May 2020

A Statistical Analysis of Drought and Its Global Impact

Jared Delk
Clemson University, delkja97@gmail.com

A STATISTICAL ANALYSIS OF DROUGHT AND ITS GLOBAL IMACT

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Civil Engineering

---

by
Jared Delk
May 2020

Accepted by:
Dr. Ashok Mishra, Committee Chair
Dr. Abdul Khan
Dr. Nigel Kaye

ABSTRACT

Droughts are the most ambiguous of all natural hazards and yet are often cited as the most destructive and are responsible for the most widespread damage across all sectors of society. The purpose of this study was to further understand the impact that drought has on various sectors of society, especially the economic sector, and how various regions across the United States are specifically impacted by droughts and drought effects. In order to quantify the impact that drought has on the economic sector, an analysis was performed internationally between each country's GDP and various drought indices such as PDSI, SPI, and SPEI. In order to account for exponential growth in GDP, the correlation was performed on detrended GDP using logarithmic trend free pre-whitening (TFPW) and logarithmic quadratic methods. The combination of PDSI and Log. TFPW gave the most complete understanding of negative correlation between drought and a nation's economy. In order to focus on drought impact in the United States, ARIMA modeling was used to establish a forecasting model for PDSI time series for various climatic regions around the country. The accuracy of these forecasting models was quantified through an approximate AIC method and compared to precipitation and temperature of each of the regions to determine the influence each drought component had on model accuracy. The regions with lower temperatures such as the Upper Midwest gave the more accurate drought forecasting models. The applicability of each of these climatic regions towards drought studies were tested by Severity Area Frequency curve analysis. While the Northwest region of America necessitated a need for two drought sub-regions, most of the climatic regions were affected by droughts homogenously.

ACKNOWLEDGMENTS

I appreciate all of the helpfulness and flexibility that Dr. Ashok Mishra, my advisor, expressed to me. I would not have been able to have this opportunity without him. I would also like to thank the rest of my committee, Dr. Abdul Khan and Dr. Nigel Kaye, for their expertise and cooperation throughout my year of graduate school. The rest of the faculty and staff in the Glenn Department of Civil Engineering have been extremely helpful throughout my time at Clemson and I would like to express my greatest gratitude to all of those that helped me throughout my entire Clemson career including Dr. Jennifer Ogle and Mrs. CJ Bolding.

I would also like to express appreciation to the Glenn family for extending a research fellowship to me this school year as well as the fellowship committee headed by the chair of the Civil Engineering department, Dr. Jesus de la Garza, for selecting me for the research fellows honors.

Finally I would like to think my family, specifically my mother, father, brother, and fiancée for all of their support and the sacrifice of their time as well as all of my friends in Clemson, Columbia, and Charleston. Your love and encouragement provided through all of my student career and my personal life has helped me achieve this prestigious honor.

TABLE OF CONTENTS

Page

TITLE PAGE i

ABSTRACT ii

ACKNOWLEDGMENTS iii

LIST OF TABLES vi

LIST OF FIGURES vii

LIST OF EQUATIONS xi

CHAPTER 1: A REVIEW OF DROUGHTS 1

1.1 Droughts as Natural Hazards 1
1.2 Drought Indices 7
1.3 Drought Descriptors 14
1.4 Study Objectives and Outline 15
1.5 Resources 17

CHAPTER 2: DROUGHT IMPACT ON GLOBAL ECONOMIES 19

2.1 Introduction 19
2.2 Economics Background 25
2.3 Methodology 29
2.4 Results and Discussion 35
2.5 Conclusion 47
2.6 Resources 48

CHAPTER 3: DROUGHT FORECASTING AND REGIONAL ANALYSIS 52

3.1 Introduction 52
3.2 Methodology 54
3.3 Regional Applicability of Stochastic Models on PDSI 60
3.4 Spatio-Temporal Regional Analysis 71
3.5 Conclusion 83
3.6 Resources 85

Table of Contents (Continued)

Page

CHAPTER 4: CONCLUSION 88

4.1 Summary of Results 88
4.2 Study Limitations and Future Recommendations 90

LIST OF TABLES

Table Page

3-i Candidate Models for All Regions 61

3-ii Best Fit Models for All Regions 62

3-iii Comparing Maximum Drought Severity and Intensity and Percentage of Drought between 1900-1959 & 1960-2019 69

LIST OF FIGURES

Figure		Page
1-1	Drought run theory characteristics	15
2-1	GDP Detrending Methods on United States GDP	30
2-2	GDP Detrending Methods PDSI Scatter Plots on United States GDP	34
2-3	Global Average Annual PDSI Drought Severity & Log-Quad Detrended Global GDP	35
2-4	Global Average Annual PDSI Drought Severity & Log TFPW Detrended Global GDP	36
2-5	World GDP PDSI National Correlation Values	37
2-6	World GDP SPI National Correlation Values	38
2-7	World GDP SPEI National Correlation Values	38
2-8	Continental GDP PDSI National Correlation Values	40
2-9	Global Log-Quad GDP PDSI National Correlation Values	41
2-10	Global Log TFPW GDP PDSI National Correlation Values	41
2-11	Global Log-Quad GDP SPI National Correlation Values	42
2-12	Global Log TFPW GDP SPI National Correlation Values	43
2-13	Global Log-Quad GDP SPEI National Correlation Values	44
2-14	Global Log TFPW GDP SPEI National Correlation Values	44
2-15	Global GDP PDSI National Significantly Negative Correlation Values	45
2-16	Global GDP SPI National Significantly Negative Correlation Values	46

List of Figures (Continued)

Figure Page

2-17 Global GDP SPEI National Significantly Negative Correlation Values 46

3-1 NOAA US Climatic Regions 52

3-2 SPI Data Stations and Associated Regions of Influence 58

3-3 ACF Correlogram for Southeast region (SC) PDSI from 1895-1989 60

3-4 PACF Correlogram for Southeast region (SC) PDSI from 1895-1989 61

3-5 Southeast Region (SC) AR(1) PDSI Model In Testing Period of 1990-2019 62

3-6 South Region (TX) AR(1) PDSI Model In Testing Period of 1990-2019 63

3-7 Southwest Region (CO) AR(1) PDSI Model In Testing Period of 1990-2019 63

3-8 West Region (CA) AR(1) PDSI Model In Testing Period of 1990-2019 64

3-9 Northwest Region (WA) AR(1) PDSI Model In Testing Period of 1990-2019 64

3-10 Northern Rockies & Plains Region (SD) AR(1) PDSI Model In Testing Period of 1990-2019 65

3-11 Upper Midwest Region (WI) AR(1) PDSI Model In Testing Period of 1990-2019 65

3-12 Ohio Valley Region (IL) AR(1) PDSI Model In Testing Period of 1990-2019 66

3-13 Northeast Region (NY) AR(1) PDSI Model In Testing Period of 1990-2019 66

List of Figures (Continued)

Figure Page

3-14 Regional Average Monthly Precipitation and Stochastic Model Accuracy 68

3-15 Regional Average Monthly Temperature and Stochastic Model Accuracy 68

3-16 Regional Average Monthly Precipitation/Temperature and Stochastic Model Accuracy 68

3-17 Maximum Drought Severity between 1900-1959 & 1960-2019 70

3-18 Maximum Drought Intensity between 1900-1959 & 1960-2019 70

3-19 Maximum Percentage of Drought between 1900-1959 & 1960-2019 70

3-20 Southeast PDSI SAF Curves 72

3-21 South PDSI SAF Curves 72

3-22 Southwest PDSI SAF Curves 73

3-23 West PDSI SAF Curves 73

3-24 Northwest PDSI SAF Curves 74

3-25 Northern Rockies & Plains PDSI SAF Curves 74

3-26 Upper Midwest PDSI SAF Curves 75

3-27 Ohio Valley PDSI SAF Curves 75

3-28 Northeast PDSI SAF Curves 76

3-29 Climate Regions 50 Year PDSI SA Curves 76

3-30 Climate Regions 100 Year PDSI SA Curves 77

List of Figures (Continued)

Figure Page

3-31 Southeast SPI SAF Curves 78

3-32 South SPI SAF Curves 78

3-33 Southwest SPI SAF Curves 79

3-34 West SPI SAF Curves 79

3-35 Northwest SPI SAF Curves 80

3-36 Northern Rockies & Plains SPI SAF Curves 80

3-37 Upper Midwest SPI SAF Curves 81

3-38 Ohio Valley SPI SAF Curves 81

3-39 Northeast SPI SAF Curves 82

3-40 Climate Regions 50 Year SPI SA Curves 82

3-41 Climate Regions 100 Year SPI SA Curves 83

LIST OF EQUATIONS

Equation		Page
1-A	Thornthwaite PET Estimate	9
1-B	Monthly Heat Index	9
1-C	Thornthwaite Heat Index Exponent	9
1-D	Thornthwaite Correction Coefficient	9
1-E	Maximum Number of Sun Hours	9
1-F	Hourly Angle of Sun Rising	9
1-G	Solar Declination	9
1-H	Potential Recharge	10
1-I	Potential Losses	11
1-J	Surface Layer Potential Losses	11
I-K	Underlying Layer Potential Losses	11
1-L	Potential Runoff	11
1-M	Evapotranspiration Monthly Coefficient	11
1-N	Recharge Monthly Coefficient	11
1-O	Runoff Monthly Coefficieent	11
1-P	Losses Monthly Coefficient	11
1-Q	Water Balance Deficit	12
1-R	Palmer Z-Score	12
1-S	Palmer Weighting Factor	12
1-T	Palmer Weighting Factor Coefficient	12

List of Equations (Continued)

Equation		Page
1-U	Moisture Demand to Supply Ratio	12
1-V	Palmer Drought Severity Index	13
2-A	Thiel-Sen Slope Analysis	28
2-B	Trend Free Processing	28
2-C	Pre-Whitening Processing	28
2-D	Quadratic Trend Free Processing	29
2-E	Log-Pearson Type III Distribution Probability Density Function	31
2-F	Log-Pearson Type III Distribution Shape Parameter	31
2-G	Log-Pearson Type III Distribution Scale Parameter	32
2-H	Log-Pearson Type III Distribution Location Parameter	32
2-I	Log-Logistic Distribution Probability Density Function	32
2-J	Log-Logistic Distribution Mean Derivation	32
2-K	Log-Logistic Distribution Standard Deviation Derivation	32
2-L	Log-Logistic Distribution Skewness Derivation	32
2-M	Log-Logistic Distribution Gamma Based Function	32
2-N	Pearson Product Moment Correlation Coefficient	33
2-O	Sample Covariance Between x & y	33
2-P	Sample Standard Deviation According to x	33
2-Q	Sample Standard Deviation According to y	34

List of Equations (Continued)

Equation Page

3-A Autocovariance Equation 55

3-B Autocorrelation Equation 56

3-C Partial Autocorrelation Estimation 56

3-D Approximate Akaike's Information Criterion 57

3-E Gamma Distribution Probability Density Function 59

3-F Gamma Distribution Mean Derivation 59

3-G Gamm Distribution Variance Derivation 59

3-H Scaled PDSI Cumulative Area Percentage 60

3-I Scaled SPI Cumulative Weighted Area Percentage 60

CHAPTER 1: A REVIEW OF DROUGHTS

1.1 Droughts as Natural Hazards

The World Health Organization (WHO) defines a natural disaster as "an act of nature of such magnitude as to create a catastrophic situation in which the day-to-day patterns of life are suddenly disrupted and people are plunged into helplessness and suffering and as a result need food, clothing, shelter, medical, and nursing care and other necessities of life and protection against unfavourable environmental factors and conditions" (Assar, 1971). A natural hazard is the threat of a natural disaster that has not necessarily created the anthropomorphic impact that classifies as a disaster. Ultimately, natural hazards and disasters are determined by their impact on human activity and are even declassified as hazards if they have no human impact (Maybank et al., 1995). Natural hazards can include earthquakes, floods, tornadoes, storms, and droughts. Of all of these natural hazards, droughts arguably have the highest human impact over the entirety of its impact (Wilhite 2001). This is due to the cyclical negative system of water shortage creating higher water demand ultimately creating significant damage to societies as a whole.

