Louisiana State University
LSU Digital Commons

LSU Master's Theses | Graduate School

2016

A Statistical, Data-driven Assessment of Climate Extremes and Trends for the Continental U.S.

Xinbo Huang
Louisiana State University and Agricultural and Mechanical College, forhuang8804@gmail.com

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses

 Part of the Social and Behavioral Sciences Commons

Recommended Citation
Huang, Xinbo, "A Statistical, Data-driven Assessment of Climate Extremes and Trends for the Continental U.S." (2016). *LSU Master's Theses*. 4471.
https://digitalcommons.lsu.edu/gradschool_theses/4471

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

A STATISTICAL, DATA-DRIVEN ASSESSMENT OF CLIMATE EXTREMES AND TRENDS FOR THE CONTINENTAL U.S.

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science

in

The Department of Geography and Anthropology

by
Xinbo Huang
M.S. Wuhan University, 2013
December 2016

Acknowledgements

First, I would like to express my sincere thanks to my supervisor Dr. David Sathiaraj. Without his patient guidance and encouragement during the process of researching and writing, I could not have written this thesis. It is my honor to complete my master's study under his supervision.

Additionally, I also want to thank my advisor Dr. Lei Wang and Dr. Barry Keim for serving in my thesis committee and their valuable suggestions and precious time spent on improving my thesis research.

I also want to thank the Department of Geography & Anthropology. During my study in the master's program, I gained valuable skills and knowledge in academics and communication. Besides, I want to thank the NOAA Southern Regional Climate Center for providing me the opportunity to work as a graduate assistant (GA). The working experience as a GA in climate extremes related projects gave me practical training on solving real world problems such as managing and analyzing massive climate data sets.

Last, but not the least, my heartfelt appreciation goes to my parents, Haitao Huang and Yan Hong, for supporting me spiritually throughout writing this thesis and my life in general. Finally, I want thank all my close friends for their help and encouragement through my academic life.

Table of Contents

List of Tables

List of Figures

Abstract

Climate extremes are meteorological events that can have significant impacts on human and natural systems. Weather hazards, such as heat waves, drought, heavy thunderstorms, floods, hurricanes, occur frequently, and are a threat to human lives and property.

Climate data observations spanning over 100 years are an important asset in understanding climate extremes and trends. This research uses daily climate data observations from more than 3000 climate stations in the continental U.S. to assess the climate trends and extremes, including temperature, precipitation, and snowfall.

The climate data measurement sites were grouped by climate divisions and each climate division was statistically assessed for the following elements: maximum and minimum temperature, precipitation and snowfall. Furthermore, by dividing the climate data time series into 2 time intervals (1946-1980 and 1981-2015). Application of a host of non-parametric, statistical tests, provided insights on whether the recent time period is experiencing increased, decreased or similar frequencies of the climate extremes threshold being analyzed.

The study also analyzed trends of climate extremes on a regional basis by breaking up the continental US into western, high plains, southern, midwestern, northeast and southeast regions. A data visualization system was also developed to assess and analyze the results from this data-intensive study. The visualization system includes intuitive choropleth maps and charts that depict climate trends.

Chapter 1
Introduction

1.1 Purpose

In recent decades, global climate is changing and this change is apparent across a wide range of observations [4], and severe weather occurs frequently in recent years and causes casualties and property losses [5, 6]. Meanwhile, the linear trend of the globally averaged temperature combined land and ocean surface show a warming of 0.85 °C over the period 1880 to 2012 [1]. It is indisputable fact that climate extremes are inextricably associted with climate change.

The rule of more extreme weather and climate, coupled with increased vulnerability, highlights a need to collect, analyse, and assess extreme climate data so as to discover the trend of climate change and to prevent it further.

This research attempts to provide an assessment of trend in climate extremes for the the continental United States in recent decades by analyzing daily resolution climate data, including maximum temperature, minimum temperature, precipitation, and snowfall. In addition, in this study, a data visualization system is established to provide a portal to help users access these climate data more easily and intuitively.

1.2 Problem Statement

Due to the nature of climate data, some research problems addressed in this thesis includes

- Continental United States is vast with variable climate types and different land cover. The amount of climate data analyzed for the Continental United

States is enormous and analyzing a wide range of data is becoming a problem to process extreme climate data.

- Climate is defined as long-term averages and variations in weather measured over a period of several decades [4], and thus climate data are easily analyzed as time-series nature.However, it is a problem to tranform climate data into the threshold exceeding frequencies data set (TEF).

- The trend of climate data is an abstract concept and climate data itself contains numerous nuances. It is a problem to extract trend of climate extremes from numerous time-series climate data.

- The result and dataset involved in this study are volumonous and complex. It is necessary to establish an intuitive, friendly interface to help users grasp the trend and relationship in extreme climate data.

1.3 Organization of the Thesis

To achieve the goals outlined above, the thesis is divided into three sections:

- The first section, includes Chapter 3 and it details information regarding sources for the daily climate data and metadata, information on the 3000 climate stations in the continental U.S., the data preprocessing routines used to transform raw climate data observations into climate divisions based aggregations and deriving frequencies information, which is then tested using non-parameric statistics.

- The second section, includes Chapter 4, that describes the data visualization system and analysis and insights derived from the choropleth map-based visualizations.

- The last section, including Chapter 5, display general results about extreme high maximum temperature, extreme high minimum temperature, extreme low minimum temperature, extreme high precipitation, total annual precipitation, and total annual snowfall in the continental U.S. and regional result in six regions, including western region, high plains region, midwestern region, northeast region, southern region, and southeast region.

Chapter 2
Literature Review

This chapter reviews the literature about climate extremes and climate change.

Meehl [7] gave a definition and conceptual discussion of climate extremes. Houghton [8] focused on earlier climate model studies of global warming, and he started to analyze possible climate changes of future weather and climate extremes [9].

2.1 Temperature

Overall, a global warming of approximately 0.85 °C has occurred over the past century [1, 10, 11]. Human activities are at least partially responsible for the observed warming in the 20th century and particularly for that warming which has occurred in the latter half of the century, a view supported by numerous authors [12, 11, 13, 14, 15, 16, 7, 17]. These global trends may continue through this century, resulting in a global warming of 1 to 3.5 $^{\circ}C$ over the next century [13, 18, 19].

Hennessy [20] established a high-resolution regional model over southeastern Australia nested in a global model run under a transient enhanced greenhouse scenario shows that the frequency of minimum temperatures below freezing was roughly halved when the mean minimum temperature increased by nearly 2 $^{\circ}C$. Karl and Knight [21] discovered that increases in minimum temperature have appeared consistently in a number of different climate models and also are associated with an observed decrease in diurnal temperature range in some areas. The greatest change in the 20-year return values of daily maximum temperature is found in central and southeast North America, central and southeast Asia, and tropical Africa where there is a decrease in soil moisture content. Large extreme temperature increases also are seen over the dry surface of north Africa. In contrast,

the west coast of North America is affected by increased precipitation resulting in moister soil and more moderate increases in extreme temperature [17].

Representative Concentration Pathways (RCPs) Scenarios that include time series of emissions and concentrations of the full suite of greenhouse gases (GHGs) and aerosols and chemically active gases, as well as land use/land cover [22]. The word representative signifies that each RCP provides only one of many possible scenarios that would lead to the specific radiative forcing characteristics. The term pathway emphasizes that not only the long-term concentration levels are of interest, but also the trajectory taken over time to reach that outcome [23].

RCPs usually refer to the portion of the concentration pathway extending up to 2100, for which Integrated Assessment Models produced corresponding emission scenarios. Extended Concentration Pathways (ECPs) describe extensions of the RCPs from 2100 to 2500 that were calculated using simple rules generated by stakeholder consultations and do not represent fully consistent scenarios.

In Climate Change 2014 Synthesis Report [24], it is described that the increase of global mean surface temperature by the end of the 21st century (2081-2100) relative to 1986-2005 is likely to be 0.3 °C to 1.7 °C under RCP2.6, 1.1 °C to 2.6 °C under RCP4.5, 1.4 °C to 3.1 °C under RCP6.0 and 2.6 °C to 4.8 °C under RCP8.59. The Arctic region will continue to warm more rapidly than the global mean.

Four RCPs produced from Integrated Assessment Models were selected from the published literature and are used in the present IPCC Assessment as a basis for the climate predictions and projections presented n WGI AR5 Chapters 11 to 14 [24]

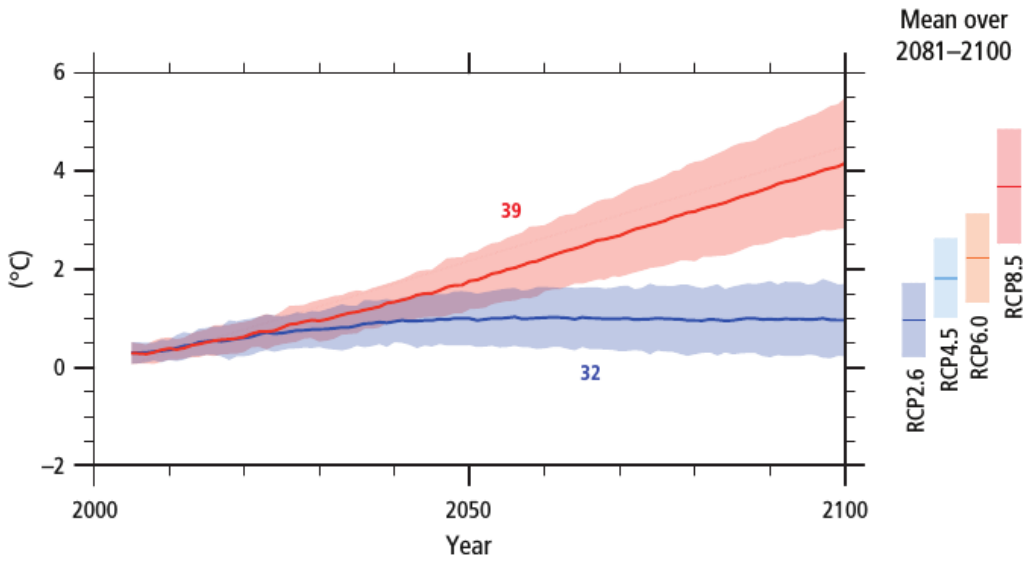

FIGURE 2.1. Global average surface temperature change with RCP2.6 and RCP8.5 (relative to 1986-2005) [1]

- RCP2.6 One pathway where radiative forcing peaks at approximately 3 W/m^2 before 2100 and then declines.

- RCP4.5 and RCP6.0 Two intermediate stabilization pathways in which radiative forcing is stabilized at approximately 4.5 W/m^2 and 6.0 W/m^2 after 2100.

- RCP8.5 One high pathway for which radiative forcing reaches >8.5 W/m^2 by 2100 and continues to rise for some amount of time.

In Figure 2.1, Global average surface temperature change from 2006 to 2100 as determined by multi-model simulations. All changes are relative to 1986-2005. Time series of projections and a measure of uncertainty (shading) are shown for scenarios RCP2.6 (blue) and RCP8.5 (red). The mean and associated uncertainties averaged over 2081-2100 are given for all RCP scenarios as coloured vertical bars at the right hand side of each panel.

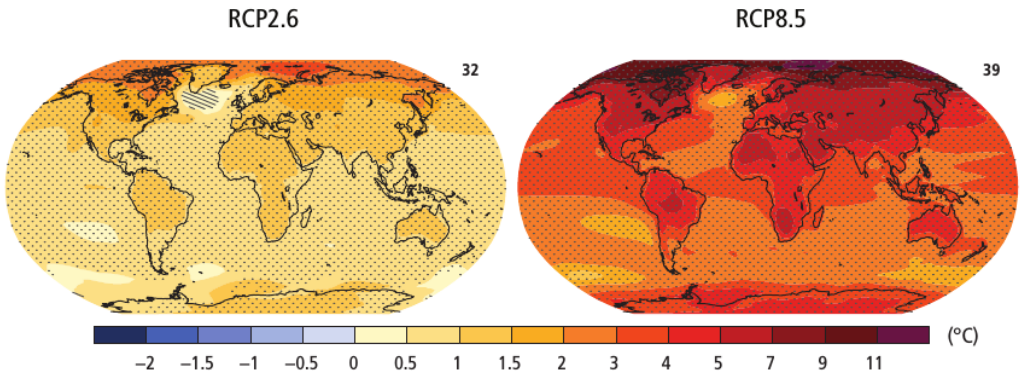

FIGURE 2.2. Change in average surface temperature (1986-2005 to 2081-2100) [1]

Figure 2.2 reflects change in average surface temperature based on multi-model mean projections for 2081-2100 relative to 1986-2005 under the RCP2.6 (left) and RCP8.5 (right) scenarios. The number of models used to calculate the multi-model mean is indicated in the upper right corner of each panel. Stippling shows regions where the projected change is large compared to natural internal variability and where at least 90% of models agree on the sign of change. Hatching shows regions where the projected change is less than one standard deviation of the natural internal variability.

2.2 Precipitation

Changes in precipitation will not be uniform. The high latitudes and the equatorial Pacific are likely to experience an increase in annual mean precipitation under the RCP8.5 scenario. In many mid-latitude and subtropical dry regions, mean precipitation will likely decrease, while in many mid-latitude wet regions, mean precipitation will likely increase under the RCP8.5 scenario. Extreme precipitation events over most of the mid-latitude land masses and over wet tropical regions will very likely become more intense and more frequent.

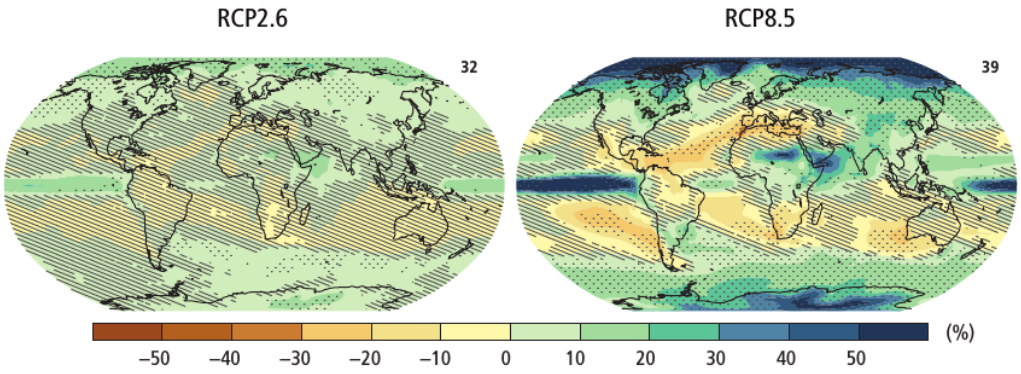

FIGURE 2.3. Change in average precipitation (1986-2005 to 2081-2100) [1]

Increased precipitation intensity (albeit with certain regional variations) in a future climate with increased greenhouse gases was one of the earliest model results regarding precipitation extremes, and remains a consistent result with improved, more detailed models [25, 26]. There are also some indications from observations that such changes of precipitation intensity are already being seen in some regions [21]. There have been questions regarding the relatively coarse spatial scale resolution in climate models being able to represent essentially mesoscale and smaller precipitation processes. However, the globally averaged increase in moisture capacity of a warmer atmosphere is physically consistent with increases in precipitation and, potentially, with increases of precipitation rate in some regions.

It has been recognized recently that changes in precipitation intensities could have a geographical dependence. For example, Bhaskharan and Mitchell [27] note that the range of precipitation intensity over the south Asian monsoon region broadens in a future climate experiment with increased greenhouse gases, with decreases prevalent in the west and increases more widespread in the east. Increases in extreme precipitation events recently have been projected in nested regional models over Australia [20] and the United States [28], and in a high-resolution

nested hurricane model over the northwest tropical Pacific [29]. In a recent global model simulation with doubled CO_2, precipitation extremes increase more than the mean daily precipitation (the mean increase is 4%; 20-yr extreme precipitation event return values increase 11%) with a consequent decrease in return period for the 20-yr extreme precipitation events almost everywhere [30].

Chapter 3
Data and Methologies

This chapter introduces the data and methodologies and related problems. Figure 3.1 is the flowchart for data processing in the study. First, daily climate data are obtained from ACIS. The data are then grouped by climate divisions, and extreme frequency data are generated by setting some thresholds. Grouping by climate divisions is conducted as follows: An average was computed using data from at least 3 climate measurement sites - each of the sites included data that spanned the time period 1946-2015 and each of the climate measurement sites included less than 10% missing values per year. So for example, to compute for a say New York's climate division 1, annual frequencies of minimum temperatures exceeding 75 °F are collected from at least 3 climate measurement sites and a mean annual frequency value is computed. If less than 3 climate measurement sites were available for a climate division (likely due to excessive missing values), then that climate division was excluded from the study.

The time-series extreme frequency data between 1946 and 2015 is divided into two independent samples, which can be compared using non-parametric statistical hypothesis test. Finally, p-value is obtained to be the representation of significance of whether the two time-series climate data have a similar distribution. In addition, the difference between the means of the 2 time-series data for each of the thresholds is also evaluated, to indicate an increasing [1] or decreasing [2] trend.

