

Winkler, J. A. 1988. Climatological characteristics of summertime extreme rainstorms in Minnesota. *Annals of the Association of American Geographers* 78:57–73.

Yang, W. Y., Z. Li, T. Sun, and G. H. Ni. 2016. Better knowledge with more gauges? Investigation of the spatiotemporal characteristics of precipitation variations over the Greater Beijing Region. *International Journal of Climatology* 36:3607–19.

Zolina, O., C. Simmer, A. Kapala, P. Shabanov, P. Becker, H. Mächel, S. Gulev, and P. Groisman. 2014. Precipitation variability and extremes in central Europe: New view from STAMMEX results. *Bulletin of the American Meteorological Society* 95:995– 1002.