While there are many aspects of droughts which are similar to other natural hazards such as the varying degrees of intensity, duration, and areal extent, there are many ways in which droughts differ which has created an ambiguous nature surrounding droughts and their impact. While most other hazards are a sudden event that occurs very rapidly, the onset and end of droughts are usually difficult to determine and have made

droughts considered a creeping phenomenon. This makes drought monitoring difficult to determine and also leaves mild local scale or short term droughts as mostly unidentifiable. And while all natural hazards can vary in scale, the maximum range of the areal extent of droughts is much larger than any other natural hazard. This makes the impacts of droughts more widespread but these direct impacts have been difficult to estimate in total. This is because the damages that drought cause are not often structural which is usually the standard which the damages of other natural hazards are estimated by. The cost of human life is not directly impacted as much as other natural hazards but the indirect impact of droughts have had tolls on human health and livelihood.

Finally, the main difference between droughts as opposed to other natural hazards is due to their lack of consensus definition. Most definitions revolve around a lack of precipitation while some include an aspect of temperature but it is important to explicitly differentiate a drought from a heat wave. Not only do heat waves lack the facet of precipitation but they also take place over a time scale of weeks rather than the months or years that a major drought event can last (Mishra & Singh, 2010). Droughts have different causes in varying locations and regions throughout the world, especially regions where precipitation follows a seasonality in magnitude and intensity. While other natural hazards are clearly defined by their natural structure, droughts have been defined and classified by their impacts rather than the natural or meteorological makeup (Lloyd-Hughes, 2014). Multiple studies previously performed on droughts have put droughts into four different categories of increasing effect magnitude including (1) meteorological, (2)

hydrological, (3) agricultural, (4) groundwater, and (5) socio-economical droughts (Hisdal & Tallaksen, 2000). An expansion of these drought classifications follows.

1.1.1 Meteorological Droughts

A meteorological drought is based purely on precipitation and climatological norms. These droughts are defined by their degree of dryness and below average rainfall. In order to determine if lack of precipitation is due to drought, it is important to take location and local climate into account. In areas where the climate is already arid and dry, long periods of little rainfall may not be accurately classified as drought as this may just be a natural part of the climate (Wilhite & Glantz, 1985). Areas with seasonal rainfall such as the Indian subcontinent may also have periods with little rainfall and extreme rainfall such as the dry and monsoon seasons, but the dry seasons may not be classified as drought related since it is an annual and expected event. A special exception to these climate related lack of precipitation are those areas affected by El Niño-Southern Oscillation (ENSO) patterns as these events are inter-annual, less predictable, and extend beyond any single global region or climate. Meteorological droughts in all regions, whatever the cause, are all defined by a deficit of precipitation from the expected average over an extended time period. Because of the extent of records of precipitation data globally, most of the studies and analyses performed on droughts have centered on meteorological droughts, including the majority of the focus of this study.

1.1.2 Hydrological Droughts

Droughts are classified as hydrological droughts when the lack of precipitation affects the access of local surface water resources. Many of the records defining this class of droughts include streamflow and land runoff volumes. These types of droughts are generally associated with low streamflow but this is not the only indicator of hydrological droughts as they affect surface water availability for water resource systems as well. Low streamflow is also not always an indicator of drought as there is a larger temporal aspect to droughts that an instantaneous record such as streamflow does not fully capture (Van Loon, 2015). A streamflow deficit can be caused by human activity and while this can exacerbate effects on other sectors of society, droughts are natural hazards and as such are caused by acts of nature rather than human activity. Hydrologic droughts are greatly impacted by local geography as land runoff is the direct supply of these hydrological records. Catchments and watersheds that are highly resilient to drought include those areas with low flow speeds and greater water accumulation spread throughout the area. This means that changes in land cover and land usage can also make areas more or less prone to hydrological droughts over time.

1.1.3 Agricultural Droughts

Agricultural droughts are determined when lack of water availability begins to affect agricultural production. These droughts are not only dependent on precipitation but also soil moisture and the specific crop water demand. The

variability associated with these kinds of droughts make agricultural droughts highly localized and region specific. While all drought classifications have some kind of impact on human activity, agricultural droughts are the earliest classification where droughts directly impact human society and are largely determined by human action. Choice of crop and farming practices can directly affect the severity of agricultural droughts (Rodda & Ubertini, 2004). These kind of droughts are also highly dependent on temperature as evapotranspiration drying out the upper layers of soil directly adds to the decrease in crop yield. Ultimately agricultural droughts focus on soil moisture and its direct effect on crop productions.

1.1.4 Groundwater Drought

A groundwater drought occurs when the significant lack of groundwater recharge begins to affect the groundwater storage levels and reduces discharge. Groundwater droughts are highly dependent on media properties where the groundwater is stored, similar to the impact that the surface sub-basins have on hydrological droughts. Aquifers with low permeability do not store as much total groundwater but are not as easily affected by short term recharge deficits unlike aquifers with high permeability which store more groundwater but have highly variable groundwater levels. Overall, groundwater droughts are on a much longer time scale than others such as meteorological droughts due to the lag time associated with groundwater recharge (Peters et al., 2005). Shorter or milder

droughts generally leave groundwater unaffected because of this recharge lag time buffer which makes groundwater management a vital part of regional drought resiliency. It is important to differentiate groundwater droughts from the previous agricultural droughts as agricultural droughts deal with low soil moisture whereas groundwater droughts are focused on deeper, saturated aquifer levels that are not accessed by crop roots (Van Lanen & Peters, 2000). Also agricultural droughts are primarily measured by drop yields while groundwater droughts focus on groundwater availability.

1.1.5 Socio-economical Droughts

Droughts which cause a shortage in demand of water resource systems are classified as socio-economical droughts. These droughts focus on water as an economic good and occur when the supply for that good cannot meet the water demand. Socio-economical droughts are the most significant to human activity and involve components of all previous drought classifications. The efficiencies in system development and water resource policy determine the vulnerability that a society has to socio-economic droughts. Minimizing water demand through conservative practices help increase drought resiliency in times of water scarcity. There is often public confusion between actual socio-economic droughts and water shortages created by over-usage of the local water supply. Consistent with all other drought classifications, socio-economical droughts are caused by weather related water shortages and cannot be reversed by short term human activity.

1.2 Drought Indices

While there are many indices focusing on various factors contributing to droughts such as precipitation, temperature, soil moisture, crop yields, etc., this study primarily focuses on the most popular indices of Palmer Drought Severity Index (PDSI) and Standardized Precipitation Index (SPI). It also includes a variation of SPI which includes the calculation of evapotranspiration (ET) called the Standardized Precipitation ET Index (SPEI). The calculation and limitations of each index is discussed below.

1.2.1 Standardized Precipitation Index (SPI)

In 1993, McKee et al. introduced SPI based on standardized precipitation, or "the difference of precipitation from the mean for a specified time period divided by the standard deviation where the mean and standard deviation are determined from past records." The recommendations from that 1993 study state that the previous precipitation records used should be based on at least 30 years of precipitation data within the area of interest. The use of standardized precipitation implies that the precipitation data fits a normal distribution but as this is not always the case, a method was also determined to find the probability of the desired rainfall from a gamma distribution and then using an inverse normal function on this probability to determine the corresponding SPI value. The gamma distribution was chosen as an alternate method as this distribution is commonly associated with precipitation statistical distributions and modeling (Husak et al., 2007). SPI is commonly used over monthly periods but could also expand to include

multi-month, annual, or multi-year accumulations of rainfall. SPI is often used because of its simplicity and flexibility in terms of regionality and time scale (McKee et al., 1995). The main limitations associated with SPI are the dependency on distribution fit of the data and length of precipitation records used to derive the index value. Multiple statistical distributions have been used in association with SPI including the Log-Pearson Type III (LP3) distribution used later in this study. The precipitation data must have a high goodness-of-fit to the applied statistical distribution in order to produce accurate SPI values. The length of previous precipitation records used can directly affect this statistical distribution as a robust record of past precipitation creates much more reliable SPI values that accurately reflect the rainfall or lack thereof within the region of interest.

1.2.2 Standardized Precipitation Evapotranspiration Index (SPEI)

Similar to SPI, SPEI also involves a standardization process where any statistical distribution used to determine probability of monthly data is converted into a normally distributed Z-score but its standardized value includes the calculation of loss of runoff due to ET (Vicente-Serrano et al., 2010). A common method used to predict ET or potential ET (PET) is the Thornthwaite method developed in 1948 which only requires the monthly average temperature for the study location. This method for calculating PET is shown in Equations 1-A – 1-G. The T in Equation 1-A represents the average temperature for the specified month in degrees Celsius. The I in Equation 1-A is a heat index which is a summation of all monthly heat indices i shown in Equation 1-B. The m

in Equation 1-A represents a coefficient based on the heat index *I* of the area and is calculated through Equation 1-C. *K* in Equation 1-A is calculated in Equation 1-D and is a correction coefficient based on the maximum number of sun hours *N* and the number of days within the specified month *NDM*. The maximum number of sun hours *N* shown in Equation 1-E is based on the hourly angle of sun rising ϖ_s. Equation 1-F shows the calculation of the hourly angle of sun rising ϖ_s which is based on the latitude of the location of interest φ and the solar declination δ, both in radians. The solar declination δ in Equation 1-G is based on the average Julian day of the specified month *J*. After deriving the PET, this value is directly removed from the monthly precipitation depth to give the rainfall deficit or surplus for that month.