[1] In this thesis, ***increasing trend*** denotes that the more recent time period (1981-2015) is experiencing higher mean frequency of days for the threshold being analyzed, when compared to the prior time period (1946-1980)

[2] ***decreasing trend*** denotes that the more recent time period (1981-2015) is experiencing lower mean frequency of days for the threshold being analyzed, when compared to the prior time period (1946-1980)

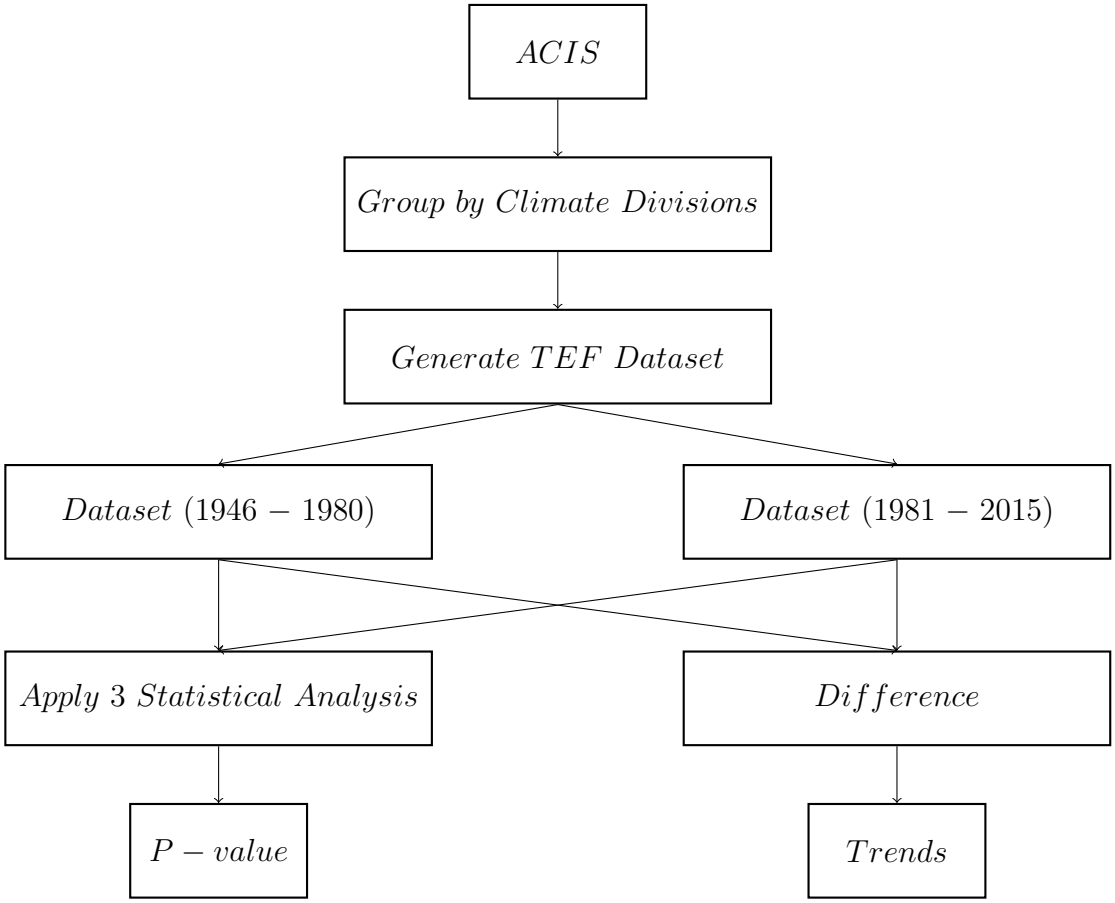

FIGURE 3.1. Flowchart for data processing

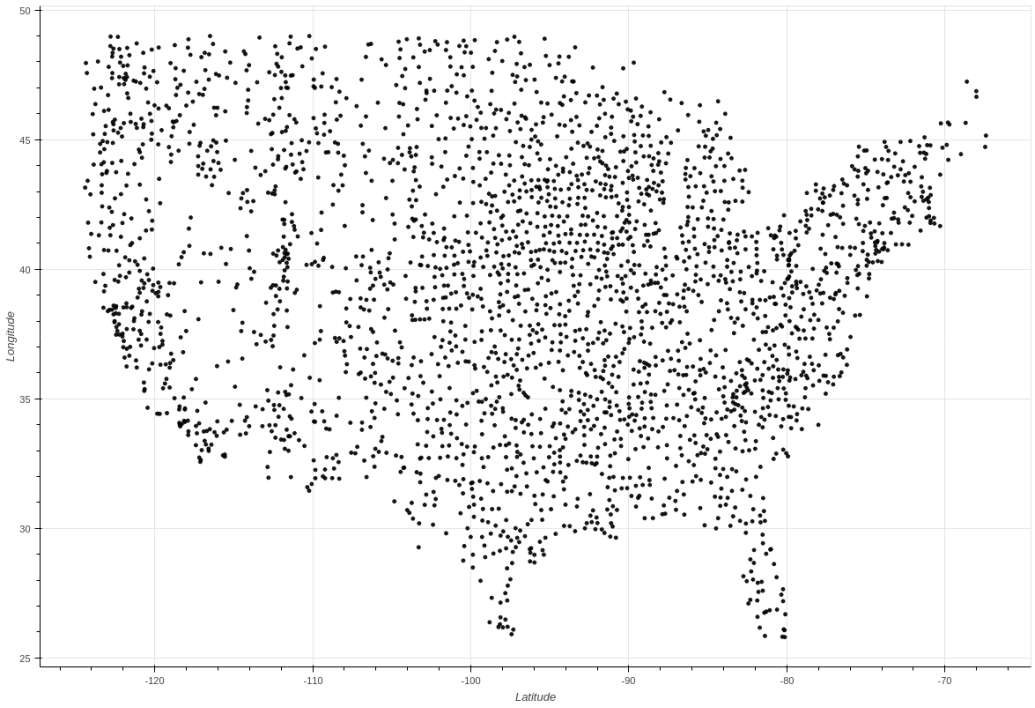

FIGURE 3.2. Spatial distribution of climate data observing stations

3.1 Data Source

This research used the daily climate data from the Applied Climate Information System (ACIS), an Internet-based system designed to facilitate the generation and dissemination of climate data products to users. ACIS is developed by the NOAA Regional Climate Centers (RCCs) to manage the complex flow of information from climate data collectors to end users of climate data information.

ACIS accepts and returns climate information in JavaScript Object Notation (JSON), which uses structures that are similar to those used in many coding languages, including C, C++, Java, JavaScript, Perl, and Python. For each call, users specify a set of parameter to describe the data being requested. After passing these parameters to the server and accessing these climate data, a climate data product is returned to users.

As shown in Figure 3.2, 3210 climate stations in the continental United States were used in this study. There are more than 26000 GHCN climate data measure-

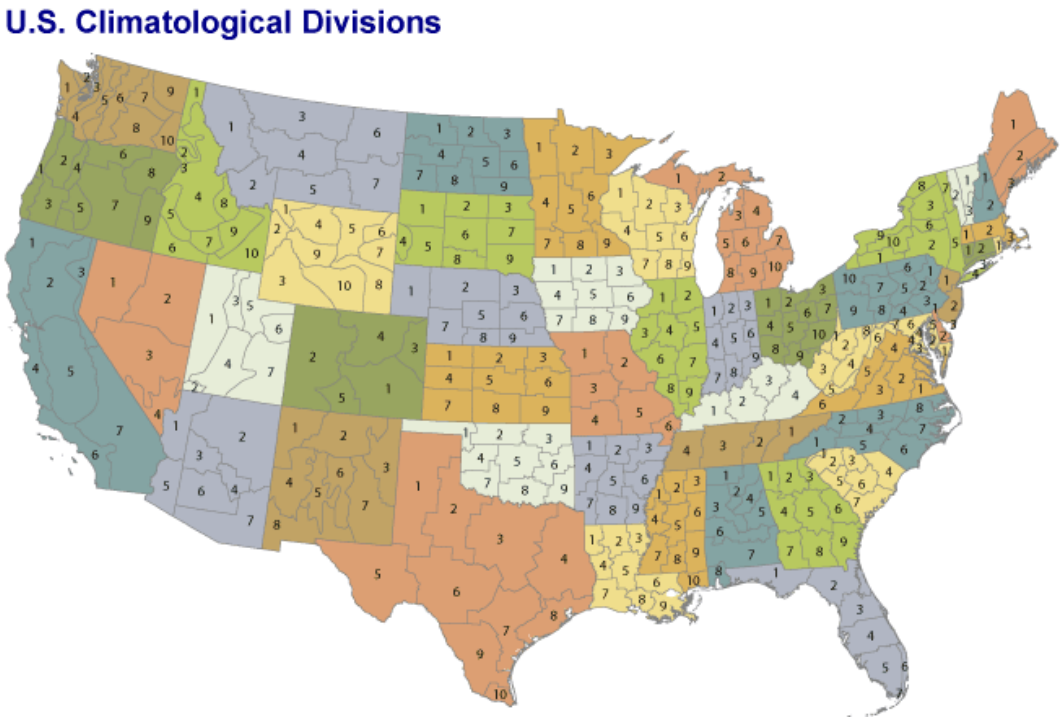

FIGURE 3.3. Climate divisions of the continental U.S. [2]

ment sites. However not all span the entire time period of 1946-2015. Additional criteria used for this data analysis and study included the following: allow for less than 10% missing values for a station per year and every climate division should have at least 3 climate measurement sites. Once this criteria was applied, the number of valid stations that fit these criteria reduced to 3210. These 3210 stations are distributed to cover most of land in the continental United States. By analyzing the climate data from these stations, the trends of climate extremes can then be obtained for the continental United States.

3.2 Data Preprocessing

As daily climate data from ACIS is raw data which cannot be processed further, data preprocessing is necessary to generate extreme climate dataset.

Require: $T_{s,v}$
1: Initialize int $NumClimdiv = 327$
2: Initialize array $ArrSum[NumClimdiv] = \phi$
3: Initialize array $ArrNum[NumClimdiv] = \phi$
4: Initialize array $ArrClimdiv[NumClimdiv] = \phi$
5: Initialize array $ArrRes[NumClimdiv] = \phi$
6: **for** each dataset dc in $T_{v,s}$ **do**
7: $ClimdivCode = dc[ClimdivCode]$
8: **if** $ClimdivCode$ in $ArrClimdiv$ **then**
9: $ArrSum[ClimdivCode] += dc[Value]$
10: $ArrNum[ClimdivCode] += 1$
11: **else**
12: Insert $dc[Value]$ into $ArrSum$
13: Insert 1 into $ArrNum$
14: Insert $ClimdivCode$ into $ArrClimdiv$
15: **end if**
16: **end for**
17: **for** i := 0 to $NumClimdiv - 1$ **do**
18: $ArrRes[i] = ArrSum[i]/ArrNum[i]$
19: **end for**
20: Export $ArrRes$

FIGURE 3.4: Algorithm used to group data by climate division

3.2.1 Climate Division

As shown in Figure 3.3, the continental United States (U.S.) is subdivided into 344 climate divisions by The National Climatic Data Center (NCDC) [31]. For each climate division, which represents nearly homogenous climatic regions, monthly station temperature and precipitation values are computed from the daily observations [32], and their monthly temperature, monthly water equivalent precipitation, Palmer Drought Severity Index, and Palmer Hydrological Drought Index values have been generated back to 1895 [33].

Numerous applications have used these climate divisional data, e.g., they are used to monitor the U.S. climate by the NCDC, the Climate Prediction Center, the National Drought Mitigation Center, and others. These divisional data sets are also used frequently in applied research [34, 35].

```
Require: T_{s,v}
Initialize int NumYears = 70
Initialize array ArrRes[NumYears] = φ
Initialize array ArrYears[NumYears] = φ
Initializa array Arr
for each dataset dc in T_{v,s} do
    if dc[Value] then
        if dc in ArrYears then
            ArrRes[dc[Year]]+ = 1
        else
            Insert 1 into ArrRes
            Insert dc[Year] into ArrYears
        end if
    end if
end for
Export ArrRes
```

FIGURE 3.5: Algorithm used to groupe data by time

The NCEI climate divisions shape file (geographical dataset) includes 327 out of the 344 climate divisions in the continental U.S.. Hence for this study, climate measurement data from 3210 data sites were grouped into 327 climate divisions to derive the Threshold Exceeding Frequencies (TEF) Dataset. To group a set of climate measurement sites into a climate division, a minimum of 3 climate measurement stations were required and the frequency measurement (number of days exceeding a threshold) was averaged for the climate division. Each of the climate data sites in a climate division was also required to have less than 10% missing values per year. Figure 3.4 is pseudocode to group daily climate data by climate divisions and compute the means in every climate divisions.

3.2.2 Extreme Frequencies Dataset

The daily climate data from 3210 climate measurement sites was then transformed into the Threshold Exceeding Frequency (TEF) data set. The threshold used for this study was based on a combination of thresholds used in the *CLIMDEX - Datasets for Indices of Climate Extremes* [36] and the Southeast chapter of the *US National Climate Assessment* document, released in 2014 [37]. The pseudocode or

TABLE 3.1. Thresholds to generate extreme climate frequency dataset

Element	Thresholds
Maximum Temperature	≥105, ≥100, ≥95, ≥85
Minimum Temperature	≥80, ≥75, ≥70, ≥65, ≤36, ≤32, ≤28, ≤24, ≤15, ≤10, ≤5, ≤0
Precipitation	≥2, ≥4, sum
Snowfall	sum

procedure for grouping the data by year and number of days exceeding a threshold is depicted in Figure 3.5.

There are some threshold is chosen, including that maximum temperature is greater 105 °F, 100 °F, 95 °F, or 85 °F, minimum temperature is greater than 80 °F, 75 °F, 70 °F, or 65 °F, minimum temperature is lower than 36 °F, 32 °F, 28 °F, 24 °F, 15 °F, 10 °F, 5 °F, 0 °F, precipitation is greater than 2 inches or 4 inches, total annual precipitation (in inches), and total annual snowfall (in inches), as shown in Table 3.1. In addition, to ensure availability of the frequencies dataset, climate divisions should have no more than 10% missing values per year.

3.3 Non-Parametric Test

The determination of the distribution form which a sample is drawn is an important problem in many statistical applications [38]. If the distribution is not known, or is known to not follow a particular form, then non-parametric statistics are appropriate.

The verification of the compatibility of a set of observed sample values with a hypothesized distribution is carried out by a goodness-of-fit test. Various studies have shown that for continuous populations several tests based on the empirical distribution function (edf), including the Kolmogorov-Smirnov test are used.

Thus, the three non-parametric test, Wilcoxon Signed-Rank Test, Mann-Whitney U Test, and Kolmogorov-Smirnov Test, are chosen to compare distributional form of the population in two time-series datasets.

3.3.1 Wilcoxon Signed-Rank Test

The Wilcoxon Test of the hypothesis H_0, two continuous distribution functions, F and G, are equal, is typically based on independent random samples, $X_1, ..., X_m$ from F, $Y_1, ..., Y_n$ from [39]. In order to test whether X_1 is stochastically larger than Y_1 [40], Wilcoxon [41] introduced the statistic

$$W_{m,n} = sum\ of\ the\ ranks\ of\ the\ X_i's\ in\ the\ combined\ sample$$

The algorithm is the core of Wilcoxon Signed-Rank Test. Wilcoxon Test focus on accessing whether the population mean ranks differ.

3.3.2 Mann-Whitney U Test

Mann-Whitney (MW) statistical test is a test for assessing the signficance of a difference in median or central tendency or mean of two series. By comparing with parametric statistical tests such as a t-test, the nonparametric test is more suitable for non-normally distributed data and censored data, which are frequently encountered in hydrological time serie [42, 43].

Mann and Whitney [44] introduced the equivalent statistic

$$M_{m,n} = \sum_{i=1}^{m} \#\{j : Y_j < X_i\}$$

The equivalence of these statistics can be seen as follows. Let $X_{(i)}$ denote the ith-order statistic of $X_1, ..., X_m$. Then for $i = 2, ..., m$, the ranks of $X_{(1)}, ..., X_{(i-1)}$ are included in $W_{m,n}$, but not in $M_{m,n}$. Hence, $W_{m,n} = M_{m,n} + \sum_{i=1}^{m} i = M_{m,n} + \frac{1}{2}m(m+1)$ [40].

Mann-Whitney U Test is different from Wilcoxon Signed-Rank Test. The Mann-Whitney U test is applied to independent samples, and the Wilcoxon signed-rank test is applied to matched or dependent samples.

3.3.3 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov test [45] is the best-known and most widely used goodness-of-fit test based on the empirical distribution function (edf). In a random sample of size n, the edf, denoted by $S_n(x)$, is defined by

$$S_n(x) = \begin{cases} 0, & x < X_{(1)}, \\ i/n, & X_{(i)} \leq x < X_{(i+1)} \ for \ i = 1, 2, ..., n-1, \\ 1, & x \geq X_{(n)}, \end{cases}$$

where $X_{(i)}$ represents the i-th order statistic in the random sample. As in Harter, Khamis and Lamb (1984), we define

$$d_n^+ = \max_{1 \leq i \leq n} [(i - 0.5)/n - F_i]$$

$$d_n^- = \max_{1 \leq i \leq n} [F_i - (i - 0.5)/n]$$

$$d_n = max(d_n^+, d_n^-) = \max_{1 \leq i \leq n} [|(i - 0.5)/n - F_i|]$$

and base the KS test on these statistics. The relationship between these test statistics and the one-sided (D_n^+, D_n^-) and two-sided (D_n) KS test statistics is

$$D_n^+ = d_n^+ + 0.5/n$$

$$D_n^- = d_n^- + 0.5/n$$

$$D_n = d_n + 0.5/n$$

where F_i, is the theoretical (population) cdf, $F(X_{(i)})$, corresponding to the i-th order statistic [38].

Compared with Wilcoxon Signed-Rank Test, Mann-Whitney U Test, Kolmogorov-Smirnov Test is more sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples.

Chapter 4
Data Visulization

4.1 Structure of Data Visualization System

To demonstrate the results of this study and to intuitively help users access the climate extremes frequencies data set, a data visualization system was developed. Typical users can span domains such as climate science, agriculture, finance and economics (commodity markets), recreation site managers (such as those in the ski resort industry) and actuarial science. This system will help users to query the climate threshold exceedance frequencies derived in this study and provides an interface to depict the statistical analysis conducted.

The data visualization system contains a low latency, robust memory database, a flexible real-time query system, and a user-friendly web interface. This system is depicted in Figure 4.1. The process flow is as follows: the querying options in the interface receive a set of meta data and threshold parameters from the user and this information is transmitted to the database. The database caches the threshold exceeding frequencies (TEF) data set. Once the data is retrieved from the database, it is sent to the interface that includes a map and chart-based visualizations.

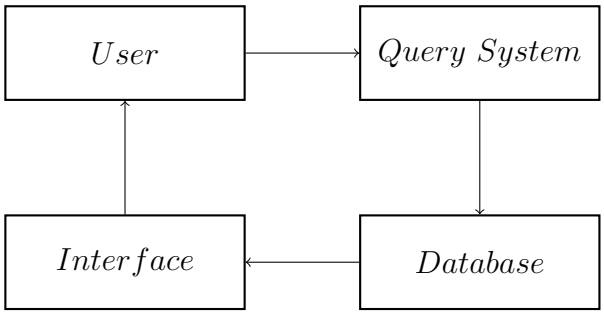

FIGURE 4.1: Pipeline of data visualization system

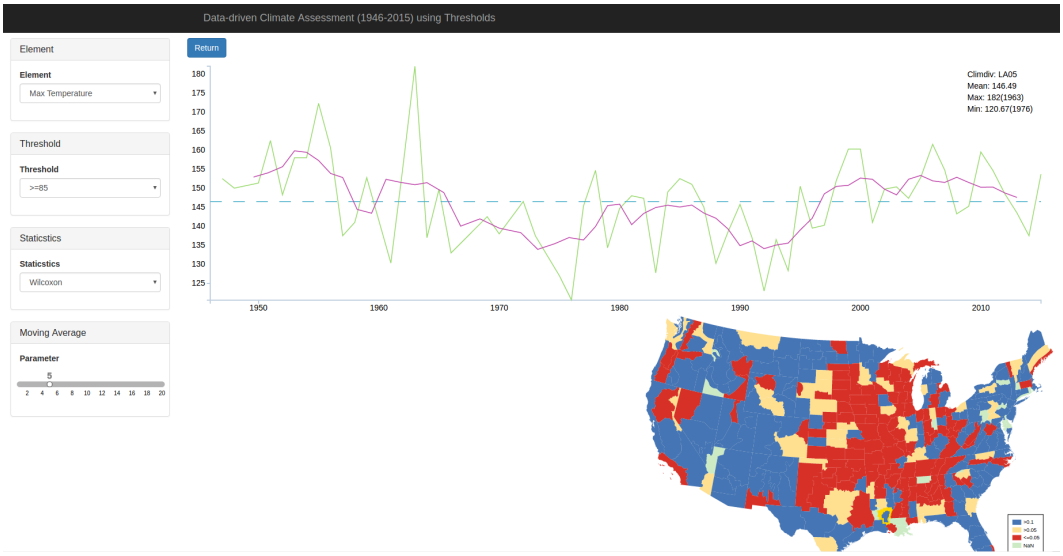

FIGURE 4.2: Screenshot of data visualization system with chart of maximum temperature

Figure 4.2 displays the interface of the data visualization system. It contains a query panel that can set the query parameter of the produce, choropleth map to show the distribution of p-values and trends in the continental United States, and produce a line chart to display the climate extremes frequencies dataset of climate divisions.

The query panels includes an element selector which can select climate elements including maximum temperature, minimum temperature, precipitation, and snow, a threshold selector, a statistic selector which can select statistic methods including Wilcoxon, Mann-Whitney, K-S, and difference, and a moving average parameter slider which can select integers from 1 to 20. These query panels are interacted automatically with users. The data visualization system would refresh all the climate data when users operate the query panels.