$$PET = K\left(\frac{10*T}{I}\right)^{m} \tag{1-A}$$

$$i = \left(\frac{T}{5}\right)^{1.514} \tag{1-B}$$

$$m = 6.75 * 10^{-7} * I^3 - 7.71 * 10^{-5} * I^2 + 1.79 * 10^{-2} * I + 0.492 \tag{1-C}$$

$$K = \left(\frac{N}{12}\right)\left(\frac{NDM}{30}\right) \tag{1-D}$$

$$N = \left(\frac{24}{\pi}\right)\varpi_s \tag{1-E}$$

$$\varpi_s = \arccos(-\tan(\varphi)\tan(\delta)) \tag{1-F}$$

$$\delta = 0.4093\sin\left(\frac{2\pi J}{365} - 1.405\right) \tag{1-G}$$

1.2.3 Palmer Drought Severity Index (PDSI)

The PDSI is one of the most widely used drought index internationally but it is also one of the most complex indices (Jacobi et al., 2013). The index, developed by Palmer in 1965, is based on a water balance theory where precipitation *P* follows one of four different outcomes: evapotranspiration (*ET*), recharge (*R*), runoff (*RO*), or losses (*L*). The maximum values of each of these outcomes are estimated by "potential" values corresponding to each of the estimates. The value for potential ET (*PE*) is estimated using the Thornthwaite method shown previously. Potential recharge (*PR*) is estimated through Equation G and is based on the available water capacity (*AWC*) of the underlying layer of soil or the depth that can hold water before reaching the wilting point in the region. In his water balance method, Palmer split the soil of the study location into two layers with the surface layer being the depth necessary to hold 25 mm of moisture and the underlying layer being dependent on all other characteristics of the soil. The depth of this underlying layer is dependent on the AWC of the area. PR is also based on the available moisture already stored in both the surface and underlying layers represented by S_s and S_u in Equation 1-H respectively.

$$PR = AWC - (S_s + S_u) \tag{1-H}$$

The potential loss (*PL*) of the soil is defined as the summation of the potential losses in both the surface and underlying layers of soil or PL_s and PL_u in Equation 1-I. PL_s is defined as the minimum of either the PE or the S_s, shown in Equation 1-J. The PL_u is given by Equation 1-K.

$$PL = PL_s + PL_u \tag{1-I}$$

$$PL_s = \min(PE, S_s) \tag{1-J}$$

$$PL_u = \frac{(PE - S_s)S_u}{AWC} \tag{1-K}$$

This leaves potential runoff (PRO) to be defined as the potential precipitation that was not soaked into the soil. In Palmer's original study, he estimated potential precipitation to be equivalent to AWC giving the definition of PRO shown in Equation 1-L.

$$PRO = AWC - PR = S_s + S_u \tag{1-L}$$

These four potential values are averaged for each month using previous records along with the average of the estimated actual values of ET, recharge, losses, and runoff to create the four local monthly coefficients used to describe the area of interest shown in Equations 1-M – 1-P.

$$\alpha_j = \frac{\overline{ET}_j}{\overline{PE}_j} \tag{1-M}$$

$$\beta_j = \frac{\bar{R}_j}{\overline{PR}_j} \tag{1-N}$$

$$\gamma_j = \frac{\overline{RO}_j}{\overline{PRO}_j} \tag{1-O}$$

$$\delta_j = \frac{\bar{L}_j}{\overline{PL}_j} \tag{1-P}$$

These coefficients are combined with the estimated actual values of the four outcomes for the specified month to create the Climatically Appropriate For Existing Conditions (*CAFEC*) precipitation which is then compared to the actual precipitation for the specified month to determine the surplus or deficit of precipitation for that month. This deficit (d) is shown in Equation 1-Q.

$$d = P - CAFEC = P - (\alpha * PE + \beta * PR + \gamma * PRO - \delta * PL) \quad \text{(1-Q)}$$

The deficit is then multiplied by a weighting factor created for the specified location and month (K_j) in order to give the moisture anomaly index (Z) shown in Equation 1-R, commonly referred to as the Z-index. The weighting factor, shown in Equation 1-S, is based on the average of the absolute values of the deficits for that month in previous records as well as the ratio T_j of the average moisture demand to the average moisture supply for that month calculated in Equation 1-U.

$$Z = K_j * d \quad \text{(1-R)}$$

$$K_j = \frac{17.67\widehat{K_J}}{\sum_{i=1}^{12} \overline{D_i} * \widehat{K_i}} \quad \text{(1-S)}$$

$$\widehat{K_J} = 1.5 \log_{10}\left(\frac{T_J + 2.8}{\overline{D_J}}\right) + 0.5 \quad \text{(1-T)}$$

$$T_j = \frac{\overline{PE_J} + \overline{R_J} + \overline{RO_J}}{\overline{P_J} + \overline{L_J}} \quad \text{(1-U)}$$

After determining the Z-index for the given month i and location, the PDSI value for these specifications is dependent on the previous month's PDSI value and the Z-index

for the specified month. The final equation for calculating PDSI is shown in Equation 1-V.

$$PDSI_i = 0.897PDSI_{i-1} + \frac{Z_i}{3} \quad (1\text{-V})$$

While PDSI is one of the most widely used drought indices, there have been many limitations and critiques of the method that have been previously established and studied. The first of these critiques is typically a comment on the complexity of the PDSI method in comparison with other drought indices. While it does give the advantage of accounting for precipitation, temperature, and soil characteristics, these also require a large amount of data to be available for the desired area of study. Another critique of the Palmer method is the somewhat arbitrary values and coefficients it gives for some of the steps in the process, especially the definition of the surface and underlying layers of soil being 25 mm of moisture (Alley, 1984). While there have been other methods that have since developed to correct some of these arbitrary values such as the Self-Corrected PDSI (SC-PDSI), there are still some aspects of the Palmer method that are derived empirically and were defined through limited sample size of the original study. One of the other limitations to PDSI that has arisen is the assumption in the water balance method used that runoff only occurs after all of the AWC is saturated which underestimates total runoff. It also assumes that all precipitation that produces runoff is rain and does not account for regions with lower temperatures which are severely affected by snowfall. Lastly, another critique to PDSI is the autoregressive portion that is integrated into the methodology. While it is beneficial to incorporate the outstanding conditions of the

region into the calculation of PDSI for the specified month, the uneven weight that PDSI gives to its previous values allows a certain precipitation or drought event to have effects far into future values, with a single value holding weight for future values up to almost three and a half years into the future. This weight may cause certain drought events to be classified as droughts longer than they are actually affected the given region.

1.3 Drought Descriptors

When the various drought indices previously mentioned are plotted with time for a certain region, they create a time series that can describe the region's drought history. Run theory can then be applied to this time series compared to a certain threshold X_0, often at zero to incorporate any negative value, to quantify certain characteristics of a drought (Fu & Koutras, 1994). The characteristics shown in Figure 1-1 are severity, duration, and intensity (Yevjevich, 1967). The duration of a drought is the run length of a drought index or the number of consecutive time steps, usually months, which the drought index value is found to be below the threshold. The drought severity is the run sum of a drought index or the summation of all consecutive drought index values that are found to be below the threshold. While the intensity in Figure 1-1 is shown to be the average index value during the drought event, the definition of intensity used in this study is the maximum index value found within the specific drought event.

1.4 Study Objectives and Outline

The objectives of this study are to analyze the impact droughts have on various sectors of society, gain a further understanding of what regions are especially affected by

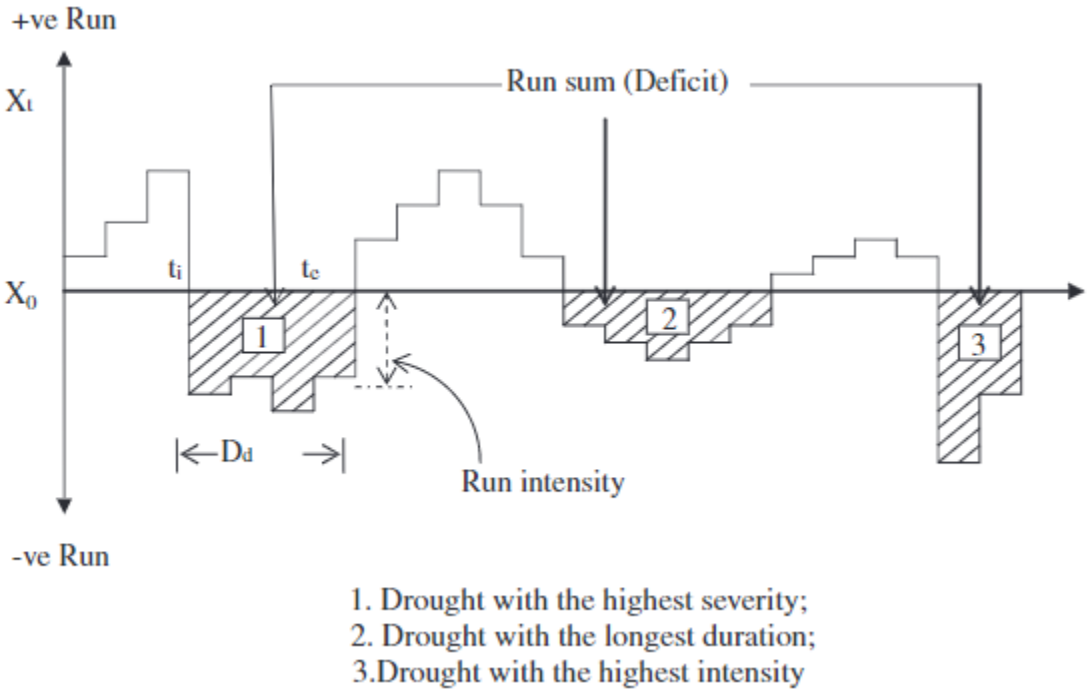

Figure 1-1: Drought run theory characteristics
(Mishra, Ashok K., and Vijay P. Singh. "A Review of Drought Concepts." Journal of Hydrology, July 2010)

drought, and determine the accuracy of drought forecasting in certain regions. In order to determine both the direct and indirect impact that droughts have on local and global scales, a study of the correlation between droughts and GDP is performed in Chapter 2. This chapter includes a summary of examples of drought studies affecting local GDP, a background review of GDP as an economic indicator, methods for transforming GDP values into a time series that can be comparable to drought index time series, the

methodology used to perform a correlation analysis between various drought indices and GDP, and the results of the international economic-drought correlation analysis. The specific objectives of the comparison of drought indices and GDP is to determine the regions of the world most impacted economically by drought as well as determine which drought index and GDP transformation method each nation should use to best indicate the impact that drought has on their own nation's economy.

Chapter 3 is focused on the applicability of drought indices to US national climatic regions. This chapter includes an introduction to time series stochastic modeling, a review of the use of Severity Area Frequency (SAF) curves as a spatio-temporal analysis, the methodology used in both ARIMA stochastic modeling and SAF curves, the results of the regional ARIMA modeling for PDSI, a brief test of the stationary assumption of stochastic modeling, and the results of the SAF curve analysis on all national climatic regions for various indices. The specific objective of these statistical analyses was primarily to determine if these national climatic regions were applicable to drought studies. Other objectives included which regions were more reliable to use stochastic modeling for drought forecasting, which aspects of drought most affected the accuracy drought modeling, and which drought index was most useful for summarizing the effects that droughts had on the various national climatic regions. The final Chapter 4 is a conclusion of the study. This includes a summary of all results, limitations of the study and recommendations for future studies.

1.5 Resources

Alley, William M. “The Palmer Drought Severity Index: Limitations and Assumptions.” Journal of Climate and Applied Meteorology, vol. 23, no. 7, July 1984, pp. 1100-1109., doi: 10.1175/1520-0450(1984)023<1100:tpdsil>2.0.co;2.

Assar, Mohamed M. Guide to Sanitation in Natural Disasters. World Health Organization, 1971.

Donald A. Wilhite & Michael H. Glantz (1985) Understanding: the Drought Phenomenon: The Role of Definitions, Water International, 10:3, 111-120, DOI: 10.1080/02508068508686328

Hisdal, H. and L. M. Tallaksen. “Drought Event Definition.” ARIDE Project, No. 6, December, 2000.

Husak, G.J., Michaelsen, J. and Funk, C. (2007), Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. Int. J. Climatol., 27: 935-944. doi:10.1002/joc.1441

J. C. Fu & M. V. Koutras (1994) Distribution Theory of Runs: A Markov Chain Approach, Journal of the American Statistical Association, 89:427, 1050-1058, DOI: 10.1080/01621459.1994.10476841

J. Maybank, B. Bonsai, K. Jones, R. Lawford, E.G. O'Brien, E.A. Ripley & E. Wheaton (1995) Drought as a natural disaster, Atmosphere-Ocean, 33:2, 195-222, DOI: 10.1080/07055900.1995.9649532

Jacobi, J., Perrone, D., Duncan, L. L., and Hornberger, G. (2013), A tool for calculating the Palmer drought indices, Water Resour. Res., 49, 6086– 6089, doi:10.1002/wrcr.20342.

Lloyd-Hughes, B. The impracticality of a universal drought definition. Theor Appl Climatol 117, 607–611 (2014). https://doi.org/10.1007/s00704-013-1025-7

McKee, T.B., Doesken, N.J., Kleist, J., 1993. The Relationship of Drought Frequency And Duration to Time Scales, Paper Presented at 8th Conference on Applied Climatology. American Meteorological Society, Anaheim, CA.

McKee, T.B., Doesken, N.J., Kleist, J., 1995. Drought Monitoring with Multiple Time Scales, Paper Presented at 9th Conference on Applied Climatology. American Meteorological Society, Dallas, Texas.

Mishra, Ashok K., and Vijay P. Singh. "A Review of Drought Concepts." Journal of Hydrology, vol. 391, no. 1-2, 9 July 2010, pp. 202–216., doi:10.1016/j.jhydrol.2010.07.012.

Palmer, W.C., 1965. Meteorologic Drought. US Department of Commerce, Weather Bureau, Research Paper No. 45.

Peters, E., et al. "Drought in Groundwater—Drought Distribution and Performance Indicators." Journal of Hydrology, vol. 306, no. 1-4, May 2005, pp. 302–317., doi:10.1016/j.jhydrol.2004.09.014.

Rodda, J. C., and L. Ubertini. The Basis of Civilization--Water Science? IAHS, 2004.

Thornthwaite, C. W. "An Approach toward a Rational Classification of Climate." Geographical Review, vol. 38, no. 1, 1948, pp. 55–94. JSTOR, www.jstor.org/stable/210739.

Van Lanen H.A.J., Peters E. (2000) Definition, Effects and Assessment of Groundwater Droughts. In: Vogt J.V., Somma F. (eds) Drought and Drought Mitigation in Europe. Advances in Natural and Technological Hazards Research, vol 14. Springer, Dordrecht

Van Loon, A.F. (2015), Hydrological drought explained. WIREs Water, 2: 359-392. doi:10.1002/wat2.1085

Vicente-Serrano, Sergio M., et al. "A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index." Journal of Climate, vol. 23, no. 7, Apr. 2010, pp. 1696–1718., doi:10.1175/2009jcli2909.1.

Wilhite, Donald A. Drought: a Global Assessment. Vol. 1, Routledge, 2001.

Yevjevich, V., 1967. An Objective Approach to Definitions and Investigations of Continental Hydrologic Drought. Hydrology Paper No. 23, Colorado State Univ., Fort Collins, Colo.

CHAPTER 2: DROUGHT IMPACT ON GLOBAL ECONOMIES

2.1 Introduction

There have been many examples throughout history of negative impacts due to droughts in all areas of the world. Much literature has been written on these various impacts, especially agricultural impacts due to droughts, but the economic impact of droughts in all sectors of society is relatively undiscovered. There is great work to be done on determining and quantifying the connection between droughts and economies of various scales across all sectors. Of the studies that have been done that have modeled drought impact on economics (Harou et al., 2010; Freire-González et al., 2017), a few significant drought events have been highlighted.

2.1.1 Catalonian Drought (2004-2008)

The first of these significant drought examples was the Catalonian drought in the mid to late 2000s which at one point caused the local government to have water shipped in from external sources such as Mersailles, France (Iceland, 2019). Catalonia, which is located in northeast Spain, is a region that contains Barcelona which has the highest population density in the country of Spain at 5.5 million people. This high population paired with the fact that the area already has inconsistent rainfall like much of the other Mediterranean cities meant that the area was already considerably vulnerable to drought (Martin-Ortega et al., 2012). When the worst drought on record in at least 90 years (with previous records being unreliable) hit the area from 2004-2008, the Catalonian region felt

the impact in all areas of society. As water reserves reached as low as 21% capacity in March 2008, a survey of the residents of Barcelona showed that around two-thirds of Barcelona residents took active water conservation methods and changed lifestyle routines (March et al., 2013). The Catalonian government made communication and conservation education a high priority as many campaigns to reduce water usage during hygiene routines such as turning off the water while brushing teeth or reducing time spent in showers were commonplace during the water shortage. This caused the Barcelona water demand to lower by 21% from 2007-2008 which created budget savings for the Ter-Llobregat river basin authority, an agency which regulates the usage of the watershed located inland in the Catalonian region, to provide relief for its residents. The total costs for the relief and regulation measures provided by the river basin authority made up 4.2% of the Spanish national budget in 2008.

Outside of the social impacts of this drought event, the direct and indirect impacts of the drought across all sectors was estimated by Martin-Ortega et al. (2012) through direct economic records and numerical estimations of impact on other sectors of society through previous input output tables of Gross Domestic Product (GDP) published by the Catalonian government. The areas most directly impacted by the regional drought were agricultural and hydroelectric production. With both of these sectors experiencing a 250 million € and 114 million € losses between 2007 and 2008, respectively, the total direct production loss was estimated around 620 million €. Of all of the indirect losses that occurred due to the direct production loss, none were hit harder than the industrial sector which experienced a financial loss equivalent to 132.7 million €/year. Ultimately the total

losses from 2007-2008, including direct and indirect impacts, were estimated to be 750 million €, corresponding of a 0.27% decrease in Spanish GDP in 2008 and a 0.48% decrease in Catalonian GDP.

2.1.2 Australia's Millennium Drought (1994-2009)

The most significant drought event that was highlighted across literature was the "Millennium Drought" that hit southeastern Australia in the late 90s and throughout the 2000s. From 2001-2009, the Murray-Darling Basin located near Victoria in southeast Australia experienced nine straight years of below median rainfall on record (Van Dijk et al., 2013). With meteorological records in the area sparse prior to the twentieth century, it is difficult to say the exact frequency of this drought but it has been estimated to have a frequency of 1500 years and could possibly have been the most severe since 1783. As is usual with any drought event, the lack of water resources caused water use restrictions in urban areas and among farmers and also increased energy prices due to lack of hydroelectric production. While the impacts of the drought have been recorded extensively, the vast majority of the economic losses came from the agricultural industry in the area. The previously mentioned water use restrictions caused agricultural water diversions within the Murray-Darling Basin to fall from 11 cubic km/year in the 1990s to 4 cubic km/year in 2009. This lack of water caused the production of year round crops, which were most impacted, to decrease significantly with rice and cotton production falling by 99% and 84%, respectively. The wheat production in the area, another year

round crop, actually increased during this time period but mainly due to the increase in crop area and the transition to wheat crops that was already occurring. Van Dijk et al. (2013) estimated that despite the increase in wheat production, the overall yield of wheat crops in the area was 20% lower than what would have been expected under normal hydrologic conditions. The effect that the drought event had on seasonal crops such as oranges did not have as immediate of an effect as year round crops, but orange production was 32% lower from 2003-2007 than production in 1999-2002. Livestock population and production in the area also fell as sheep population in the Murray-Darling Basin fell by half. While wool clip production had been decreasing since its peak in 1970, its decline was exaggerated by the drought as overall production dropped by 40% during the drought event (Kirby et al., 2012).

The economic impact of these agricultural losses was significant to the Murray-Darling Basin region and to the country of Australia. Overall, national agricultural GDP contribution fell by 16% on average during the drought period which corresponded to a 1.6% drop in national GDP. The national agricultural production dropped by $7.4 billion in 2002-2003 alone. Socially, the employment rate in the Murray-Darling Basin dropped by 3% in drought years and farm debt totals tripled over the 10 year period of 2001-2010. In order to combat the social impacts of the drought, the Australian government provided $4 billion in government relief (Heberger, 2012). The majority of this relief went to mid-size farms which were too large to have other avenues of income outside of the agricultural sector but too small to have drought resilient infrastructure in place. Although a direct and indirect analysis was not performed similar to the Catalonian

drought, the tourism economic production reduced by 5% in the Murray-Darling Basin area in 2008, representing a 0.7% drop in national tourism GDP. Effects that the drought event had on other sectors were also significant but not as easily quantified.

2.1.3 California Drought (2012-2016)

While the recent drought in California was a significant drought event that was studied extensively, the public infrastructure established to help mitigate the negative effects serves as an example of how proper governmental preparation can allow drought economic effects to be minimized overall. The drought that occurred in the Central Valley Basin in California from 2012-2016 has been widely estimated based on various assumptions in frequency analysis, with reports ranging from a 20 year to a 1,200 year drought event (Lund et al., 2018). Its impacts across all sectors were undeniable as there was a 30% of the agricultural use surface water and hydroelectric power production reduced by half. In order to reduce the impact of these losses, groundwater supplies were used to replace two-thirds of the lost surface water for agricultural use and gas-turbine power was used to replace the hydroelectric power loss. These replacements came at a cost, however, with the gas-turbine power costing a total of $2 billion more than what the hydroelectric power would have cost and increased environmental risk that was not associated with the cleaner hydroelectric power. While groundwater pumping only gave an extra expense of $600 million per year, the use of groundwater arguably had a larger effect overall due to the negative effect of land subsidence that occurred causing the San

Joaquin Valley canal capacity to reduce by 60% (Tortajada, 2017). And while groundwater replaced a majority of the surface water supply used for agriculture, 6% of cropland in California was either fallowed or idled due to the drought event. Outside of agriculture, the most well-known effect of the drought event was the wildfires that were causing massive damage environmentally and socially. The drought event caused the death of 102 million forest trees which further increased the magnitude and effects of the wildfires that were occurring over this time period.