4.2 Choropleth Map

A choropleth map is a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable being displayed on the

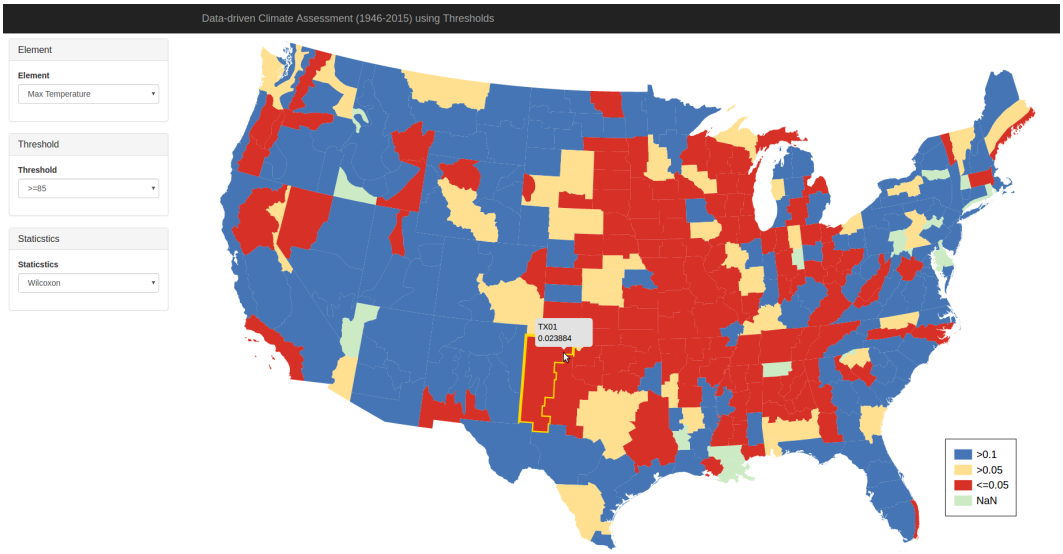

FIGURE 4.3: Screenshot of data visualization system with Wilcoxon p-value choropleth map when maximum temperature is greater than 85 °F

map [46]. The interface of the data visualization system displays a choropleth map of p-value or difference of climate dataset of climate divisions in the continental United States. For example, as shown in Figure 4.3, there is the choropleth map about wilcoxon test p-value of the climate dataset of climate divisions whose maximum temperature is greater than 85 °F. In the choropleth map, the blue represents that the p-value is greater than 0.1 and the transformation is not significant, the yellow represents that the p-value is between 0.05 and 0.1 and the transformation is not significant relatively, the red represents that the p-value is lower than 0.05 and the transformation is significant, and the green represents there are no data.

When a mouse moves over a climate division in the choropleth map, the name and p-value of the climate division would be shown in a dialog box. As shown in Figure 4.3, the name, TX01, and the p-value, 0.023884, are displayed in the dialog box.

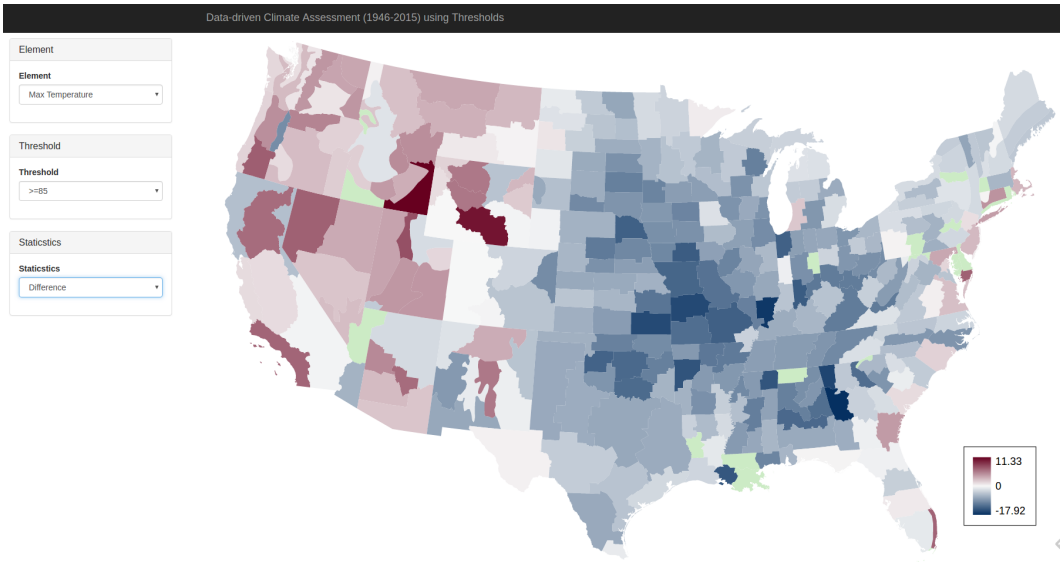

FIGURE 4.4: Screenshot of data visualization system with difference choropleth map when maximum temperature is greater than 85 °F

Figure 4.4 is the choropleth map for difference when minimum temperature is greater than 85 °F. The red areas means the frequencies in these places increase and the blue areas means that in these decrease.

4.3 Line Chart

When a climate division in the choropleth map is clicked, the line chart about the time-series extreme climate data for the climate division is displayed to help users access more detailed information. Figure 4.5 shows interface about the chart, the green line represents climate extremes frequencies data of the climate division, the blue dash line represents the mean of the data, the red line represents moving average line and users can set the parameter of moving average line in the left panel.

When a mouse is moved over the chart, the closest value is shown so that users can check all the value in the climate extremes datasets.

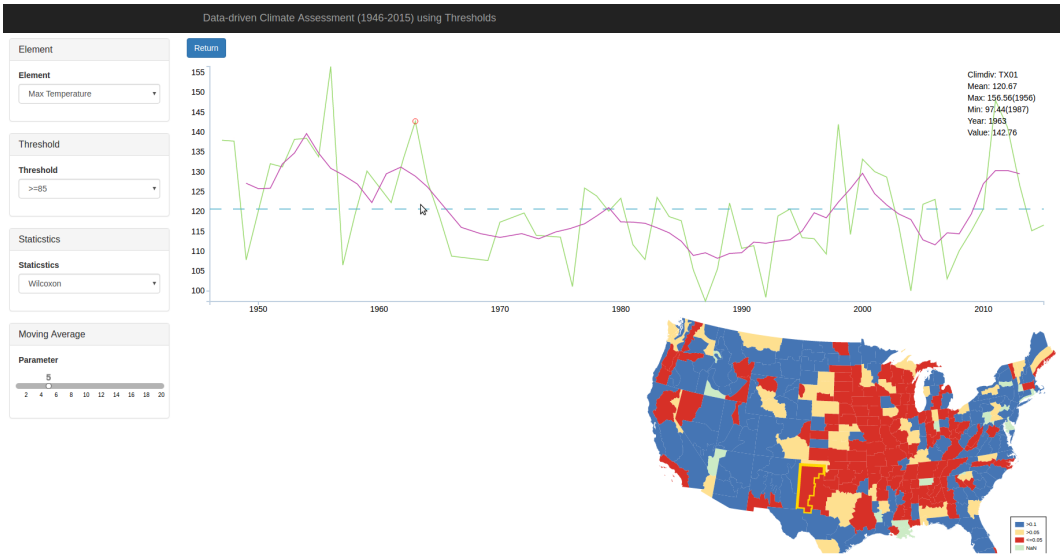

FIGURE 4.5: Screenshot of line chart with the climate division TX01 when maximum temperature is greater than 85 °F

Chapter 5
Results and Analysis

This chapter discusses the result of the study to examine the trends of climate change in recent decades.

5.1 P-value

In statistics, the p-value is a function of the observed sample results (a test statistic) relative to a statistical model, which measures how extreme the observation is [47]. The p-value is defined as the probability of obtaining a result equal to or "more extreme" than what was actually observed, when the null hypothesis is true [48]. A small p-value (typically ≤ 0.05) indicates strong evidence against the null hypothesis, so you reject the null hypothesis. A large p-value (> 0.05) indicates weak evidence against the null hypothesis, so you fail to reject the null hypothesis [49, 50, 51].

Table 5.1, Table 5.2, and Table 5.3 display the number of climate divisions with different statistic test p-value ranges, including p-value ≤ 0.05, $0.05 <$p-value≤ 0.1, and p-value> 0.1. The index names represent the climate elements (tx means maximum temperature, tn means minimum temperature, pc means precipitation, and sw means snow) and thresholds. The percentage in this table means the portion that climate divisions with increasing trends (the difference between the 2 extreme frequencies datasets is positive) take up in these climate divisions with different significence.

In this table, some issues can be observed

TABLE 5.1. Number of climate divisions grouped by wilcoxon test p-value from two ETF datasets (1946-1980 and 1981-2015) and the percentage of climate divisions with an increasing trend (NaN means that p-values are unavailable in these climate divisions)

	<=0.05	>0.05	>0.1	NaN
tx>=105	7 (57%)	10 (8%)	306 (43%)	4
tx>=100	33 (39%)	19 (63%)	271 (39%)	4
tx>=95	50 (30%)	25 (32%)	248 (38%)	4
tx>=90	64 (12%)	39 (26%)	220 (4%)	4
tx>=85	124 (1%)	43 (19%)	156 (38%)	4
tn>=80	39 (95%)	4 (75%)	280 (51%)	4
tn>=75	106 (95%)	21 (95%)	196 (66%)	4
tn>=70	124 (97%)	27 (93%)	172 (77%)	4
tn>=65	107 (96%)	25 (80%)	191 (80%)	4
tn<=36	161 (2%)	29 (14%)	133 (29%)	4
tn<=32	168 (2%)	30 (7%)	125 (32%)	4
tn<=28	132 (2%)	41 (7%)	150 (27%)	4
tn<=24	132 (1%)	41 (0%)	150 (21%)	4
tn<=15	113 (0%)	35 (0%)	175 (16%)	4
tn<=10	100 (0%)	35 (0%)	188 (17%)	4
tn<=5	78 (0%)	30 (0%)	215 (21%)	4
tn<=0	69 (0%)	22 (0%)	232 (20%)	4
pc>=2	56 (98%)	39 (100%)	235 (79%)	0
pc>=4	13 (100%)	16 (94%)	301 (61%)	0
pcsum	87 (99%)	32 (94%)	211 (78%)	0
swsum	96 (4%)	23 (26%)	211 (3%)	0

TABLE 5.2. Number of climate divisions grouped by Mann-Whitney test p-value from two ETF datasets (1946-1980 and 1981-2015) and the percentage of climate divisions with increasing trend (NaN means that p-values are unavailable in these climate divisions)

	<=0.05	>0.05	>0.1	NaN
tx>=105	24 (75%)	13 (77%)	220 (52%)	70
tx>=100	40 (38%)	26 (50%)	250 (42%)	11
tx>=95	55 (27%)	23 (35%)	249 (38%)	0
tx>=90	66 (11%)	39 (26%)	222 (4%)	0
tx>=85	127 (9%)	43 (19%)	157 (38%)	0
tn>=80	57 (91%)	17 (94%)	167 (69%)	86
tn>=75	118 (93%)	25 (92%)	159 (74%)	25
tn>=70	135 (94%)	28 (93%)	157 (79%)	7
tn>=65	113 (93%)	24 (79%)	188 (81%)	2
tn<=36	165 (2%)	28 (14%)	134 (28%)	0
tn<=32	173 (2%)	28 (7%)	126 (32%)	0
tn<=28	135 (1%)	42 (7%)	150 (27%)	0
tn<=24	136 (1%)	41 (0%)	150 (21%)	0
tn<=15	120 (0%)	34 (0%)	170 (16%)	3
tn<=10	104 (0%)	38 (3%)	179 (17%)	6
tn<=5	82 (0%)	28 (0%)	205 (22%)	12
tn<=0	74 (0%)	25 (4%)	210 (21%)	18
pc>=2	59 (98%)	39 (97%)	231 (80%)	1
pc>=4	27 (89%)	19 (79%)	253 (68%)	31
pcsum	87 (99%)	32 (94%)	211 (78%)	0
swsum	96 (4%)	26 (23%)	206 (31%)	2

TABLE 5.3. Number of climate divisions grouped by K-S test p-value from two ETF datasets (1946-1980 and 1981-2015) and the percentage of climate divisions with increasing trend (NaN means that p-values are unavailable in these climate divisions)

	<=0.05	>0.05	>0.1	NaN
tx>=105	7 (57%)	4 (50%)	312 (44%)	4
tx>=100	24 (38%)	19 (47%)	280 (41%)	4
tx>=95	44 (25%)	19 (58%)	260 (37%)	4
tx>=90	51 (18%)	28 (18%)	244 (38%)	4
tx>=85	111 (8%)	42 (29%)	170 (35%)	4
tn>=80	31 (97%)	5 (60%)	287 (53%)	4
tn>=75	99 (95%)	19 (84%)	205 (69%)	4
tn>=70	98 (97%)	21 (86%)	204 (80%)	4
tn>=65	85 (95%)	28 (89%)	210 (81%)	4
tn<=36	149 (2%)	24 (4%)	150 (28%)	4
tn<=32	158 (2%)	29 (7%)	136 (30%)	4
tn<=28	132 (1%)	28 (14%)	163 (25%)	4
tn<=24	118 (1%)	31 (0%)	174 (18%)	4
tn<=15	103 (0%)	20 (0%)	200 (14%)	4
tn<=10	91 (0%)	24 (0%)	208 (15%)	4
tn<=5	70 (0%)	19 (0%)	234 (20%)	4
tn<=0	52 (0%)	20 (0%)	251 (18%)	4
pc>=2	42 (98%)	27 (100%)	261 (81%)	0
pc>=4	11 (100%)	8 (100%)	311 (62%)	0
pcsum	77 (100%)	38 (89%)	215 (79%)	0
swsum	86 (9%)	30 (17%)	214 (29%)	0

- Compared with other climate elements, p-value about minimum temperature is smaller. It means that in recent decades, minimum temperature changed more dramatically.

- When p-value is lower than 0.05, extreme high minimum temperature frequencies dataset is on the increase.

- When p-value is lower than 0.05, extreme low minimum temperature frequencies dataset is on the decrease.

- When p-value is lower than 0.05, extreme high precipitation dataset is on the increase.

- When p-value is lower than 0.05, snowfall dataset is on the decrease.

In this study, the seven thresholds, maximum temperature $\geq 95°$F, minimum temperature $\geq 75°$F, minimum temperature $\leq 32°$F, minimum temperature $\leq 0°$F, precipitation $\geq$ 2 inches, annual total precipitation, annual total snowfall, are chosen to be analyzed further.

On the other hand, the continental U.S. is vast and contains variable climate types, such as humid subtropical climate and desert climate, and different land cover, such as mountains and cities. There are six al Climate Centers representing the continental U.S., including western region, high plains region, midwestern region, northeast region, southern region, and southeast region, as shown in Figure 5.1. To analyze trends of the climate change in different regions, results of the six regions are considered seperately. Table 5.4 lists the correspondence between the six regions and states.

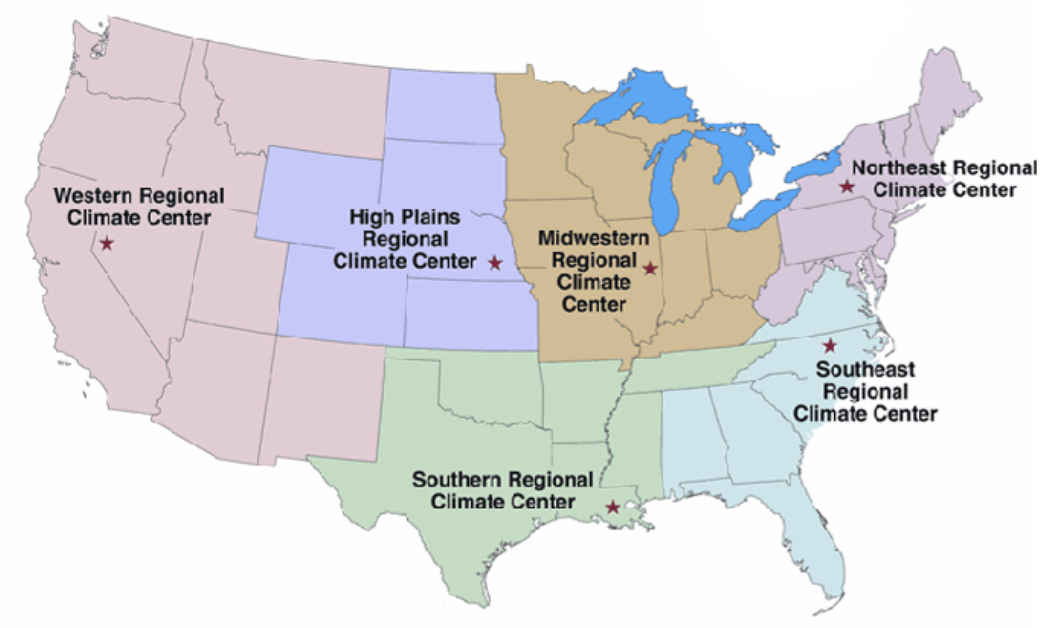

FIGURE 5.1: Regional climate center regions in the continental U.S. [3]

TABLE 5.4. Distribution of states in the six regions

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

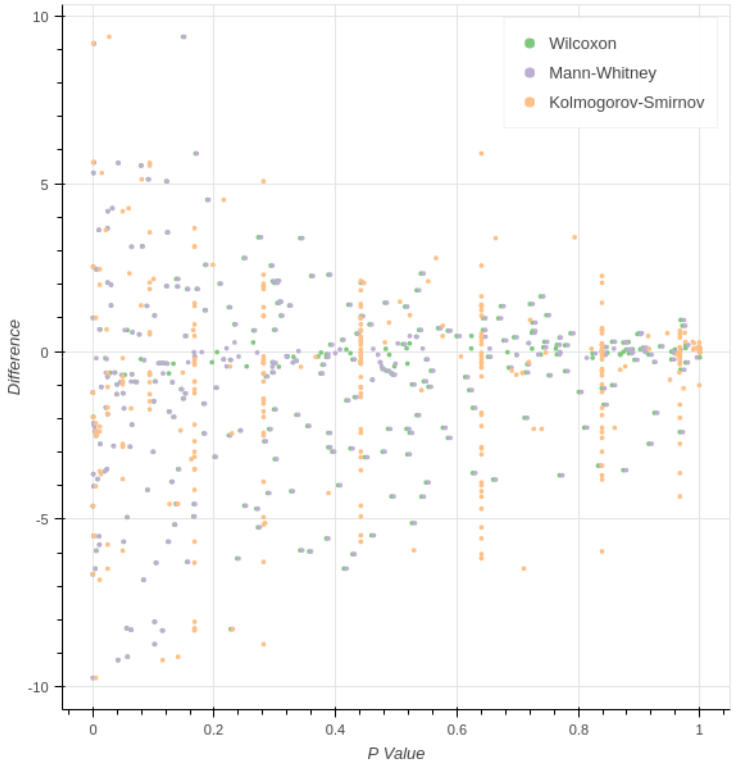

FIGURE 5.2: Distribution of difference and p-value of climate divisions in contintental U.S. when maximum tempertaure is greater than 95 °F

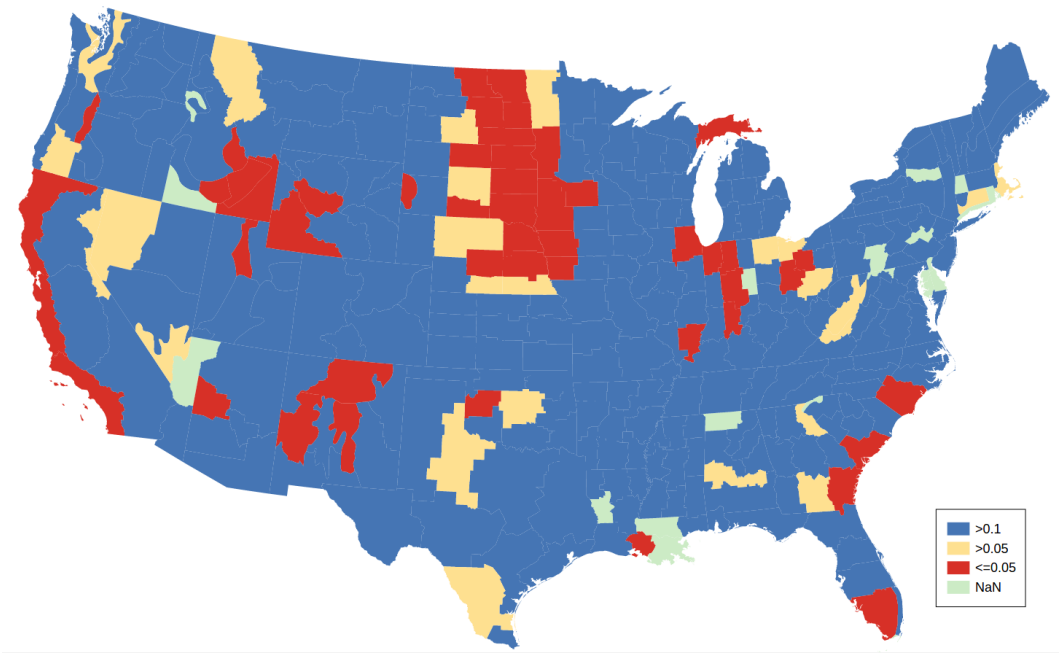

FIGURE 5.3: Distribution of Wilcoxon test p-value when maximum temperature is greater than 95 °F

5.2 Maximum Temperature is Greater than 95 °F

Figure 5.2 is the scatter graph of all climate divisions in the continental U.S. for the number of days when maximum temperature is greater than 95°F. x axis represents p-value, y axis represents the difference between the means of two time-series extreme climate dataset (1946 - 1980 and 1981 - 2015), and green points represent Wilcoxon Test, purple points represent Mann-Whitney Test, brown points represent Kolmogorov-Smirnov Test.