Knowing that the impacts of this drought were significant in their own right, it is difficult to imagine what could have occurred if the impacts were not reduced due to public action. The economic impact on an individual farm was reduced due to the transition from field and seed crop to the more lucrative fruit and not crop that was already occurring. Government infrastructure was also already established based on previous drought experience and a lower dependence on the agricultural sector. The US Department of Agriculture (USDA) established multiple relief programs to allow economic flexibility to the farmers in the drought area. During this time the Environmental Quality Incentives Program (EQIP) allocated $13.7 million in water conservation efforts for California to allow for more efficient use of water in the agricultural sector. The Noninsured Crop Disaster Assistance Program (NAP) increased the value of crop insurance due to natural disaster from 55% of value for 50% of crops lost to 100% of value for 60% of crops lost. The public action also made it easier for these farmers to take out and pay off loans with programs such as the Emergency Farm Loans Program which allowed for emergency loans to replace harmful effects due to

drought for up to $500,000. Current farm loan payments were also allowed to be delayed for up to a year with the Disaster Set-Aside Program. All of these public relief efforts contributed to the economic impact of the California drought to be minimized during this time.

2.2 Economics Background

2.2.1 Using GDP as an Indicator of Economic Strength

As nations across around the world become globally integrated, droughts that once impacted areas locally are now creating an international impact. One sector that is affected by droughts that can be quantified with relative ease is the economic sector. Using GDP as an international economic indicator, the strength of drought effects on the economic sector can be determined. GDP represents the total production value generated annually by all businesses and services located within an individual country. It was developed by the United States in complement with a national income value which eventually evolved into Gross National Income (GNI) in order to provide a comprehensive indication on the economic status of the country during the Great Depression and World War II (Landefeld, 2008). Today, virtually all nations of the world use GDP as an indicator of economic strength and most countries publish this data publically to allow for international economic studies.

GDP is an economic indicator that is widely accessible internationally, but another reason that GDP was chosen as the economic indicator in this study is its robust

temporal aspect. As economies change globally and international interests and investments shift across different markets, using a more sector specific indicator such as agricultural or energy production would be more time dependent as emphasis on these markets have grown or shrank over time (Kuznets, 1961; Foquet, 2016). Although many studies have focused on the economic impact of droughts on agriculture, this study aims to take a broader approach and determine drought impact on a nation's economy as a whole. Using a more socially oriented index such as GDP per capita allows the influence of population changes to raise or lower economic strength (Boyle & McCarthy, 1999). While GDP and economic strength of a country in general is not independent of the size of its population, focusing on GDP as a whole instead of averaging values based on population allows for the effects of events such as population booms to be dispersed throughout time instead of treatment as an isolated event (Headey & Hodge, 2009). Using GDP per capita also assumes a relatively homogenous dispersion of wealth throughout a nation and does not account for changes in wealth distributions or disparities (van den Bergh & Ankal, 2014). GDP also accounts for economic strength of a country relatively isolated from other countries as global markets continue to become interconnected since it only takes production from goods and services located domestically. Other economic strength indicators similar to GDP include Gross National Product (GNP), which account for goods and services produced by a country both abroad as well as domestic, and GNI, which reflects the total income by the citizens and businesses of a country both abroad and domestic (OECD, 2020). While these economic indicators do give a broader overview of economic wealth of a country, they allow the influence of droughts located

outside of the nation's borders to impact the economic status of the country. This study looks to reflect the impact of domestic droughts on domestic economy as some areas of the world are less resilient to droughts than others.

2.2.2 Accounting for Exponential Economic Growth

When relating an economic time series such as GDP values, it is important to take the exponential nature of macroeconomics into account (Modis 2013). In order to find a correlation between drought severity and annual economic output, this exponential trend must be taken out of the data to avoid bias towards the larger GDP values in more recent years. While there are various methods for detrending both in the sense of pure time series and economic detrending, the methods chosen for this study were Logarithmic Trend Free Pre-Whitening (Log-TFPW) to account for general time series detrending as well as Logarithmic-Quadratic detrending (Log-Quad) to represent a more economics based detrending approach. These methods were chosen for their applicability and simplicity.

Since the economic figures over time show exponential growth, both of the detrending methods took the logarithm of the data to attempt to minimize the exponential trend in the data. By taking a time series oriented detrending approach, the log. GDP numbers can be analyzed similar to other hydrologic time series (Şen, 2012). In this regard, TFPW has been proven as effective for determining and extracting trends in time series data and can be applicable to this dataset. In TFPW, the assumed intrinsic linear

trend is removed from a time series of data *x*. The value of this linear slope can be produced through the Theil-Sen Slope Analysis, which entails taking the median of the linear slope between each point in the time series, shown in Equation 2-A (Sayemuzzaman & Jha, 2014). After determining this slope value, the trend is removed in the dataset through the Trend Free (TF) process shown in Equation 2-B to produce a new time series *x'* (Yue & Wang, 2002). A 1-lag Autoregressive model (AR(1)) is then run on *x'* to determine the 1-lag autocorrelation coefficient r_1 and used to remove autocorrelation from the dataset in a process called Pre-Whitening (PW), shown in Equation 2-C. An autoregressive model analyzes the autocorrelation or linear trend value between a variable *t* and itself at a certain time lag *k*. An AR(1) model gives the autocorrelation value between *t* and *t*-1 (Anderson 1942). Removing this autoregressive behavior from a time series removes the influence that a value had on the consecutive point. In this study, the trend free, pre-whitened time series *y'* was considered acceptably detrended and stationary.

$$\beta = median\left(\frac{x_j - x_i}{j - i}\right) \text{ for all } i < j \tag{2-A}$$

$$x_i' = x_i - \beta t_i \tag{2-B}$$

$$y_i' = x_i' - r_1 x_{i-1}' \tag{2-C}$$

While keeping the simplicity of the Log TFPW model but taking a macroeconomic approach, the other detrending method used was Log-Quad detrending where an assumed quadratic trend is removed from the logarithm of the economic data (Uribe & Schmitt-Grohe, 2017). While this does not account for business cycles that are

often observed in macroeconomic studies, it does all for the data to be relatively stationary and useful for the purposes of relating droughts and economic strength. It also may be in the best interest to keep this cyclical behavior in the economic data as droughts do not necessarily occur the same time scale as these cycles and the correlation gathered will be a commentary on the effect on the country in both economic boon and depression alike (Mendoza 1991). The process to remove the quadratic trend from a time series x to create the detrended, stationary time series y', which is similar to the TF process, is shown in Equation 2-D. An example of how both of the detrending methods used transformed the raw GDP of the United States is shown in Figure 2-1.

$$y_i' = x_i - ax_i^2 - bx_i - c \tag{2-D}$$

2.3 Methodology

This study used GDP values that were found on the World Bank GDP Database. This gave a time series of annual GDP (2019 US$) values for 263 countries and regions from 1960-2018 ("GDP (Current US$)", 2019). After this data was received, the GDP time series for each country was detrended in order to create stationarity in the data set. The detrending methods used were Logarithmic-Quadratic (Log-Quad) detrending and Logarithmic Trend Free Pre-Whitening (Log-TFPW). This created three separate GDP time series for each country and region: unprocessed GDP values (Raw), Log-Quad GDP values, and Log-TFPW GDP values.

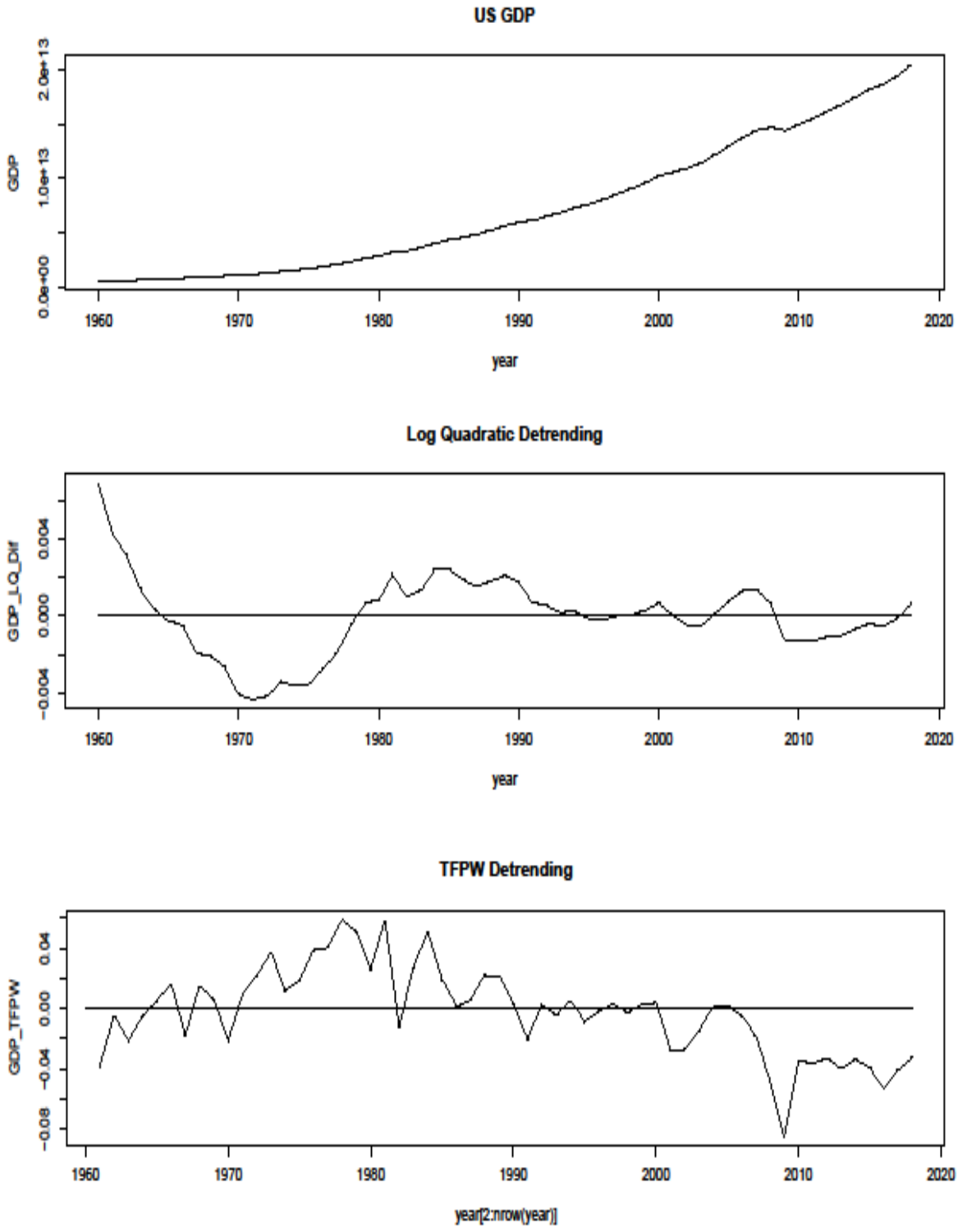

Figure 2-1: GDP Detrending Methods on United States GDP

The drought indices used in comparison with GDP were the Palmer Drought Severity Index (PDSI), 1-Month Aggregated Standardized Precipitation Index (SPI), and 1-Month Aggregated Standard Precipitation Evapotranspiration Index (SPEI). The PDSI values were provided by the Dai Dataset from the National Center for Atmospheric Research (NCAR). The dataset gave monthly time series values from 1850-2010 for each grid point on a 2.5° x 2.5° grid covering all of the global land coverage (Dai & Quian, 2004).