In the figure, p-value close to 0 means that the frequencies datasets is very significant, and difference being positive means the frequencies is indicating a increasing trend, otherwise it is in a decreasing trend.

Figure 5.3, Figure 5.4, Figure 5.5, and Figure 5.6 display p-values and differences of frequencies when maximum temperature is greater than 95 °F. It is seen easily that frequencies of extreme high maximum temperature in inland areas decreases sharply, and that in part of westren region increases relatively.

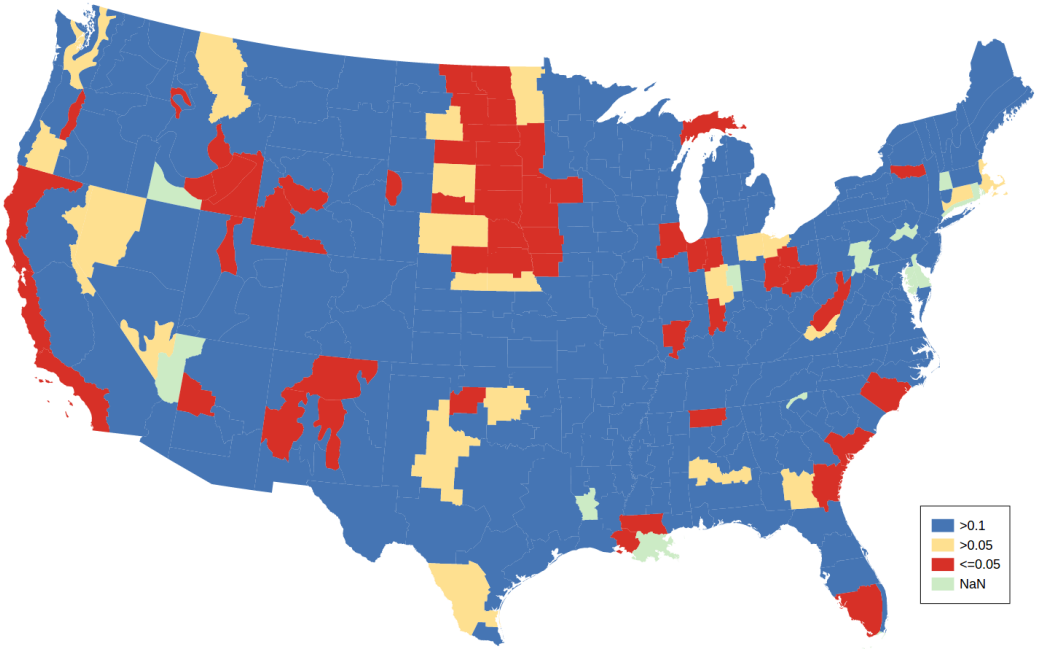

FIGURE 5.4: Distribution of Mann-Whitney test p-value when maximum temperature is greater than 95 °F

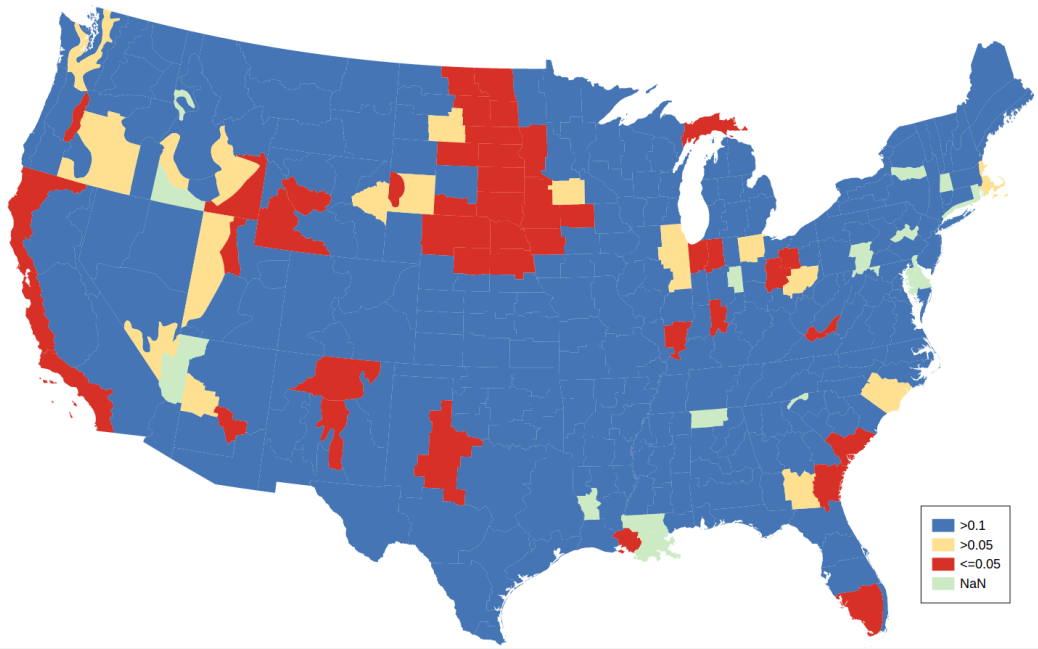

FIGURE 5.5: Distribution of Kolmogorov-Smirnov test p-value when maximum temperature is greater than 95 °F

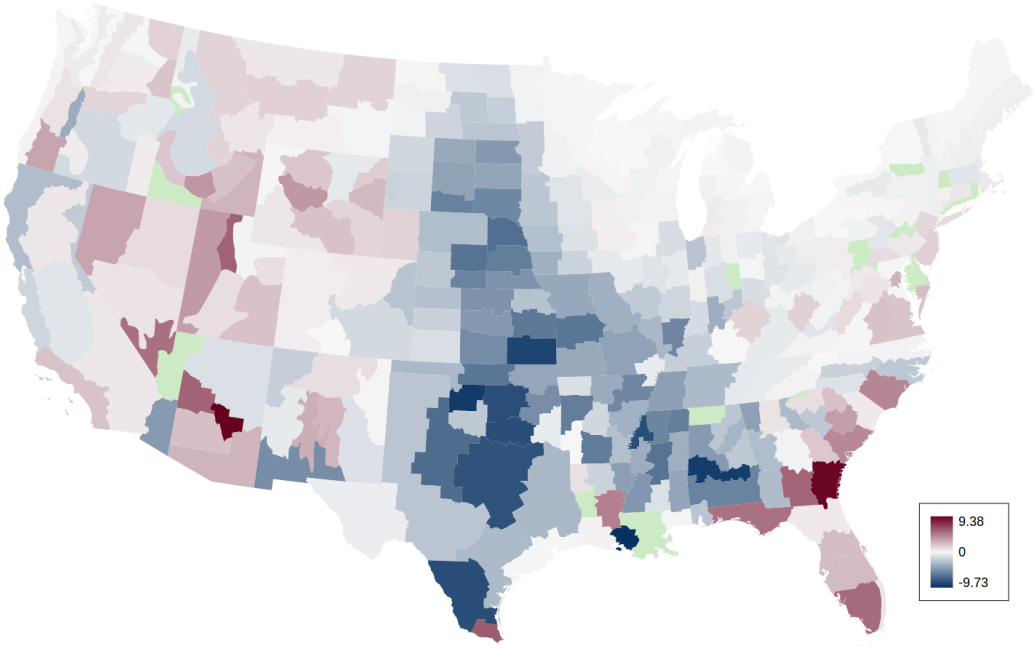

FIGURE 5.6: Distribution of difference when maximum temperature is greater than 95 °F

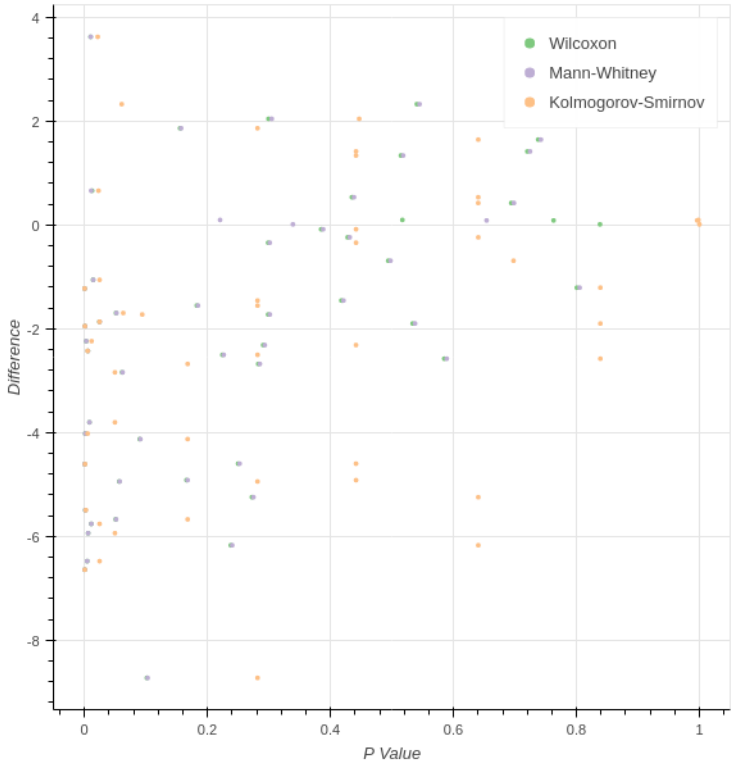

FIGURE 5.7: Distribution of difference and p-value of climate divisions in high plains region when maximum temperature is greater than 95 °F

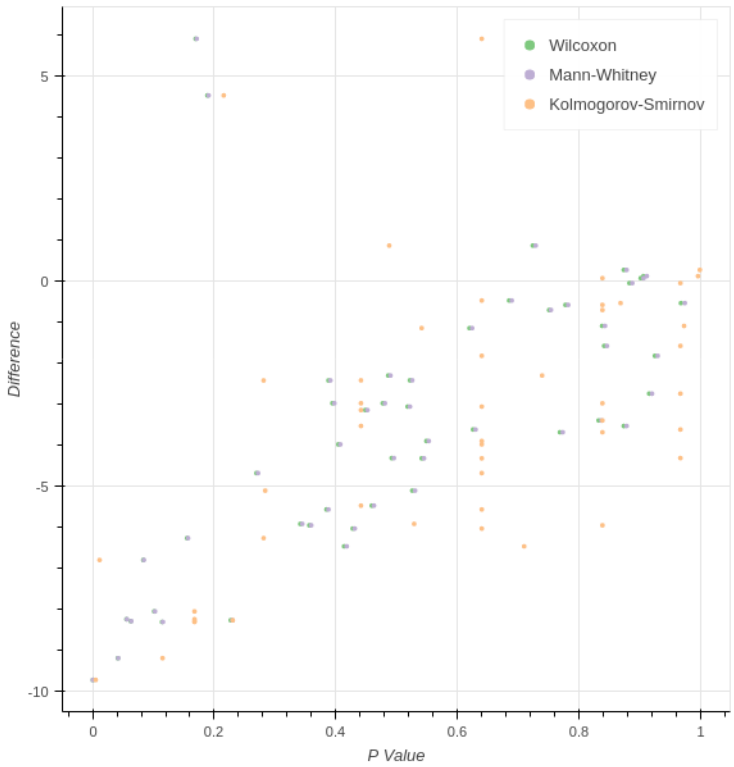

FIGURE 5.8: Distribution of difference and p-value of climate divisions in southern region when maximum temperature is greater than 95 °F

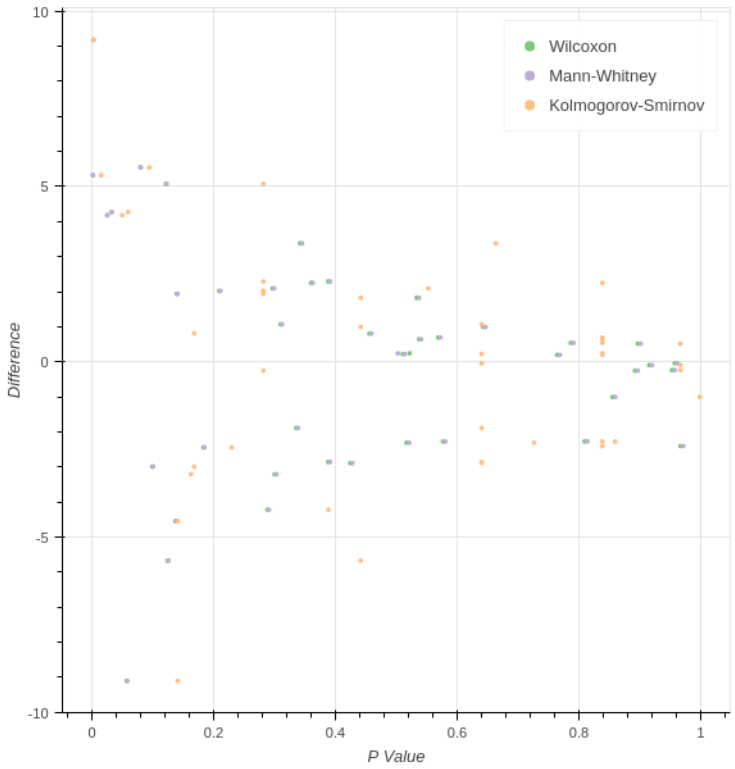

FIGURE 5.9: Distribution of difference and p-value of climate divisions in southeast region when maximum temperature is greater than 95 °F

Figures 5.7 and 5.8 are the scatter figures in high plains region and southern region. In these two region, the frequencies of extreme high maximum temperature decreases very sharply. In other words, the frequencies of extreme high maximum temperature of inland areas incline to decrease rapidly in the continental U.S..

The Figure 5.9 and Figure 5.10 are the scatter figure in southeast region and western region. In these two region, the frequencies of extreme high maximum temperature increases relatively.

The change of maximum temperature is uneven in the continental U.S. and has negative growth generally. especially, inland areas have strong downward trends and in contrast the frequencies in west coastal areas and southeast region increase relatively.

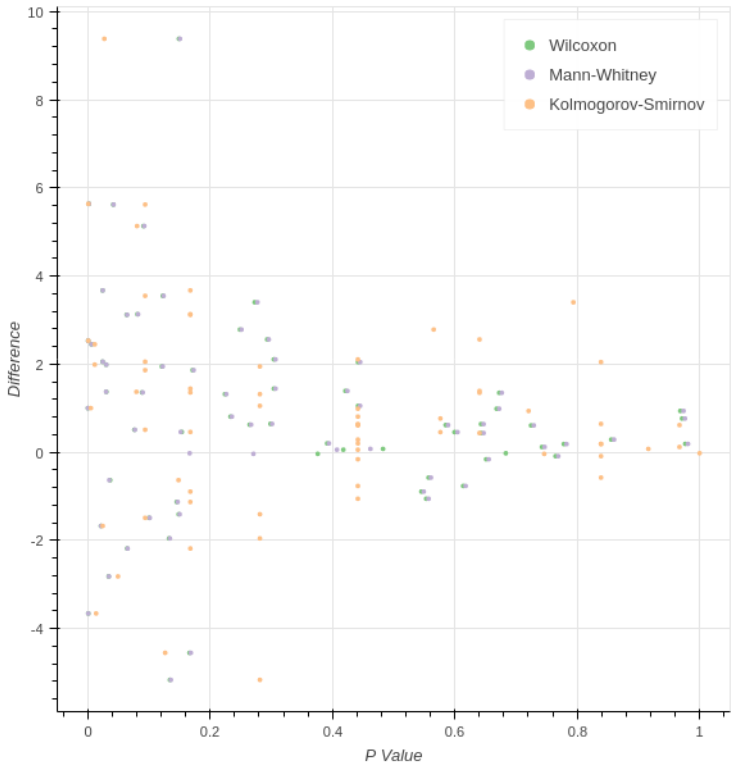

FIGURE 5.10: Distribution of difference and p-value of climate divisions in western region when maximum temperature is greater than 95 °F

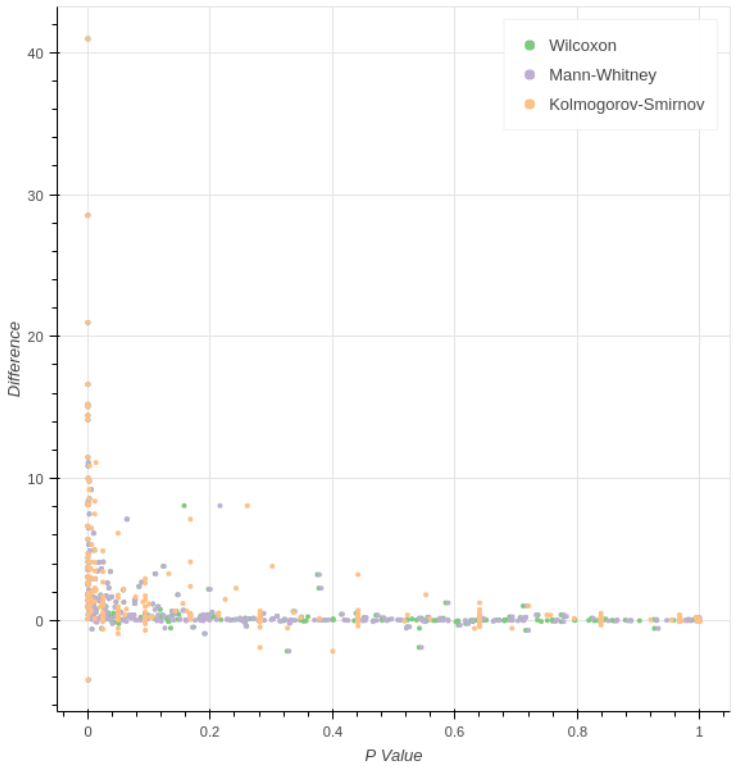

FIGURE 5.11: Distribution of difference and p-value of climate divisions in contintental U.S. when minimum temperature is greater than 75 °F

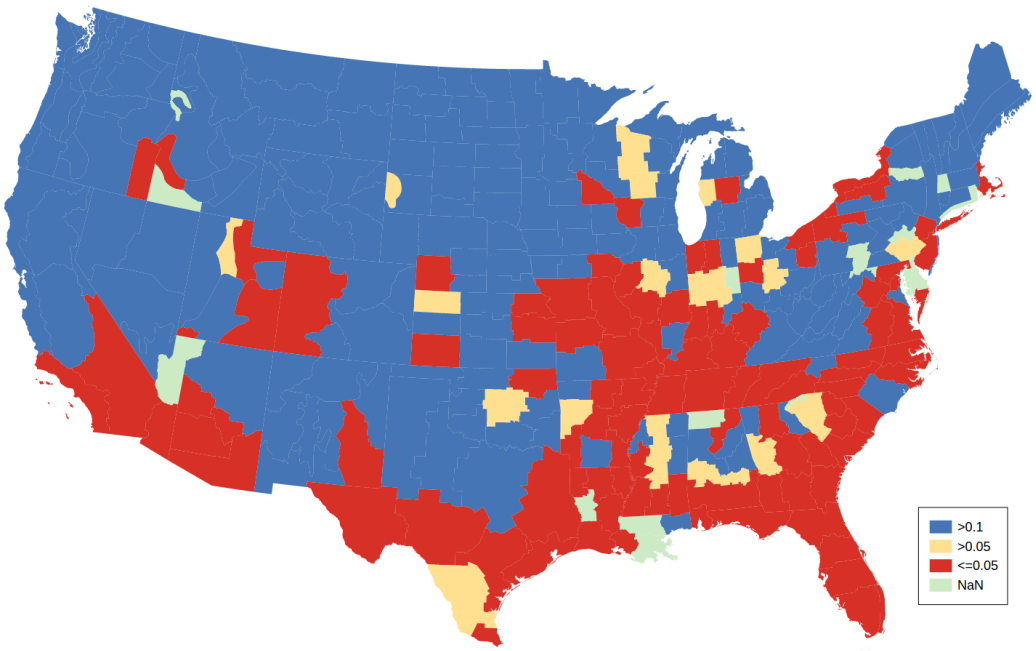

FIGURE 5.12: Distribution of Wilcoxon test p-value when minimum temperature is greater than 75 °F

5.3 Minimum Temperature is Greater than 75 °F

Figure 5.11 is the scatter graph of the all climate divisions in the continental U.S. for frequencies when minimum temperature is greater than 75 °F. In the figure, most of p-values is close to 0, and most of differences is greater than 0. It is shown that the frequencies of extreme high minimum temperature increase sharply.