The SPI data was monthly time series values from 1948-2018 for each grid point on a 0.5° x 0.5° global grid. This SPI dataset was found using the International Research Institute for Climate and Society (IRI) online database, organized by Columbia University ("Data: IRI Analyses SPI SPI-PRECL0p5_1-Month.", 2020). According to the documentation of the dataset, the SPI values for each month were derived using a Log-Pearson Type III (LP3) distribution which is based on the mean μ, standard deviation σ, and skewness γ of the data (Amin et al., 2016). The method for calculating the probability for a given precipitation x for this distribution is best found using the Probability Density Function (PDF), which is shown in Equations 2-E – 2-H. After finding the probability of the monthly precipitation, the SPI value is the Z-value of the standard normal distribution which corresponds with that probability.

$$f(x) = \frac{1}{|\alpha| x \Gamma(\beta)} \left[\left(\frac{\ln(x) - \xi}{\alpha} \right)^{\beta - 1} \right] e^{-\frac{\ln(x) - \xi}{\alpha}} \quad \text{(2-E)}$$

$$\alpha = \frac{\sigma \gamma}{2} \quad \text{(2-F)}$$

$$\beta = \frac{4}{\gamma^2} \tag{2-G}$$

$$\xi = \mu - \frac{2\sigma}{\gamma} \tag{2-H}$$

The SPEI values were given by the Spanish National Research Council (CSIC) (Beguería, 2017). This dataset gave monthly time series values from 1901-2015 for each grid point on a 0.5° x 0.5° global grid. The SPEI values for this dataset were calculated using the log-logistic distribution which is also based on the mean μ, standard deviation σ, and skewness γ of the data (Ahmed, 1988). The method for calculating the probability for a given precipitation *x* for this distribution is best found using the Probability Density Function (PDF), which is shown in Equations 2-I – 2-M. After finding the probability of the monthly precipitation, the SPEI value is the Z-value of the standard normal distribution which corresponds with that probability. (Γ(x) refers to the gamma function of x.)

$$f(x) = \frac{\left(\frac{x-\rho}{\alpha}\right)^{-\frac{1}{\beta}}}{\alpha(x-\rho)\left(1+\left(\frac{x-\rho}{\alpha}\right)^{-\frac{1}{\beta}}\right)^2} \tag{2-I}$$

$$\mu = \rho + \alpha * A(1,\beta) \tag{2-J}$$

$$\sigma = \alpha^2 * A(2,\beta) - \mu^2 \tag{2-K}$$

$$\gamma = \frac{A(2,\beta) - 3 * A(2,\beta) * A(1,\beta) + 2A^3(1,\beta)}{[A(2,\beta) - A^2(1,\beta)]^{\frac{3}{2}}} \tag{2-L}$$

$$A(j,c) = \frac{\Gamma(1+jc) * \Gamma(1-jc)}{\Gamma(2)} \tag{2-M}$$

Once these drought index datasets were received, the process of aggregating the monthly values to give annual drought severity was performed. For each of the monthly values within a given year, the absolute value of the sum of the negative index values gave the annual drought severity value for that year. Positive index values throughout the year were ignored in this summation. Once the annual drought severity was determined for each of the years at each of the grid points, the points were aggregated based on their respective countries. For countries that contained more than one data point, the annual drought severities for each year were averaged among all of the contained points. This gave each country a single time series for each of the drought indices. Countries that did not include a grid point were ignored for the study.

After associating both the GDP and drought index time series with each country, the Pearson product moment correlation coefficient (R) was determined to quantify the correlation between each drought index and GDP detrending method (Hauke & Kossowski, 2011). The method for calculating R is shown in Equations 2-N – 2-Q. The detrended US GDP series were plotted on scatter plots against the annual PDSI severity values in Figure 2-2. The raw, Log-Quad, and Log TFPW GDP correlation values were 0.61, -0.02, and -0.36, respectively.

$$R = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \tag{2-N}$$

$$S_{xy} = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} \tag{2-O}$$

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} \tag{2-P}$$

$$S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n} \qquad (2\text{-Q})$$

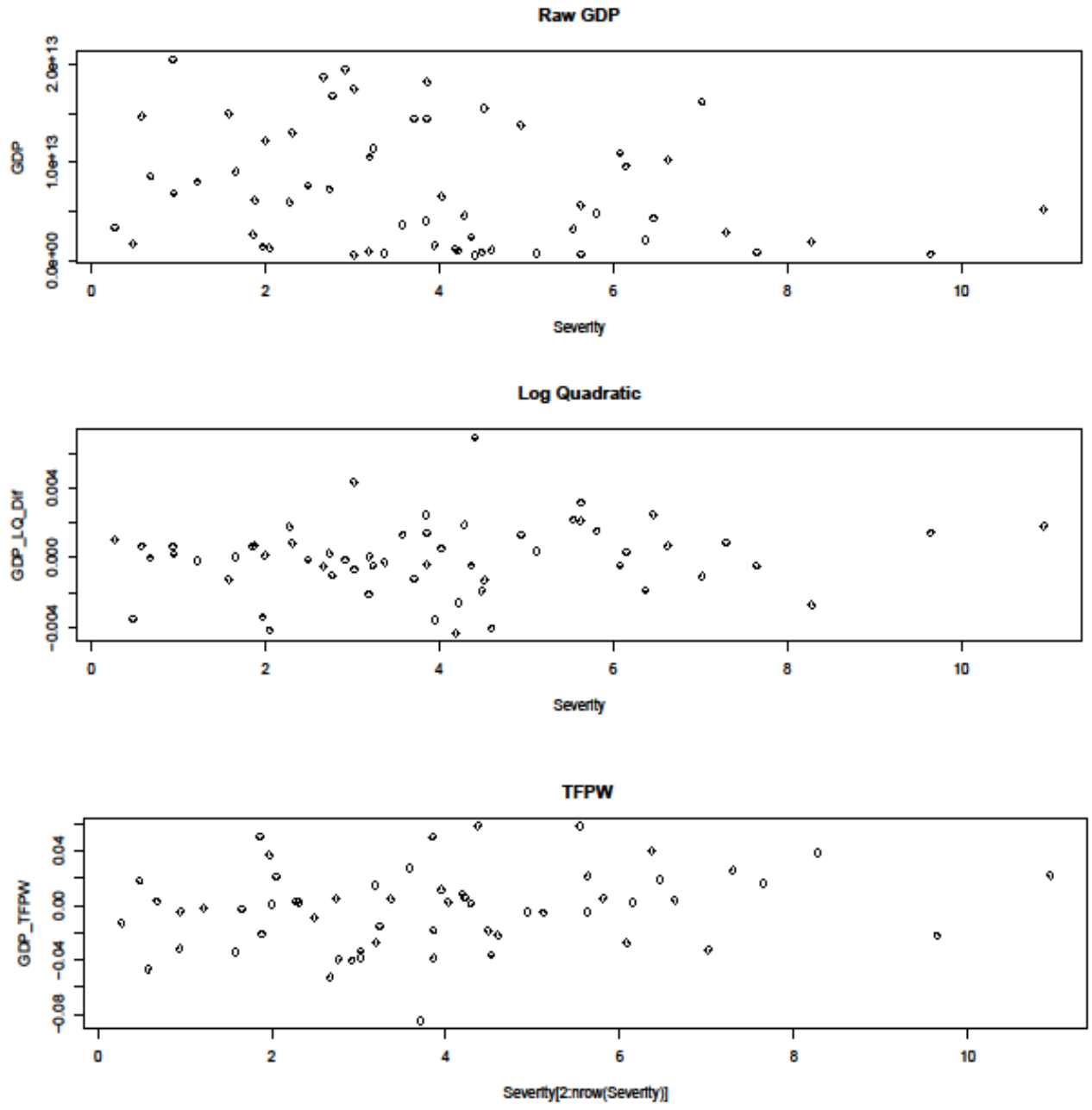

Figure 2-2: GDP Detrending Methods PDSI Scatter Plots on United States GDP

An important assumption to note when using R to determine correlation is that the results show the strength of the linear relationship between variables *x* and *y*. This means that the non-linear relationships between variables may not be completely captured by quantifying correlation in this way. However, it does provide a general sense of the trend between the two variables and is a commonly used practice in quantifying correlation.

2.4 Results and Discussion

By comparing both of the Log-Quad and Log TFPW detrended global GDP time series with the global averaged PDSI time series from 1960-2005 in Figures 2-3 – 2-4, some negative correlation is visually presented on a global scale, especially on the Log TFPW. This was an initial indicator that Log TFPW may be a more applicable detrending method to use when considering the negative effects that droughts can have on a global scale. The positive correlation that is seen in the latter half of the Log-Quad time series may show that the global economy is moving more independently of drought indices.

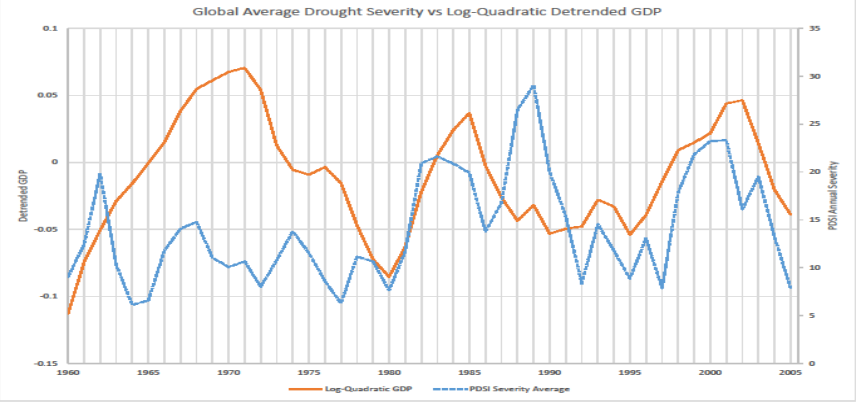

Figure 2-3: Global Average Annual PDSI Drought Severity & Log-Quad Detrended Global GDP

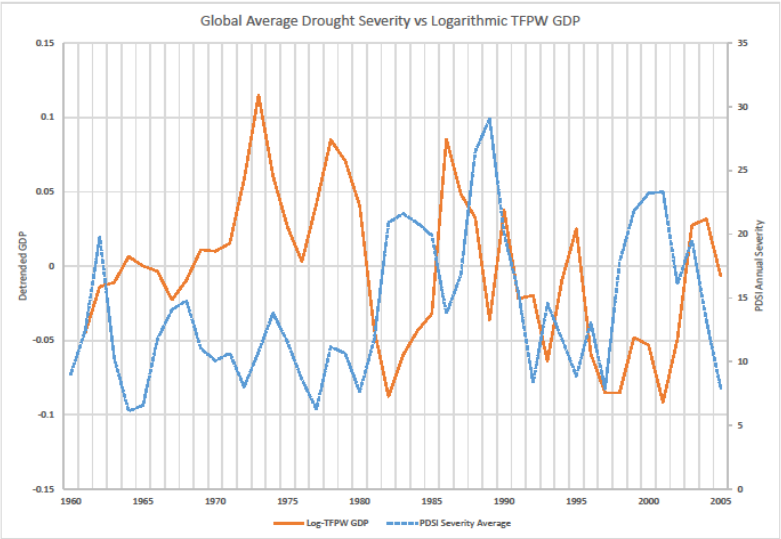

Figure 2-4: Global Average Annual PDSI Drought Severity &
Log TFPW Detrended Global GDP

In order to further investigate the applicability of these detrending method as well as test their versatility, the correlation between raw or detrended GDP and drought indices such as PDSI, SPI, or SPEI were found for each nation with applicable data. The correlation values were then summarized in boxplots. The boxplots for PDSI shown in Figure 2-5 show that while both detrending methods show a stronger negative correlation than the raw GDP values, the Log-Quad detrending correlation plot showed a wider range of negative trends than the Log TFPW detrending correlation values. This was then contradicted by Figure 2-6 representing the SPI boxplots which show that the Log TFPW method to be the only applicable method for determining negative drought effects on GDP as it was the only range of correlation values to have at least 50% of the countries tested giving a negative trend between GDP and SPI. Finally the SPEI boxplots shown in

Figure 2-7 give all GDP series, including raw values, being applicable to represent negative drought effects on GDP on a global scale with the raw GDP values giving a wider negative range and the Log TFPW GDP values giving a slightly stronger negative range among the detrending methods.