Figure 5.12, Figure 5.13, Figure 5.14, and Figure 5.15 display p-values and differences of frequencies when minimum temperature is greater than 75 °F. As shown, the frequencies of extreme high minimum temperature mostly increase between two time period, and in contrast some climdiv divisions in inland areas decrease.

The Figure 5.16 and Figure 5.17 are the scatter figure in southern region and high plains region. In these two regions, the upward trend of frequencies of extreme high minimum temperature do not show as much change as that in Figure 5.11.

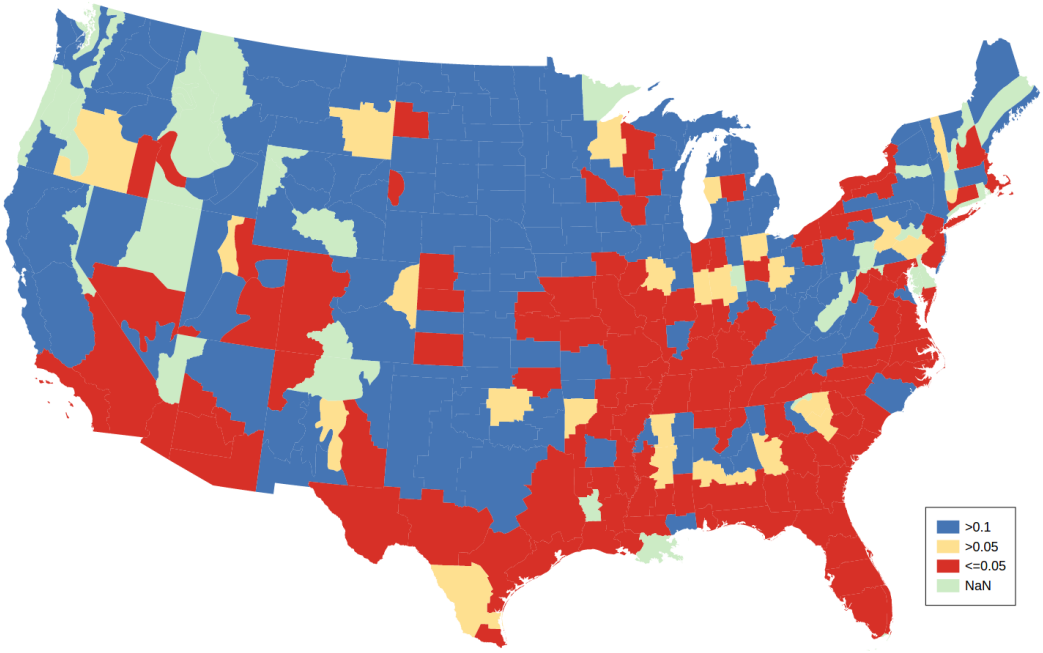

FIGURE 5.13: Distribution of Mann-Whitney test p-value when minimum temperature is greater than 75 °F

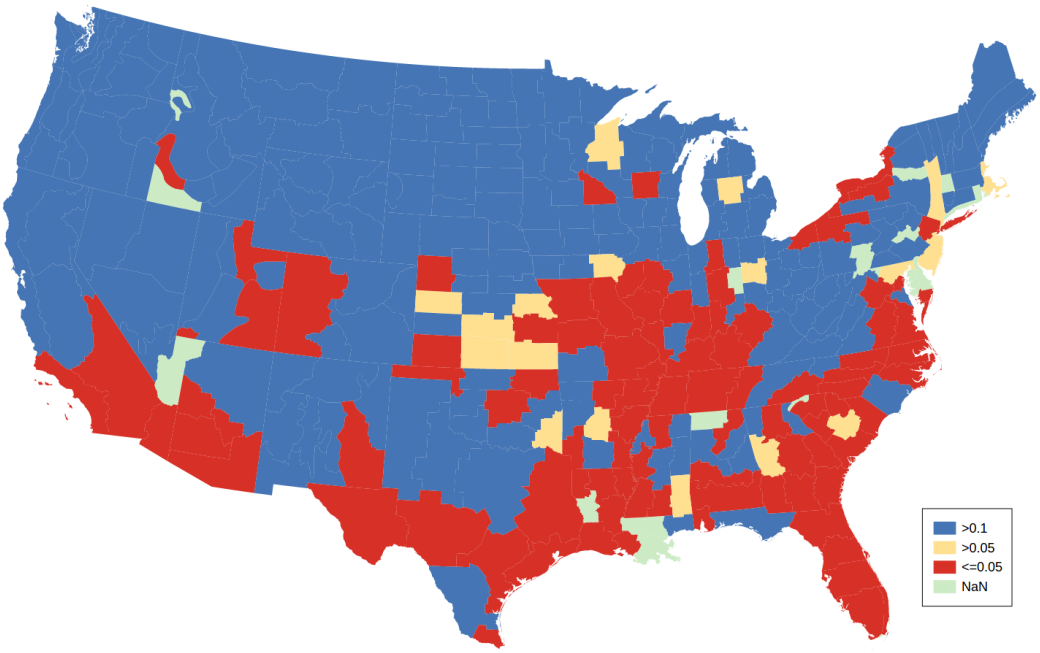

FIGURE 5.14: Distribution of Kolmogorov-Smirnov test P-value when minimum temperature is greater than 75 °F

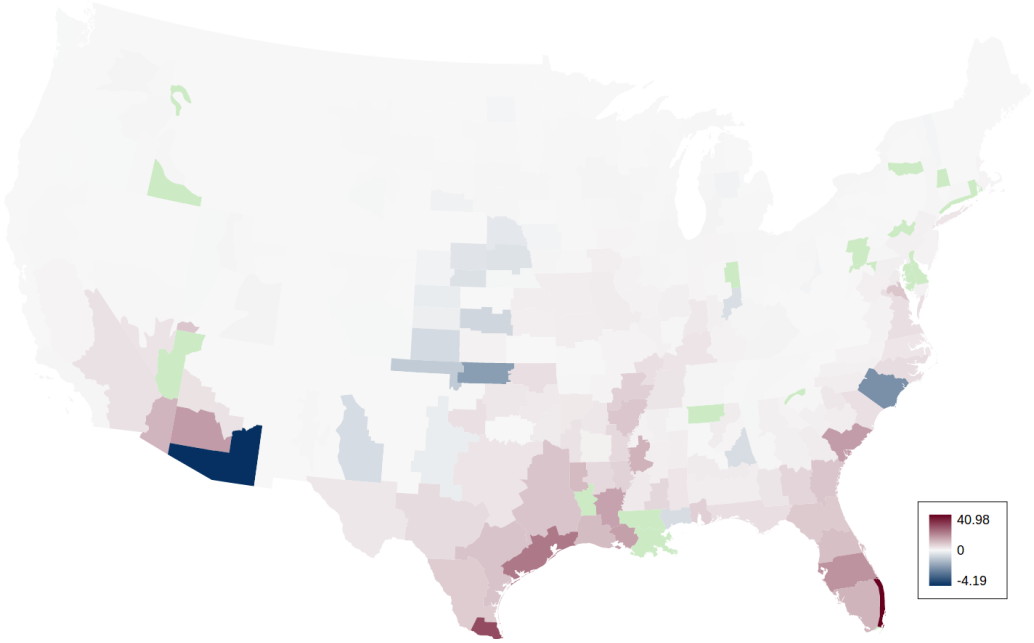

FIGURE 5.15: Distribution of difference when minimum temperature is greater than 75 °F

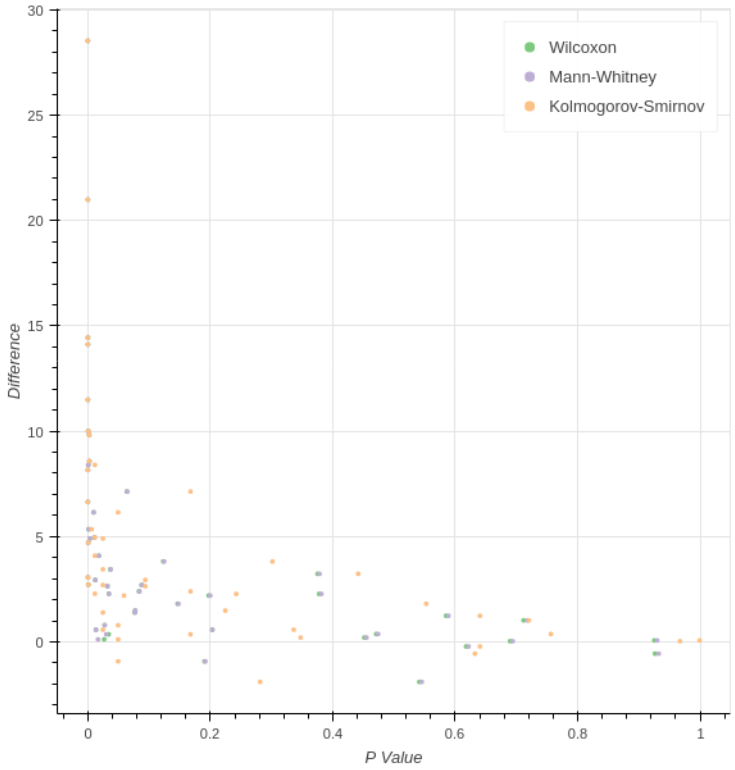

FIGURE 5.16: Distribution of difference and p-value of climate divisions in southern region when minimum tempertaure is greater than 75 °F

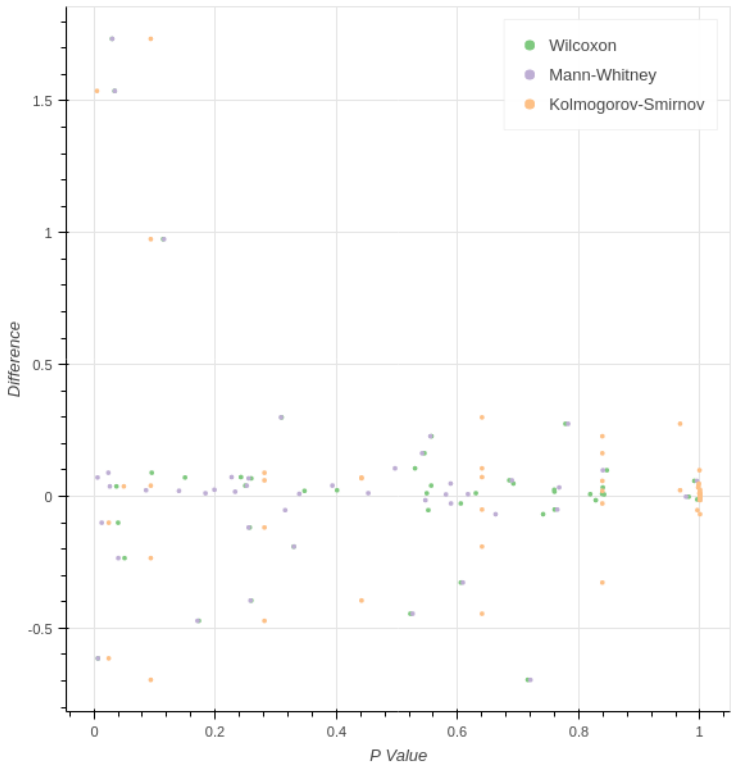

FIGURE 5.17: Distribution of difference and p-value of climate divisions in high plains region when minimum tempertaure is greater than 75 °F

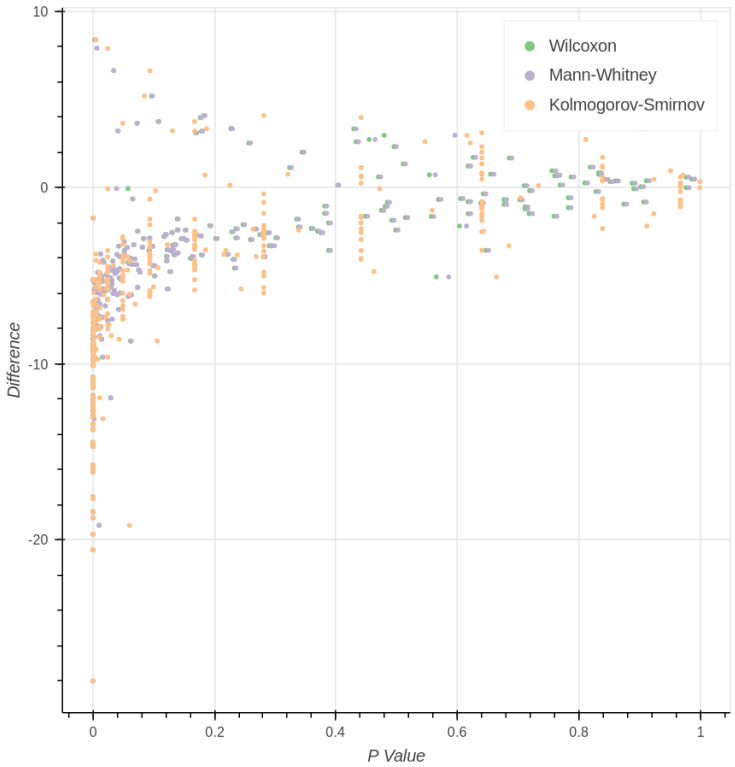

FIGURE 5.18: Distribution of difference and p-value of climate divisions in the contintental U.S. when minimum tempertaure is lower than 32 °F

Compared with extreme high maximum temperature, the frequencies of extreme high minimum temperature increase, and in contrast inlands areas remain mostly steady.

5.4 Minimum Temperature is Lower than 32 °F

Figure 5.18 is the scatter graph of the all climate divisions in the continental U.S. for frequencies when minimum temperature is lower than 32 °F. In the figure, it is shown that the number of days with extreme low minimum temperatures decrease sharply.

Figure 5.19, Figure 5.20, Figure 5.21, and Figure 5.22 display p-values and differences of frequencies when minimum temperature is lower than 32 °F. As shown, the frequencies of extreme low minimum temperature mostly decrease, and frequencies

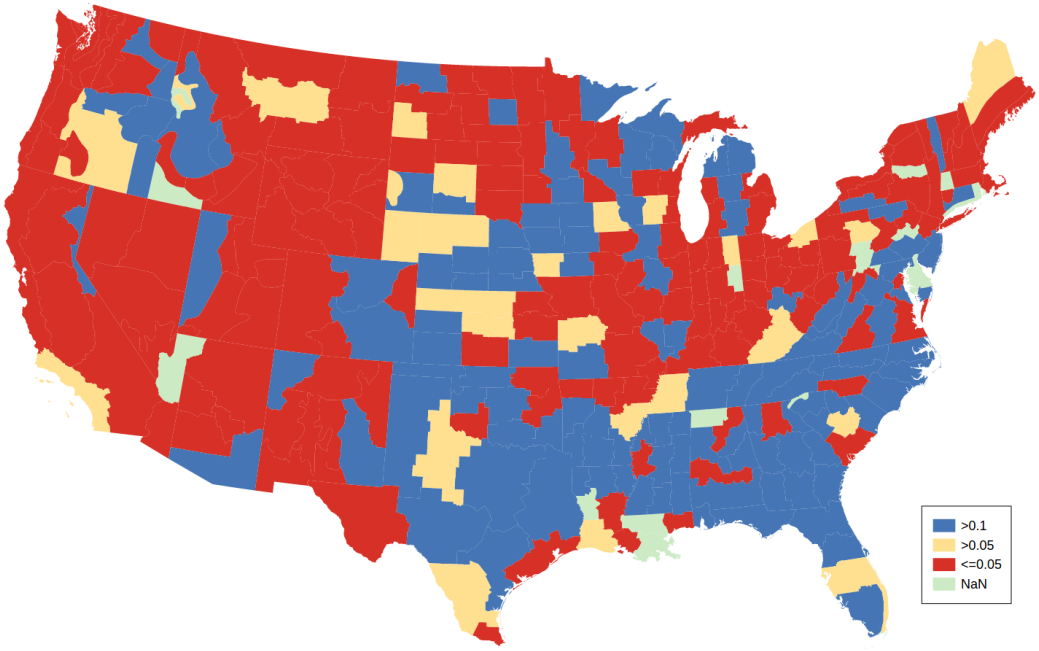

FIGURE 5.19: Distribution of Wilcoxon test p-value when minimum temperature is lower than 32 °F

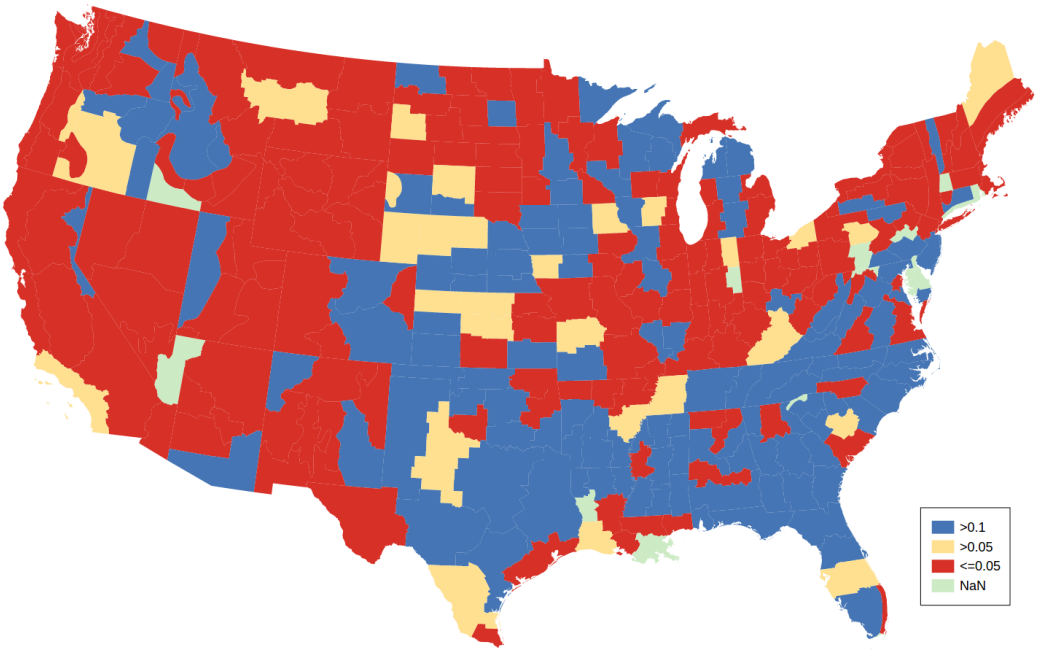

FIGURE 5.20: Distribution of Mann-Whitney test p-value when minimum temperature is lower than 32 °F

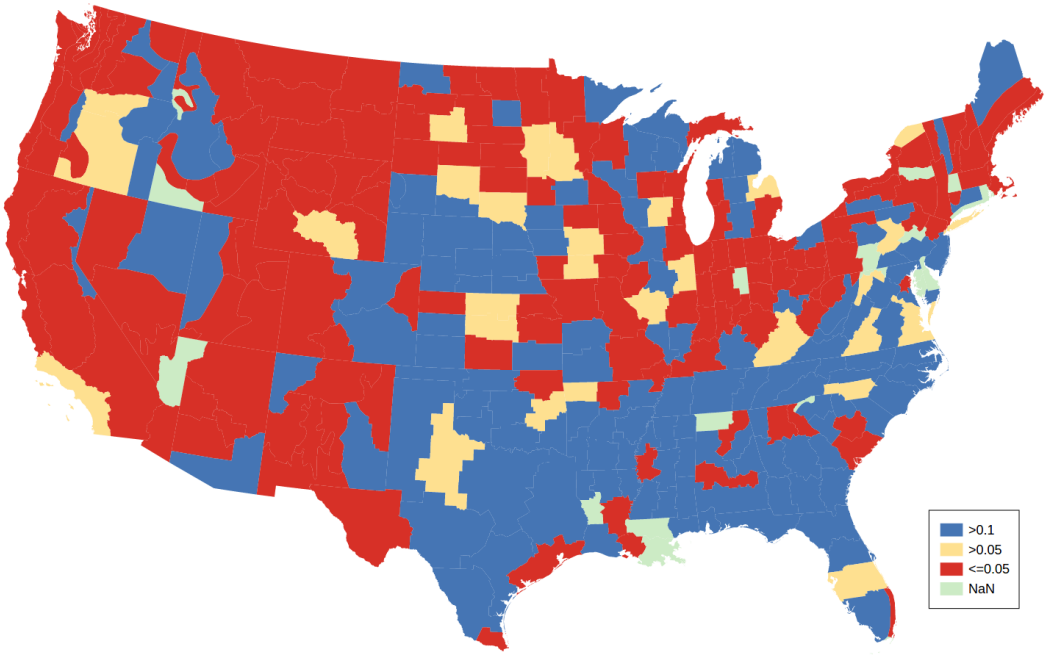

FIGURE 5.21: Distribution of Kolmogorov-Smirnov test p-value when minimum temperature is lower than 32 °F

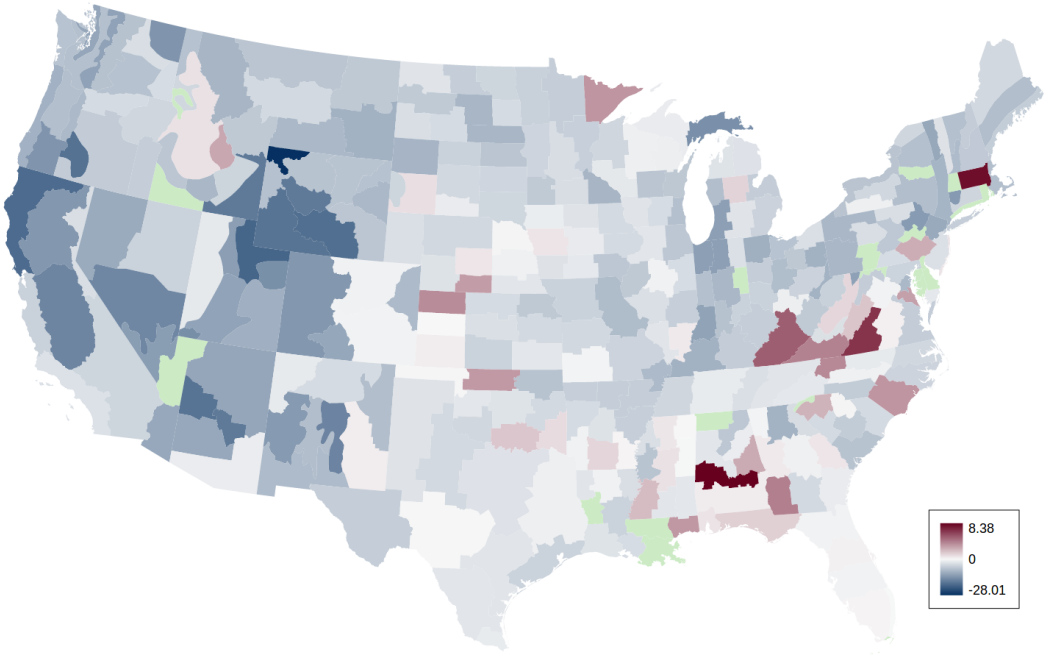

FIGURE 5.22: Distribution of difference when minimum temperature is lower than 32 °F

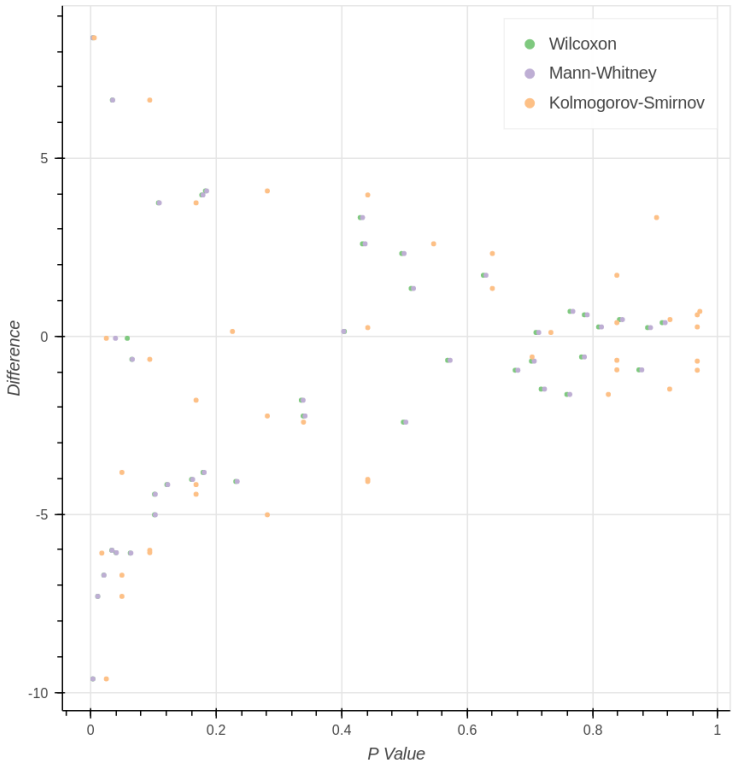

FIGURE 5.23: Distribution of difference and p-value of climate divisions in southeast region when minimum tempertaure is lower than 32 °F

of extreme low minimum temperature in west coastal areas decrease more rapidly than that in east coastal areas.

The Figure 5.23 is the scatter figure in southeast region. In this region, the number of days with extreme high maximum temperatures mostly hold steady. In contrast, in Figure 5.24, the frequencies of extreme low minimum temperature decrease rapidly. Thus, the frequencies of extreme low minimum temperature decrease.

5.5 Minimum Temperature is Lower than 0 °F

Figure 5.25 is the scatter graph of the all climate divisions in the continental U.S. for frequencies when minimum temperature is lower than 0 °F. In the figure, it is shown that the number of days with extreme low minimum temperature decreases sharply.

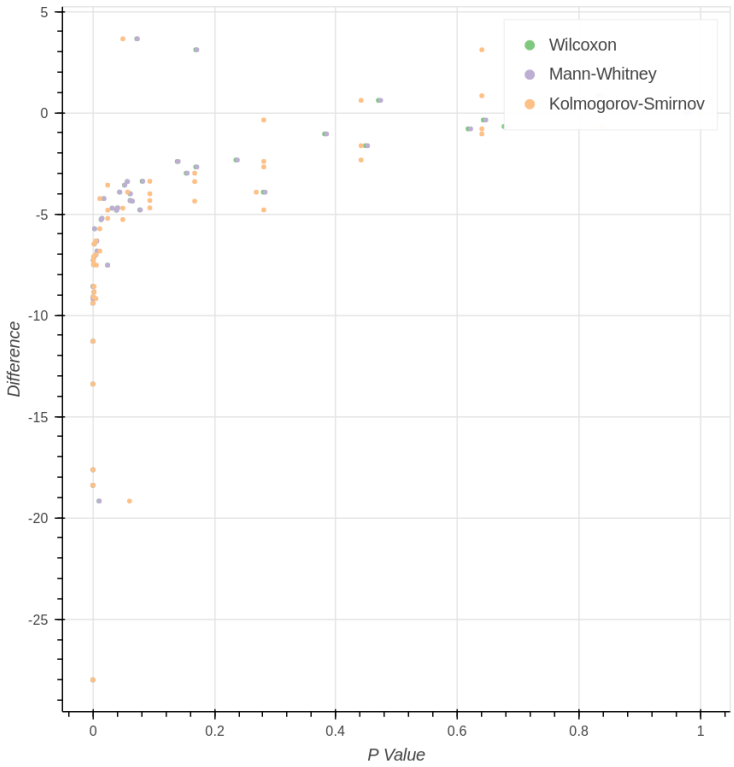

FIGURE 5.24: Distribution of difference and p-value of climate divisions in western region when minimum tempertaure is lower than 32 °F

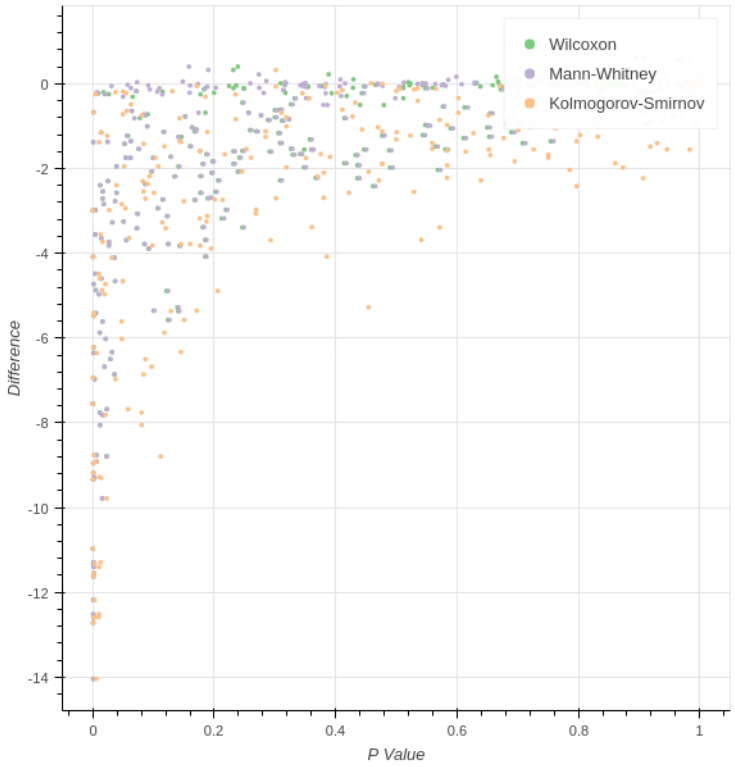

FIGURE 5.25: Distribution of difference and p-value of climate divisions in the contintental U.S. when minimum tempertaure is lower than 0 °F

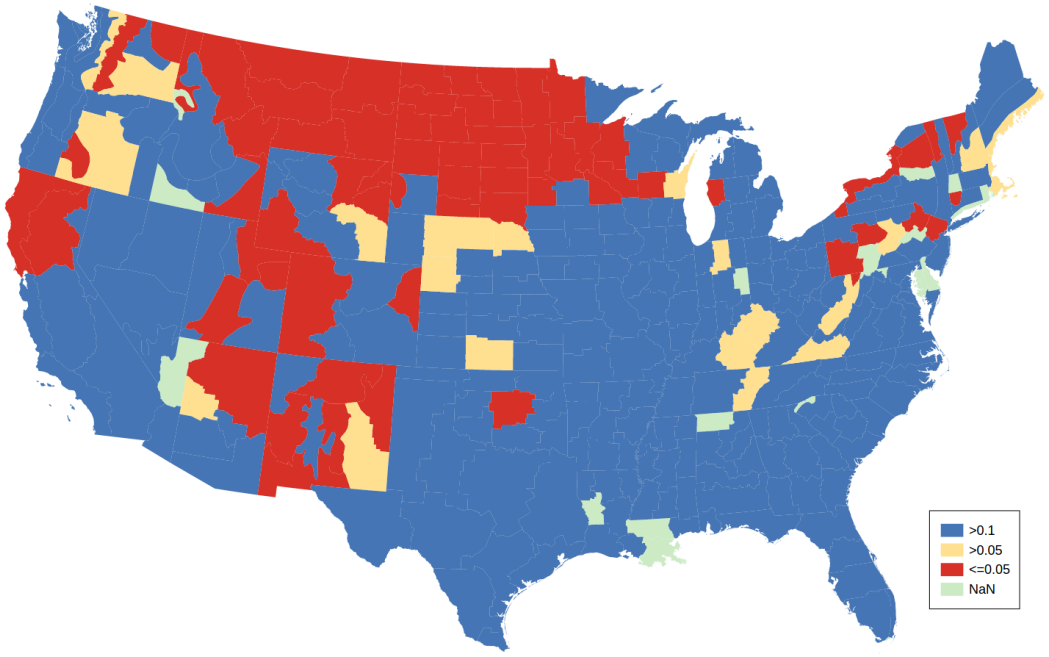

FIGURE 5.26: Distribution of Wilcoxon test p-value when minimum temperature is lower than 0 °F

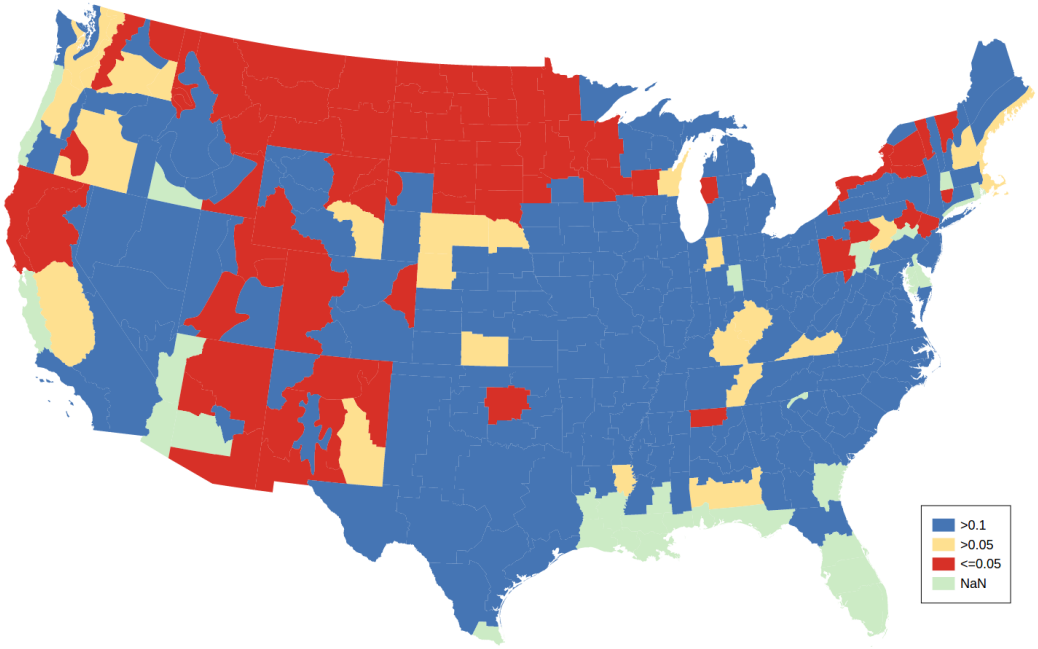

FIGURE 5.27: Distribution of Mann-Whitney test p-value when minimum temperature is lower than 0 °F

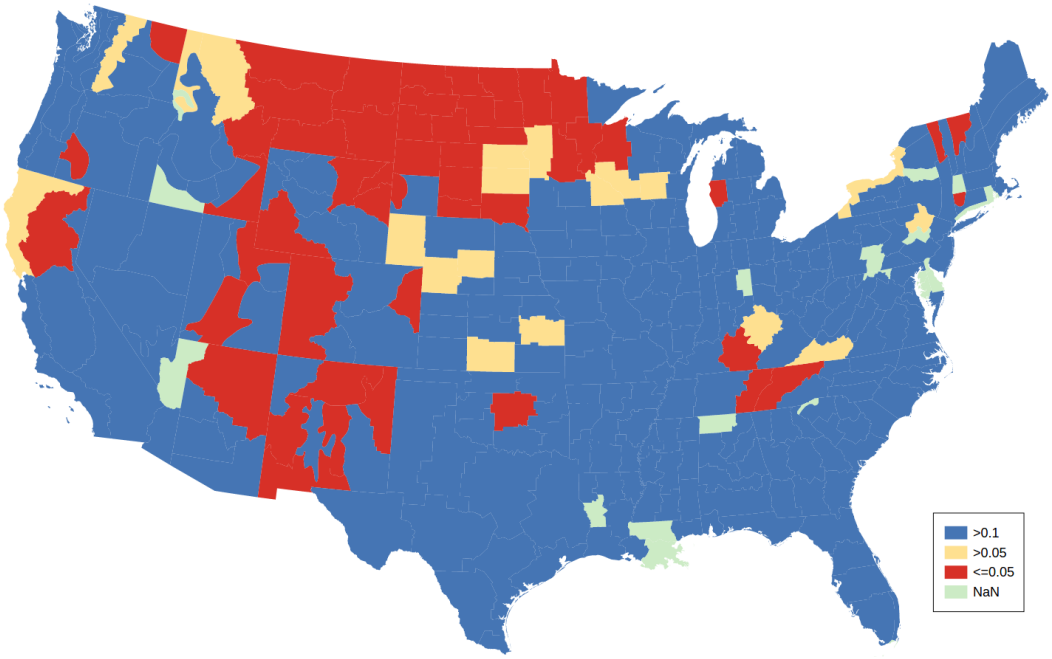

FIGURE 5.28: Distribution of Kolmogorov-Smirnov test p-value when minimum temperature is lower than 0 °F

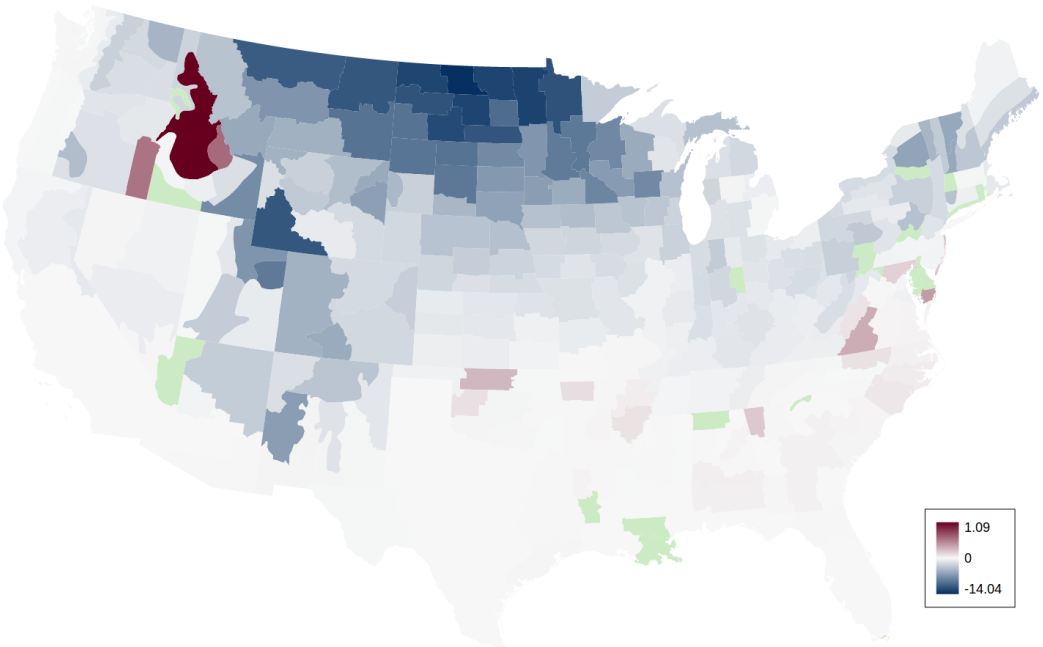

FIGURE 5.29: Distribution of difference when minimum temperature is lower than 0 °F

Figure 5.26, Figure 5.27, Figure 5.28, and Figure 5.29 display p-values and differences of frequencies when minimum temperature is lower than 0 °F. The frequencies of extreme low minimum temperature mostly show declines.