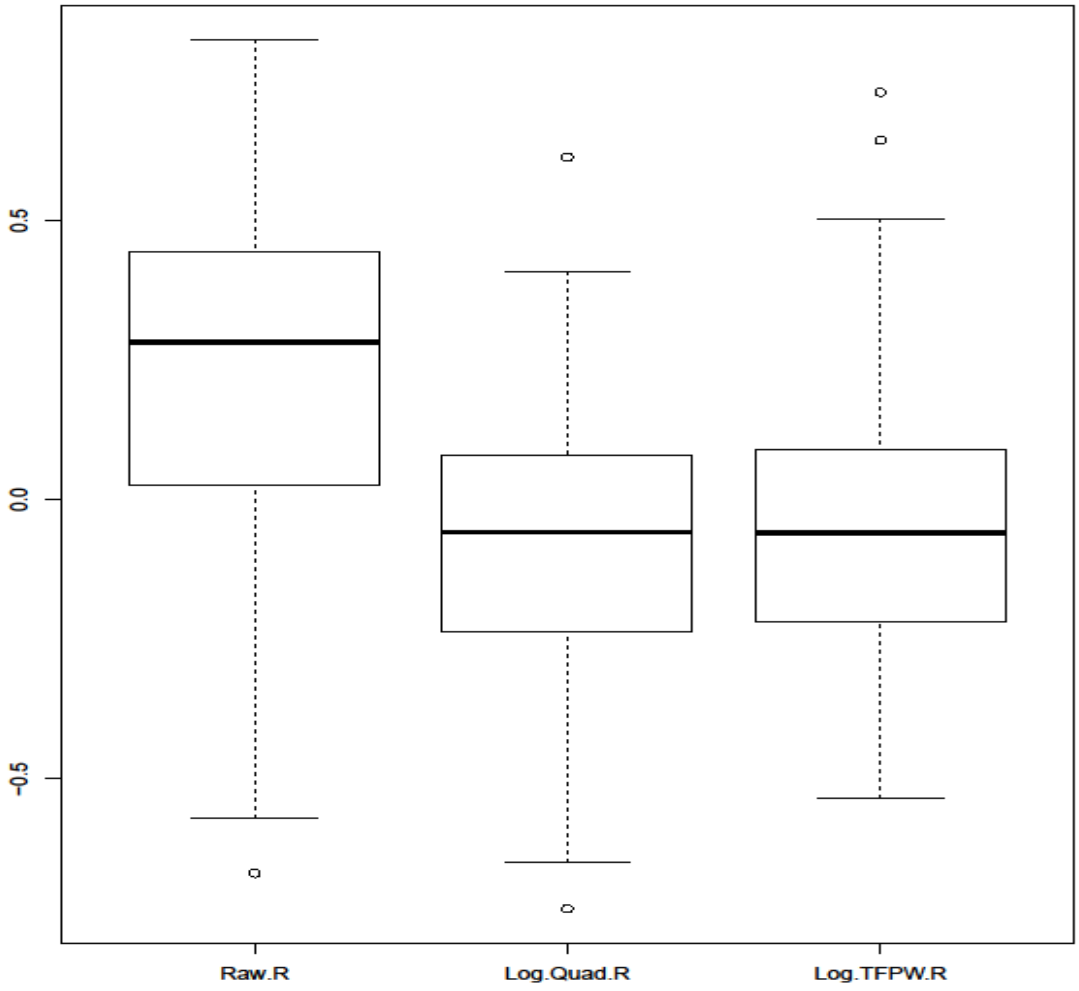

Figure 2-5: World GDP PDSI National Correlation Values

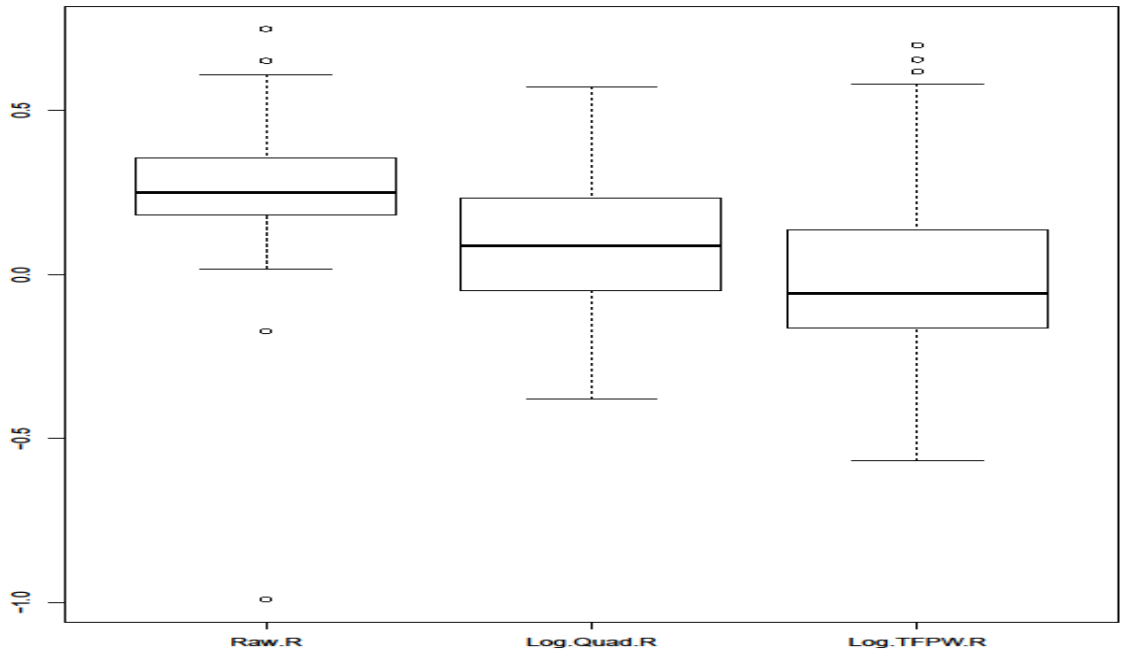

Figure 2-6: World GDP SPI National Correlation Values

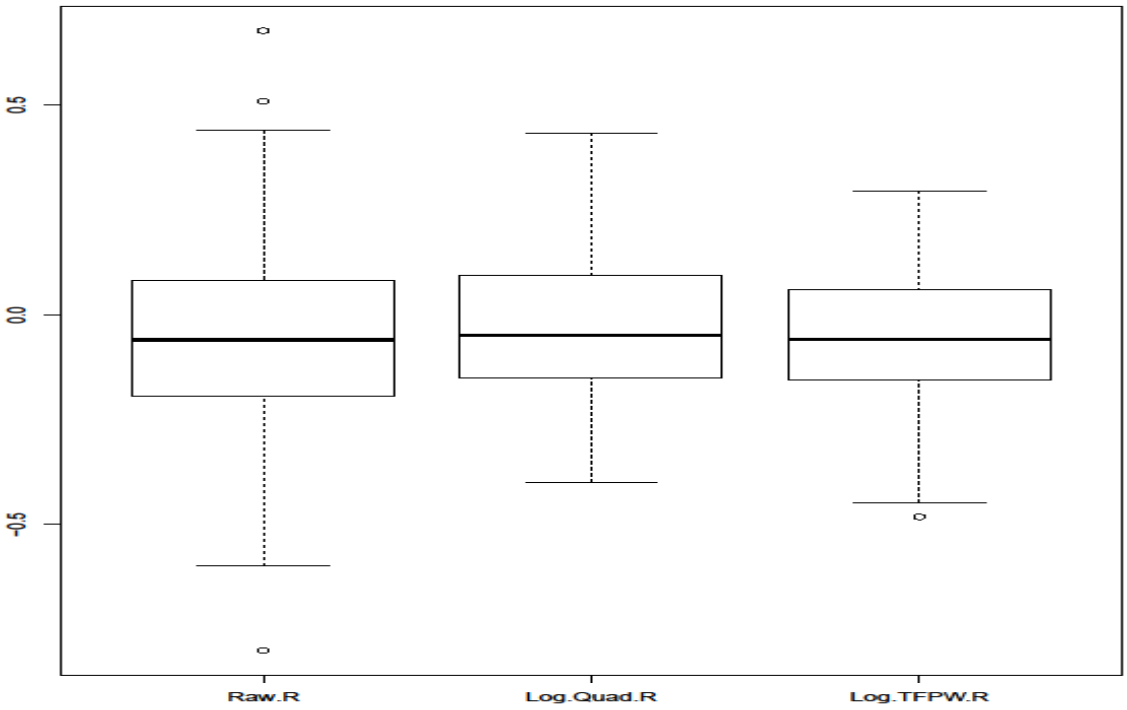

Figure 2-7: World GDP SPEI National Correlation Values

Keeping in mind the spatial component of droughts, this same process of comparing GDP and drought indices through boxplots was broken down further to show the correlation by continental region. Since the PDSI boxplots gave stronger negative correlation to the detrending methods, this continental analysis was only performed using the PDSI drought index values. Comparing the raw GDP correlations to the detrended GDP in Figure 2-8 shows that the detrending methods are useful in determining negative trends between drought and economic impact. Between the detrending methods, while Log-Quad showed stronger negative correlation in some regions, especially Europe and Asia, Log TFPW had more negative correlation values among the different regions with all continents having at least 50% of countries showing a negative trend between PDSI and GDP except for Europe. Log TFPW also gave a stronger negative correlation value for North American countries than Log-Quad GDP.