Figure 5.30 is the scatter figure in southeast region, which shows that the number of days with extreme high maximum temperature hold relatively steady. In contrast, in Figure 5.31, the frequencies of extreme low minimum temperature decrease.

Frequencies of extreme minimum temperature at as below 0 decrease dramatically over the two time periods. However, inlands areas shows little change for extreme low minimum temperature. Frequencies of extreme low minimum temperature in the west coastal area decreases more rapidly than that in the east coastal area.

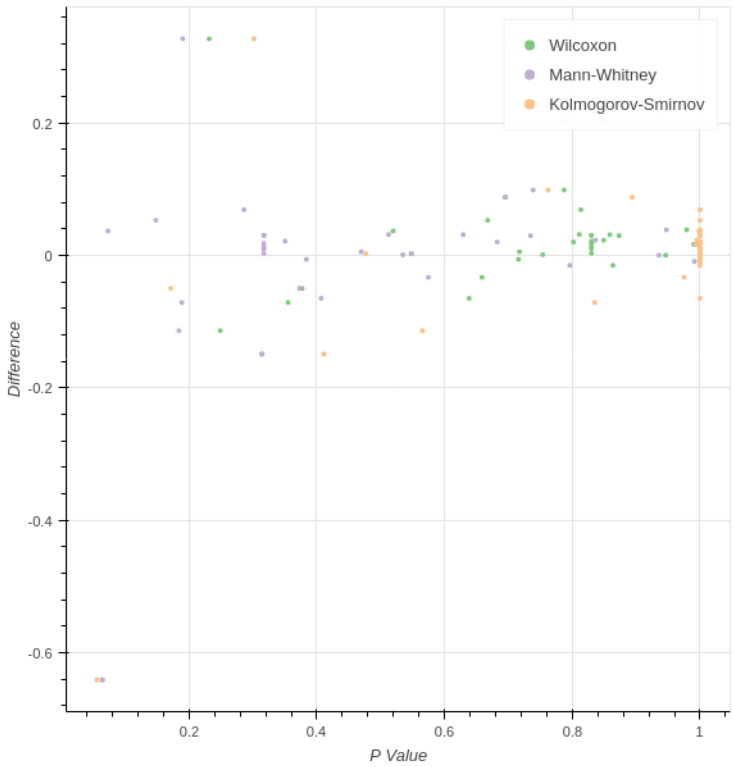

FIGURE 5.30: Distribution of difference and p-value of climate divisions in south-east region when minimum tempertaure is lower than 0 °F

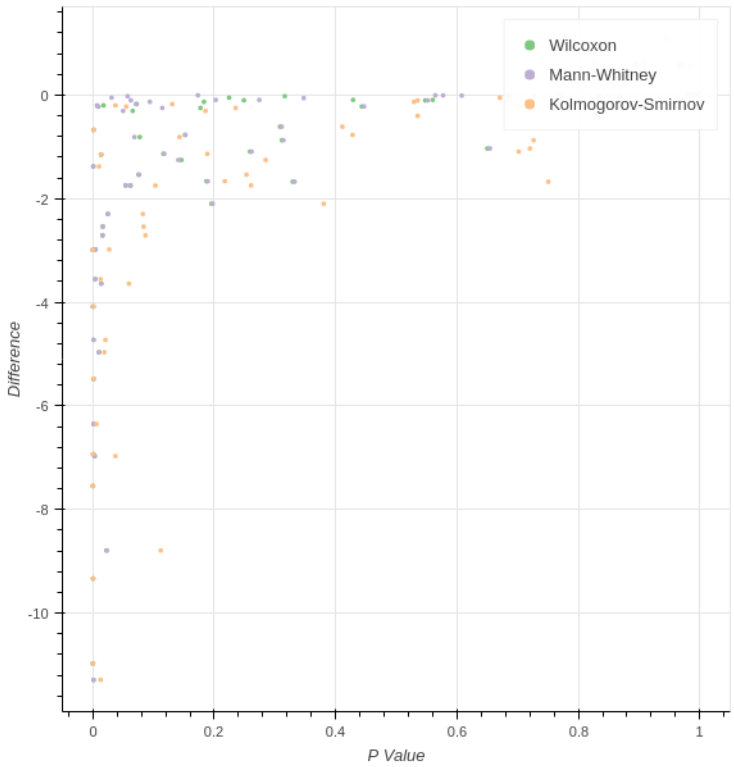

FIGURE 5.31: Distribution of difference and p-value of climate divisions in western region when minimum tempertaure is lower than 0 °F

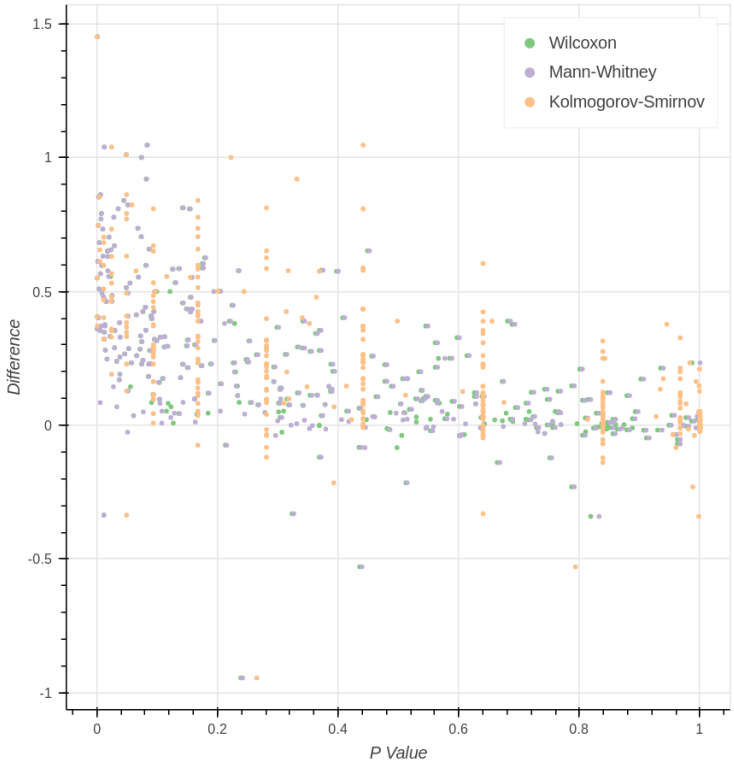

FIGURE 5.32: Distribution of difference and p-value of climate divisions in the continental U.S. when precipitation is greater than 2 inches

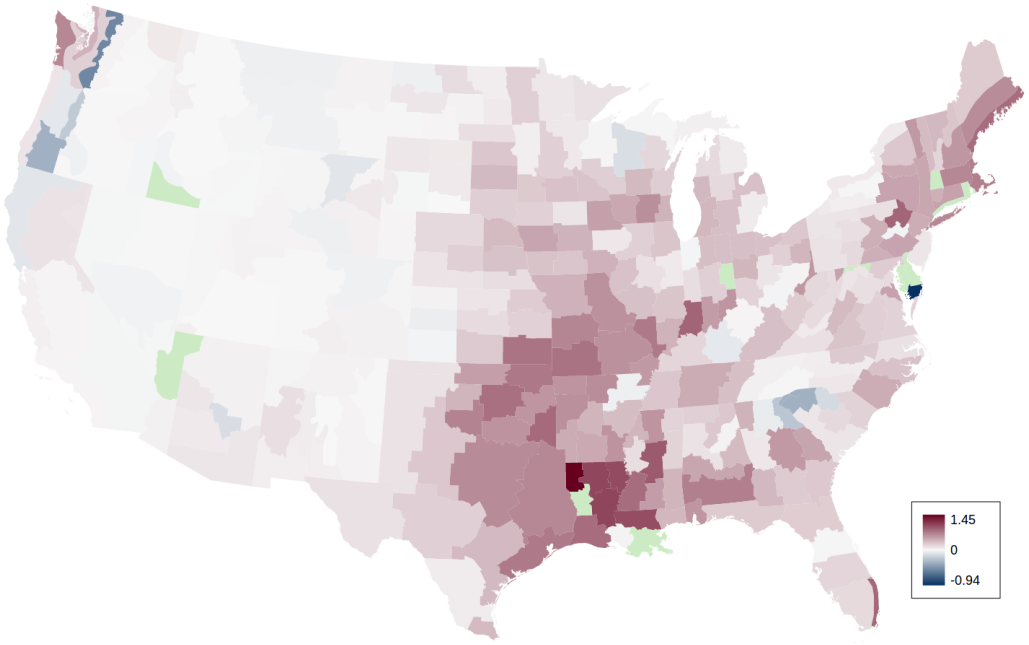

FIGURE 5.33: Distribution of difference when precipitation is greater than 2 inches

5.6 Precipitation is Greater than 2 inches

Figure 5.32 is the scatter graph of the all climate divisions in the continental U.S. for frequencies when precipitation is greater than 2 inches. In the figure, it is shown that frequencies of extreme high precipitation increases mostly in south region.

Figure 5.33 displays differences of frequencies when precipitation is greater than 2 inches. Precipitation increases clearly at most locations, and precipitation in inland areas and east coastal area increase more rapidly than that in the west coastal area.

The Figure 5.34 is the scatter figure in western region. In this region, the frequencies of extreme high precipitation show little change over time.

5.7 Annual Total Precipitation

Figure 5.35 is the scatter graph of the all climate divisions in the continental U.S. for the total annual precipitation. In the figure, it is shown that total annual precipitation is increaseing most of the country.

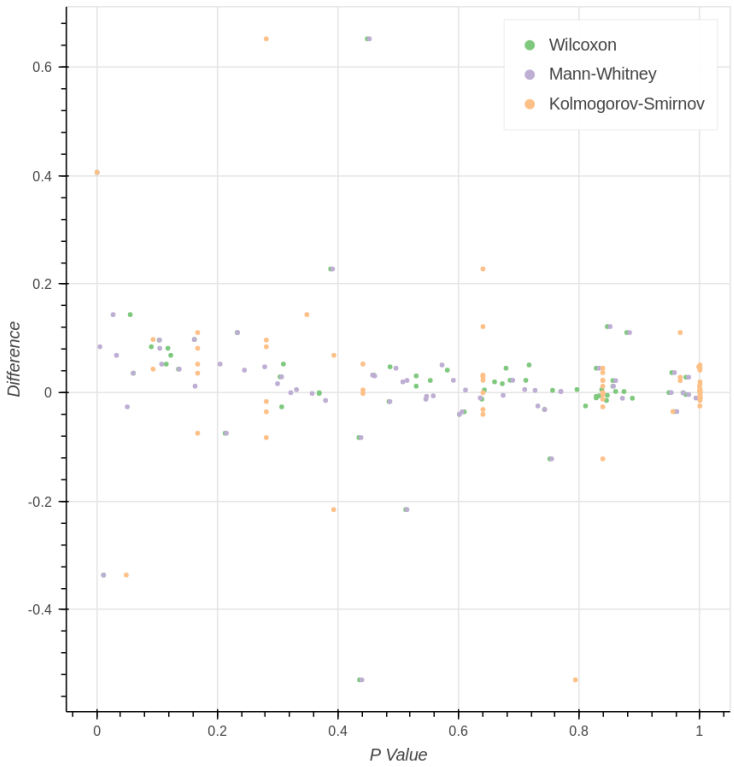

FIGURE 5.34: Distribution of difference and p-value of climate divisions in western region when precipitation is greater than 2 inches

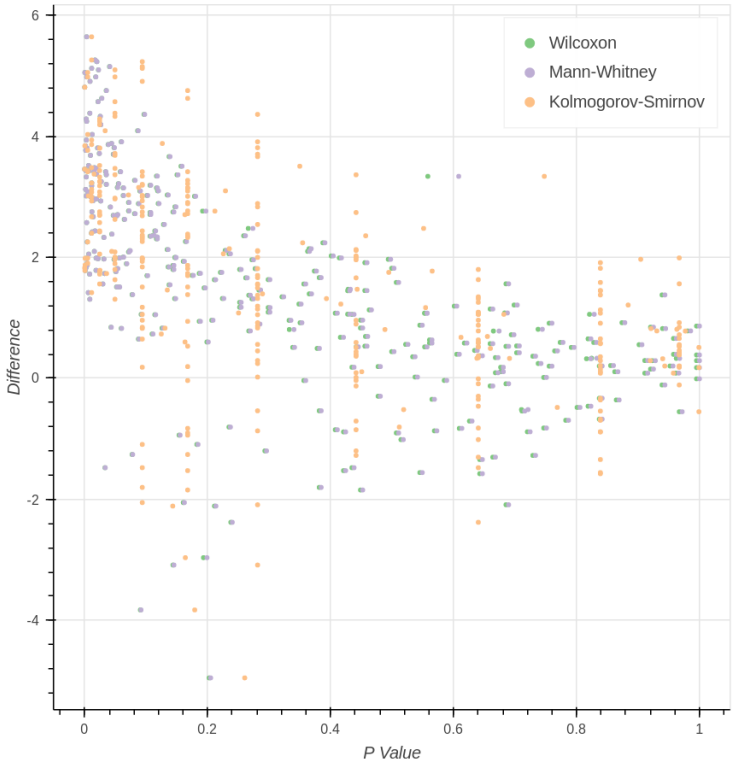

FIGURE 5.35: Distribution of difference and p-value of total annual precipitation in climate divisions in the Continental U.S.

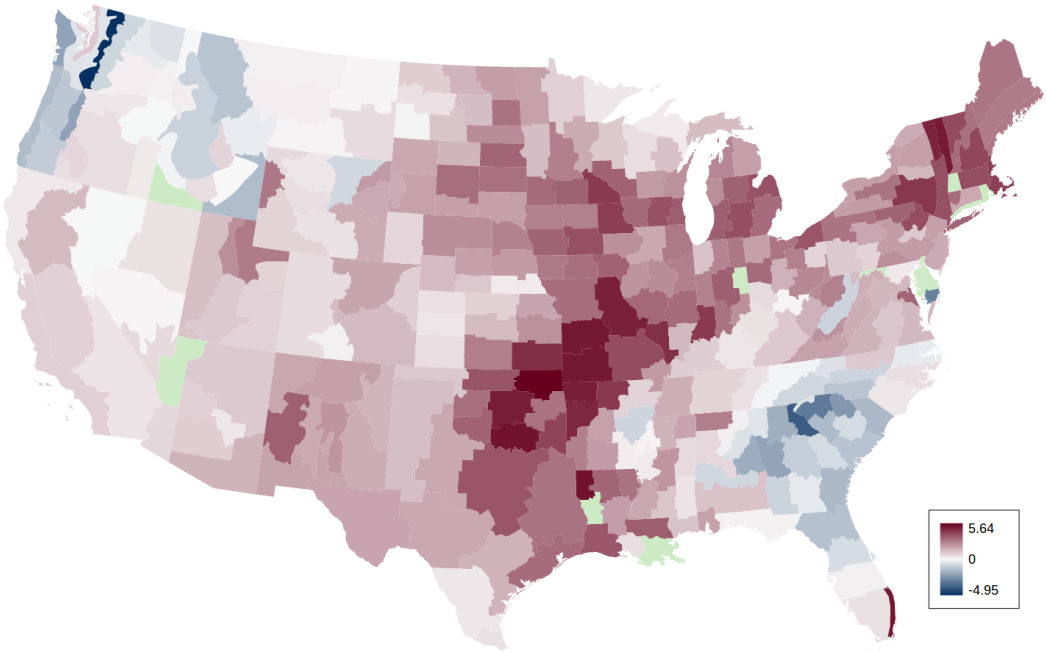

FIGURE 5.36: Distribution of difference of total annual precipitation

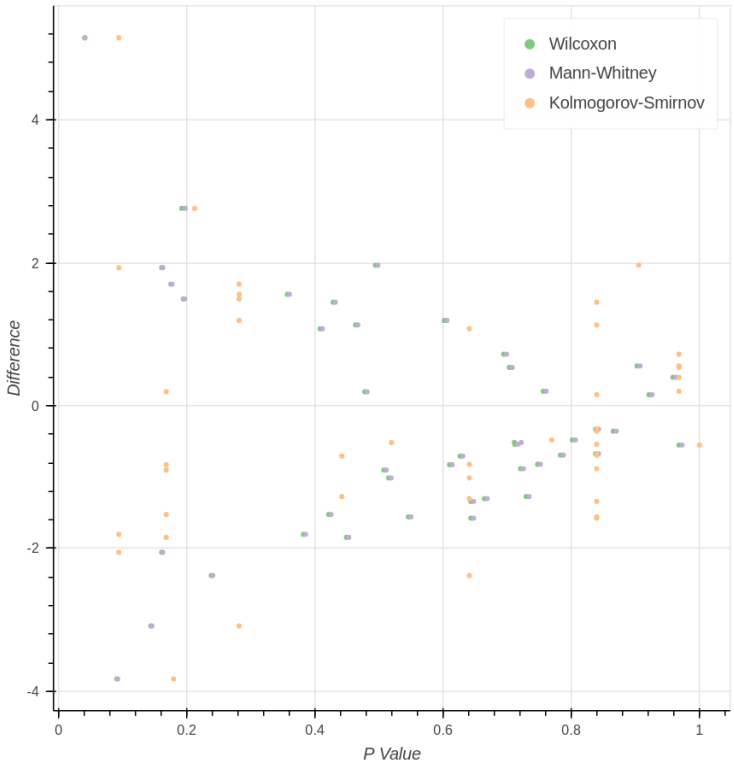

FIGURE 5.37: Distribution of difference and p-value of total annual precipitation in climate divisions in southeast region

Figure 5.36 displays the differences of total annual precipitation. As shown, precipitation increased in inland areas from Oklahoma to Missouri and northeast region increase more rapidly than that in other areas. In contrast, precipitation decreased in northwest region and southeast region.

The Figure 5.37 and Figure 5.38 are the scatter figures in southeast region and western region. In these regions, the total annual precipitation remained relatively steady.

Generally, the frequencies of extreme high precipitaion and total annual precipitation have upward trends. Compared with west coastal area and southeast region, precipitation in inland areas and northeast region increase more dramatically.

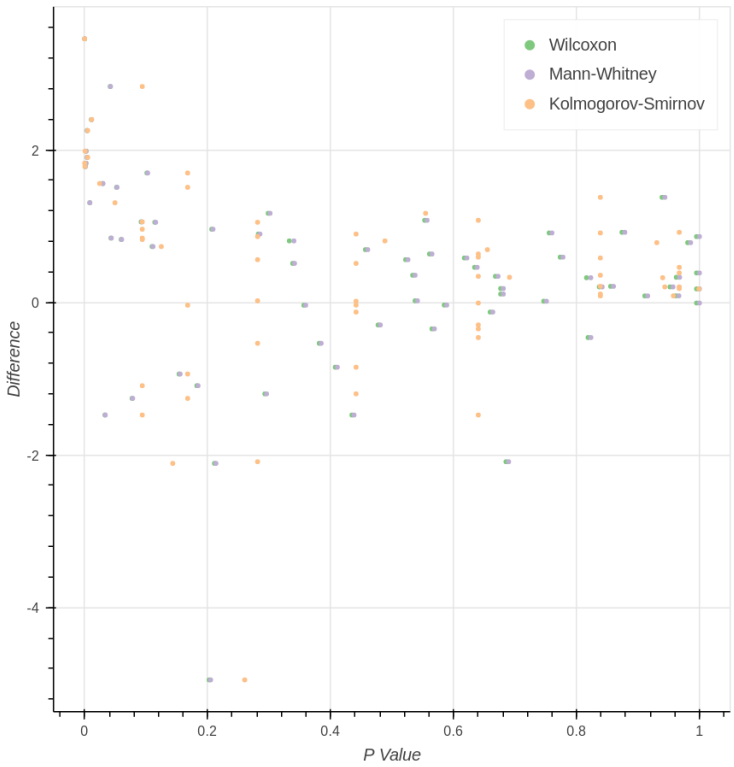

FIGURE 5.38: Distribution of difference and p-value of total annual precipitation in climate divisions in western region

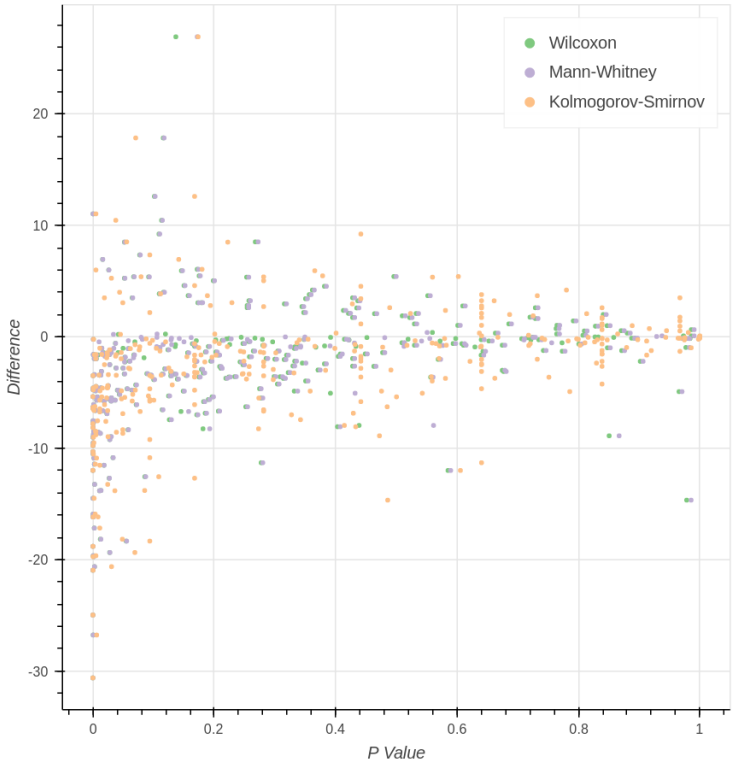

FIGURE 5.39: Distribution of difference and p-value of total annual snowfall in climate divisions in the continental U.S.

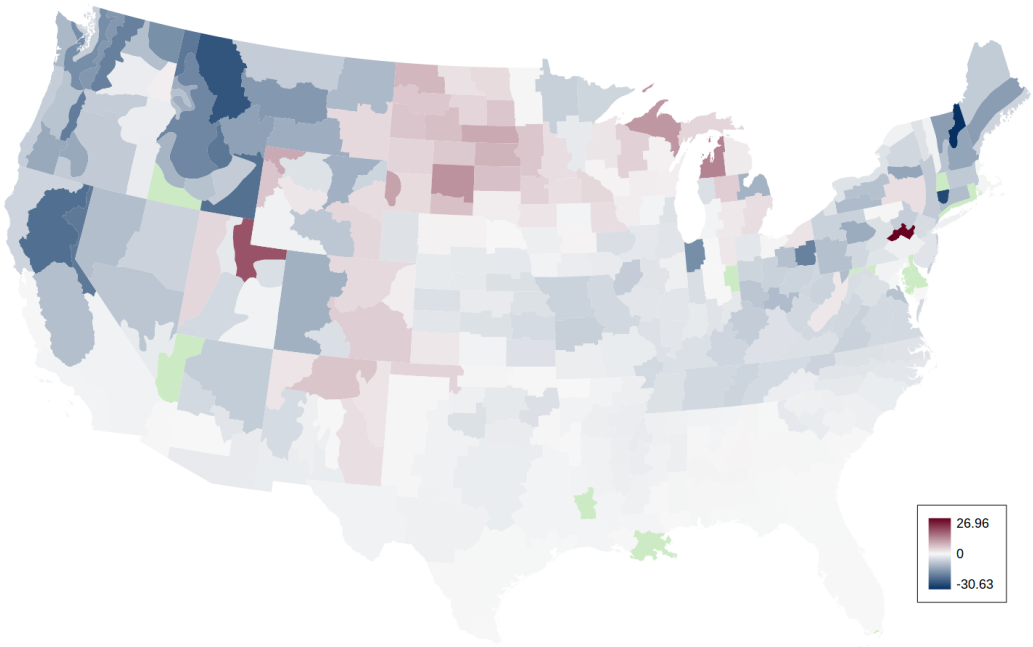

FIGURE 5.40: Distribution of difference of total annual snowfall

5.8 Annual Total Snowfall

Figure 5.39 is the scatter graph of the all climate divisions in the continental U.S. for the total annual snowfall. In the figure, it is shown that total annual snowfall decreases.

Figure 5.40 displays differences of total annual snowfall sum. As shown, snowfall decreases generally, and snowfall in the Pacific Northwest, including Montana, California, Nevada is generally increasing.

The Figure 5.41 and Figure 5.42 are the scatter figures in the high plains region and the midwestern region. In these regions, the total annual snowfall maintains a steady pattern temporally. It means that snowfall in part of the north inland areas increase.

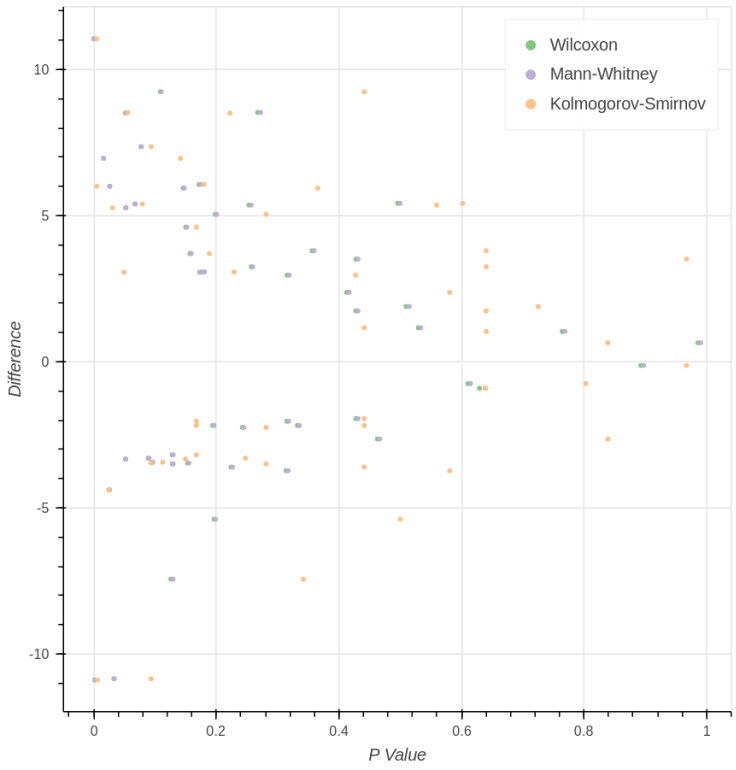

FIGURE 5.41: Distribution of difference and p-value of total annual snowfall in climate divisions in high plains region

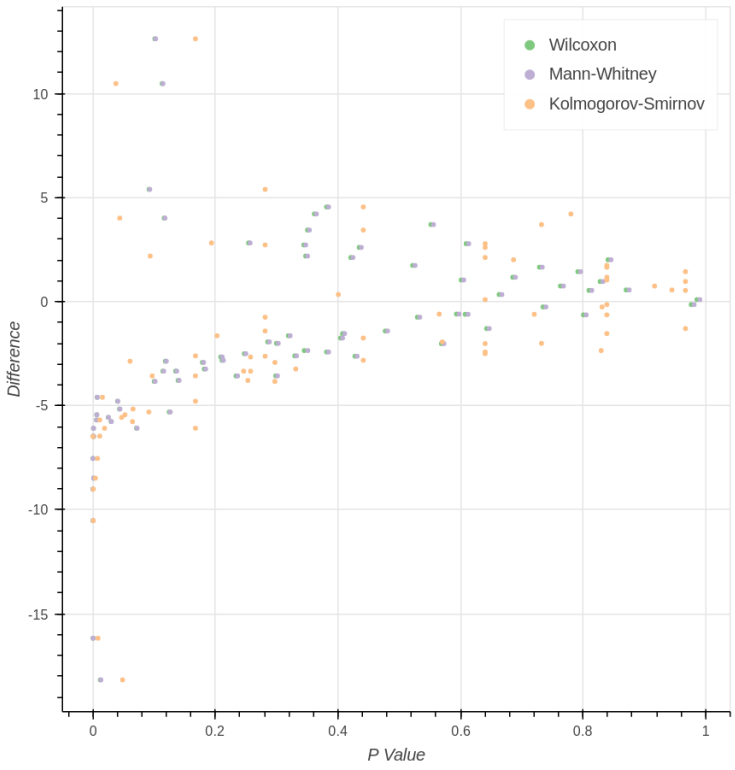

FIGURE 5.42: Distribution of difference and p-value of total annual snowfall in climate divisions in midwestern region

Chapter 6
Conclusions

Overall, by analyzing the extreme climate data from more than 3000 climate stations in the continental U.S. between 1946 and 2015, the conclusions are summarized as follows:

- In the continental United States, the frequencies of extreme high maximum temperature in inland areas have deceased dramatically, which in east coastal areas remain steady, and in contrast, the west coastal areas show increases in frequency. The frequencies of extreme high maximum temperature is generally indicating a decreasing trend.

- In the continental United States, the frequencies of extreme high minimum temperature is indicating a statistically significant, increasing trend.

- In the continental United States, the frequencies of extreme low minimum temperature is showing a statistically significant, decreasing trend.

- Due to the frequency of maximum temperatures decreasing relatively slowly and minimum temperatures increasing sharply, the diurnal temperature range narrows down across continental United States.

- In the continental United States, the frequencies of extreme high precipitation and total annual precipitation indicates an increasing trend, and precipitation in inland areas and northeast region increased more rapidly than that in other areas.

- In the continental United States, total annual snowfall decreased substantially generally.

Bibliography

[1] Rajendra K Pachauri, MR Allen, VR Barros, J Broome, W Cramer, R Christ, JA Church, L Clarke, Q Dahe, P Dasgupta, et al. Climate change 2014: Synthesis report. contribution of working groups i, ii and iii to the fifth assessment report of the intergovernmental panel on climate change. 2014.

[2] U.s. climate divisions. `https://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-divisions.php`, 2016.

[3] Regional climate centers. `https://www.ncdc.noaa.gov/customer-support/partnerships/regional-climate-centers`, 2016.

[4] Jerry M Melillo, Terese TC Richmond, and GW Yohe. Climate change impacts in the united states. *Third National Climate Assessment*, 2014.

[5] Natasha Geiline. Here's what the 2015 drought will do to california's economy. `http://thinkprogress.org/climate/2015/06/05/3666630/california-drought-economic-toll/`, 2015. Online; accessed 13-July-2015.

[6] 2014 winter storm losses expected to total close to $2.5b. `http://www.claimsjournal.com/news/national/2014/11/24/258140.htm`, 2014. Online; accessed 13-July-2015.

[7] Gerald A Meehl, Karl Thomas, David R Easterling, Stanley Changnon, et al. An introduction to trends in extreme weather and climate events: observations, socioeconomic impacts, terrestrial ecological impacts, and model projections. *Bulletin of the American Meteorological Society*, 81(3):413, 2000.

[8] John Theodore Houghton, Geoffrey J Jenkins, and JJ Ephraums. Climate change: the ipcc scientific assessment. *American Scientist;(United States)*, 80(6), 1990.

[9] John T Houghton. *Climate change 1995: The science of climate change: contribution of working group I to the second assessment report of the Intergovernmental Panel on Climate Change*, volume 2. Cambridge University Press, 1996.

[10] Phil D Jones, Mark New, David E Parker, Seelye Martin, and Ignatius G Rigor. Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics*, 37(2):173–199, 1999.

[11] Simon FB Tett, Peter A Stott, Myles R Allen, William J Ingram, and John FB Mitchell. Causes of twentieth-century temperature change near the earth's surface. *Nature*, 399(6736):569–572, 1999.

[12] Michael E Mann, Raymond S Bradley, and Malcolm K Hughes. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, 392(6678):779–787, 1998.

[13] Tom ML Wrigley. *The science of climate change: global and US perspectives*. Pew Center on Global Climate Change, 1999.

[14] Thomas J Crowley. Causes of climate change over the past 1000 years. *Science*, 289(5477):270–277, 2000.

[15] Peter A Stott, SFB Tett, GS Jones, MR Allen, JFB Mitchell, and GJ Jenkins. External control of 20th century temperature by natural and anthropogenic forcings. *Science*, 290(5499):2133–2137, 2000.

[16] Sydney Levitus, John I Antonov, Julian Wang, Thomas L Delworth, Keith W Dixon, and Anthony J Broccoli. Anthropogenic warming of earth's climate system. *Science*, 292(5515):267–270, 2001.

[17] Gerald A Meehl, Francis Zwiers, Jenni Evans, Thomas Knutson, et al. Trends in extreme weather and climate events: issues related to modeling extremes in projections of future climate change. *Bulletin of the American Meteorological Society*, 81(3):427, 2000.

[18] Reto Knutti, Thomas F Stocker, Fortunat Joos, and Gian-Kasper Plattner. Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature*, 416(6882):719–723, 2002.

[19] Peter A Stott and JA Kettleborough. Origins and estimates of uncertainty in predictions of twenty-first century temperature rise. *Nature*, 416(6882):723–726, 2002.

[20] KJ Hennessy, PH Whetton, JJ Katzfey, JL McGregor, RN Jones, CM Page, and KC Nguyen. Fine resolution climate change scenarios for new south wales annual report, 1998.

[21] Thomas R Karl and Richard W Knight. Secular trends of precipitation amount, frequency, and intensity in the united states. *Bulletin of the American Meteorological society*, 79(2):231–241, 1998.

[22] Richard H Moss, N Nakicenovic, and BC O'Neill. Towards new scenarios for analysis of emissions, climate change, impacts, and response strategies. 2008.

[23] Richard H Moss, Jae A Edmonds, Kathy A Hibbard, Martin R Manning, Steven K Rose, Detlef P Van Vuuren, Timothy R Carter, Seita Emori, Mikiko Kainuma, Tom Kram, et al. The next generation of scenarios for climate change research and assessment. *Nature*, 463(7282):747–756, 2010.

[24] IPCC Adopted. Climate change 2014 synthesis report. 2014.

[25] Zavareh Kothavala. Extreme precipitation events and the applicability of global climate models to the study of floods and droughts. *Mathematics and computers in simulation*, 43(3):261–268, 1997.

[26] KJ Hennessy, JM Gregory, and JFB Mitchell. Changes in daily precipitation under enhanced greenhouse conditions. *Climate Dynamics*, 13(9):667–680, 1997.

[27] B Bhaskaran and JFB Mitchell. Simulated changes in southeast asian monsoon precipitation resulting from anthropogenic emissions. *International Journal of Climatology*, 18(13):1455–1462, 1998.

[28] Filippo Giorgi, Linda O Mearns, Christine Shields, and Larry McDaniel. Regional nested model simulations of present day and 2× co2 climate over the central plains of the us. *Climatic Change*, 40(3-4):457–493, 1998.

[29] Thomas R Knutson and Syukuro Manabe. Time-mean response over the tropical pacific to increased c02 in a coupled ocean-atmosphere model. *Journal of Climate*, 8(9):2181–2199, 1995.

[30] Francis W Zwiers and Viatcheslav V Kharin. Changes in the extremes of the climate simulated by ccc gcm2 under co2 doubling. *Journal of Climate*, 11(9):2200–2222, 1998.

[31] Nathaniel B Guttman and Robert G Quayle. A historical perspective of us climate divisions. *Bulletin of the American Meteorological Society*, 77(2):293–303, 1996.

[32] Thomas Karl and Walter James Koss. *Regional and national monthly, seasonal, and annual temperature weighted by area, 1895-1983*. National Climatic Data Center, 1984.

[33] Barry D Keim, Adam M Wilson, Cameron P Wake, and Thomas G Huntington. Are there spurious temperature trends in the united states climate division database? *Geophysical Research Letters*, 30(7), 2003.

[34] Barry D Keim, Gregory E Faiers, Robert A Muller, John M Grymes, and Robert V Rohli. Long-term trends of precipitation and runoff in louisiana, usa. *International Journal of Climatology*, 15(5):531–541, 1995.

[35] Daniel J Leathers, Andrew J Grundstein, and Andrew W Ellis. Growing season moisture deficits across the northeastern united states. *Climate Research*, 14(1):43–55, 2000.

[36] Climdex - datasets for indices of climate extremes. `http://www.climdex.org/indices.html`, 2014.

[37] 2014 national climate assessment. `http://nca2014.globalchange.gov/highlights/regions/southeast`, 2014.

[38] Harry J Khamis. The δ-corrected kolmogorov-smirnov test for goodness of fit. *Journal of statistical planning and inference*, 24(3):317–335, 1990.

[39] Myles Hollander, Gordon Pledger, and Pi-Erh Lin. Robustness of the wilcoxon test to a certain dependency between samples. *The Annals of Statistics*, pages 177–181, 1974.

[40] A Di Bucchianico. Combinatorics, computer algebra and the wilcoxon-mann-whitney test. *Journal of Statistical Planning and Inference*, 79(2):349–364, 1999.

[41] Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6):80–83, 1945.

[42] Dennis R Helsel and Robert M Hirsch. Applicability of the t-test for detecting trends in water quality variables. *JAWRA Journal of the American Water Resources Association*, 24(1):201–204, 1988.

[43] Robert M Hirsch and James R Slack. A nonparametric trend test for seasonal data with serial dependence. *Water Resources Research*, 20(6):727–732, 1984.

[44] Henry B Mann and Donald R Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60, 1947.

[45] Andrej N Kolmogorov. *Sulla determinazione empirica di una legge di distribuzione*. na, 1933.

[46] Scott Bell. Introduction to scientific research methods in geography. *The Canadian Geographer/Le Géographe canadien*, 51(3):402–403, 2007.

[47] Bhaskar Bhattacharya and Desale Habtzghi. Median of the p value under the alternative hypothesis. *The American Statistician*, 56(3):202–206, 2002.

[48] Raymond Hubbard. Alphabet soup blurring the distinctions betweenp's anda's in psychological research. *Theory & Psychology*, 14(3):295–327, 2004.

[49] Thomas Sellke, MJ Bayarri, and James O Berger. Calibration of ρ values for testing precise null hypotheses. *The American Statistician*, 55(1):62–71, 2001.

[50] Valen E Johnson. Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences*, 110(48):19313–19317, 2013.

[51] David Colquhoun. An investigation of the false discovery rate and the misinterpretation of p-values. *Open Science*, 1(3):140216, 2014.

Vita

Xinbo Huang was born at Wuhan, Hubei Province, People's Republic of China in 1988. He earned his Master's Degree in Geomatics Engineering from Wuhan University in 2013. During this time, he was greatly influenced by outstanding faculty members such as Dr. Liangmin Liu and Dr. Zhaocong Wu.

In Fall of 2013, he entered the MS program at Louisiana State University with Dr. Lei Wang as his primary advisor. He worked as a Graduate Research Assistant at the NOAA Southern Regional Climate Center under the supervision of Dr. David Sathiaraj. During this time, he gained experience in the nation-wide climate information system, Applied Climate Information System (`www.rcc-acis.org`) and used this large data repository to conduct this data-driven study.

Xinbo Huang will be graduating with an MS degree in Geography in Fall 2016.