By showing the correlation values on an international map, the regional applicability of negative trend between droughts and economic strength can be further concentrated. Of all of the drought indices and GDP detrending methods used, the Log-Quad GDP correlated with PDSI shown in Figure 2-9 produced the negative trend with the greatest magnitude with Bosnia and Herzegovina with a correlation value of -0.73. This was in line with the surrounding area as most of the countries in Eastern Europe showed a negative trend with drought when correlating Log-Quad GDP with PDSI. While Log-Quad GDP produced the strongest negative correlations, the Log TFPW GDP with PDSI in Figure 2-10 produced the most widespread significant negative correlation values. This was especially true in North America as Canada, the United States, and

Mexico all had correlation values that were more negative than -0.25. Another significant area to note under this drought index and detrending method is near the southeast coast of Africa as one of the highest negative correlation values of this combination was produced in Zimbabwe of -0.51.

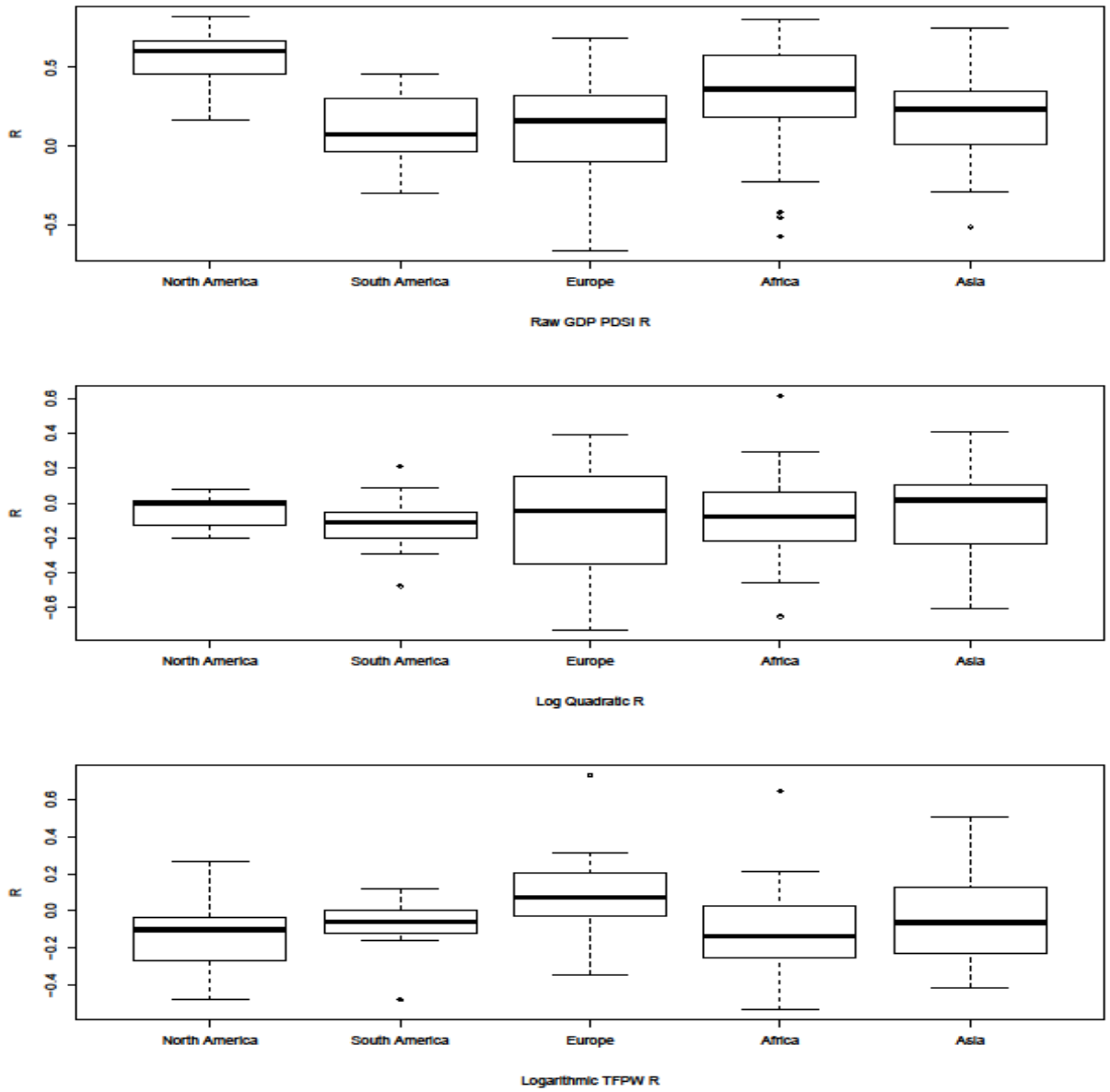

Figure 2-8: Continental GDP PDSI National Correlation Values

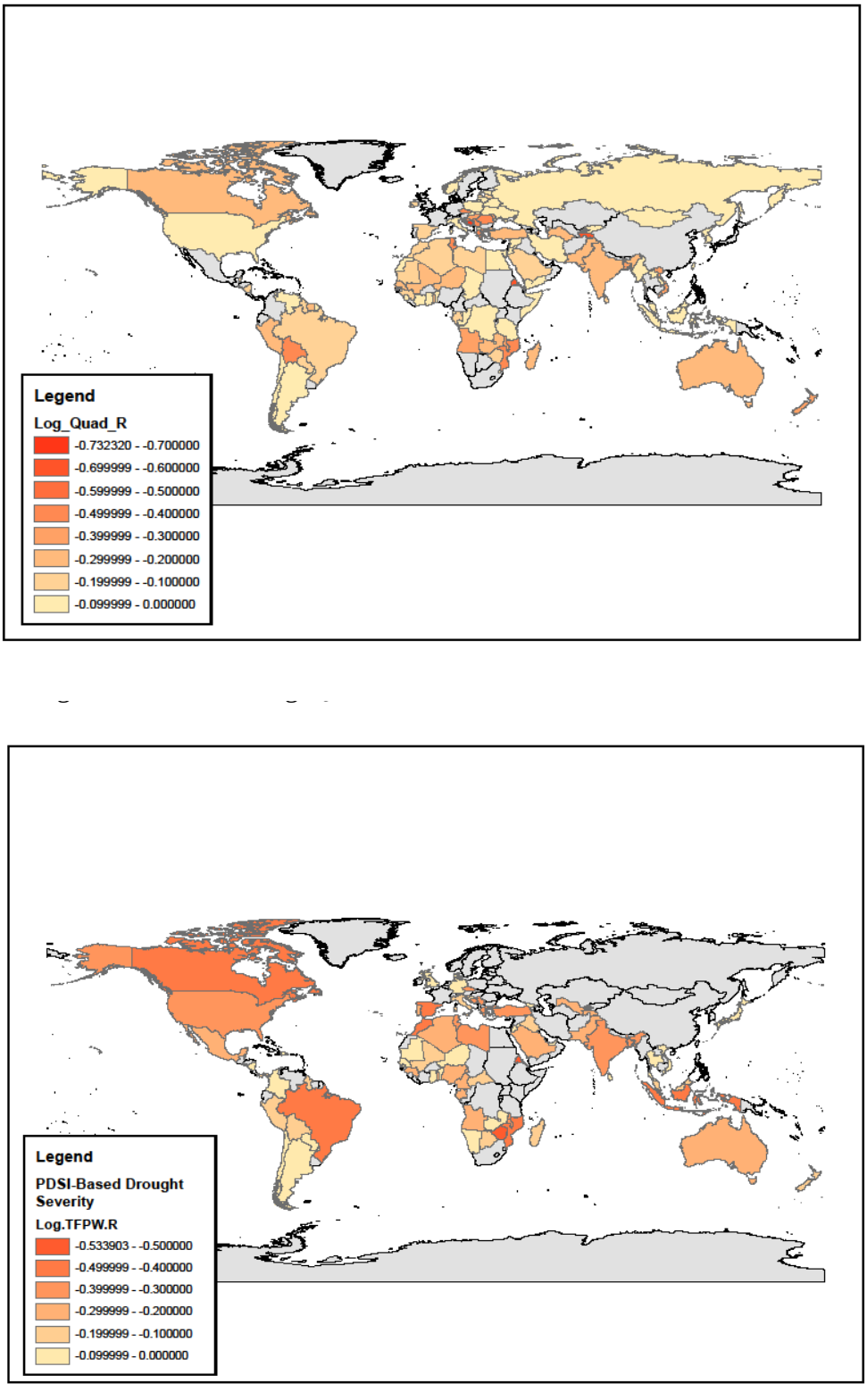

Figure 2-10: Global Log TFPW GDP PDSI National Correlation Values

Of all of the detrending methods and drought index correlations calculated, the pairing of Log-Quad detrended GDP and SPI produced the negative trends with the least magnitude and produced the least number of countries with negative correlation values, as seen in Figure 2-11. However, this combination did produce the strongest negative correlation for the Chinese economy with a value of -0.22 and the strongest negative trend under this correlation was Uganda at -0.38. Alternatively, the correlations between Log TFPW GDP and SPI in Figure 2-12 had slightly more widespread negativity in correlation across the globe and produced some of the strongest negative correlations of any combination for the region of Western Europe. It also produced a stronger maximum negative correlation with Somalia having a negative correlation of -0.57.

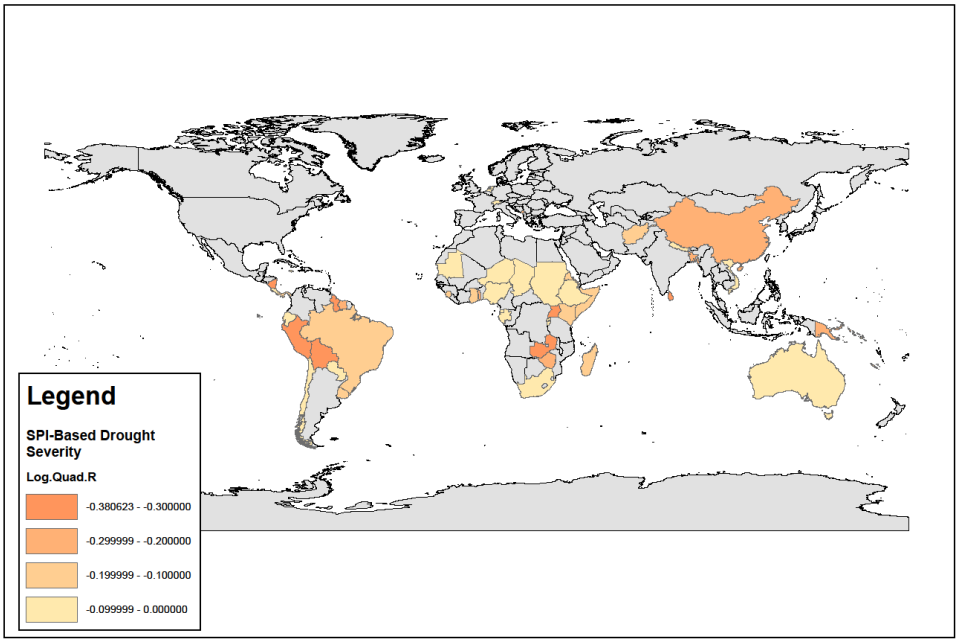

Figure 2-11: Global Log-Quad GDP SPI National Correlation Values

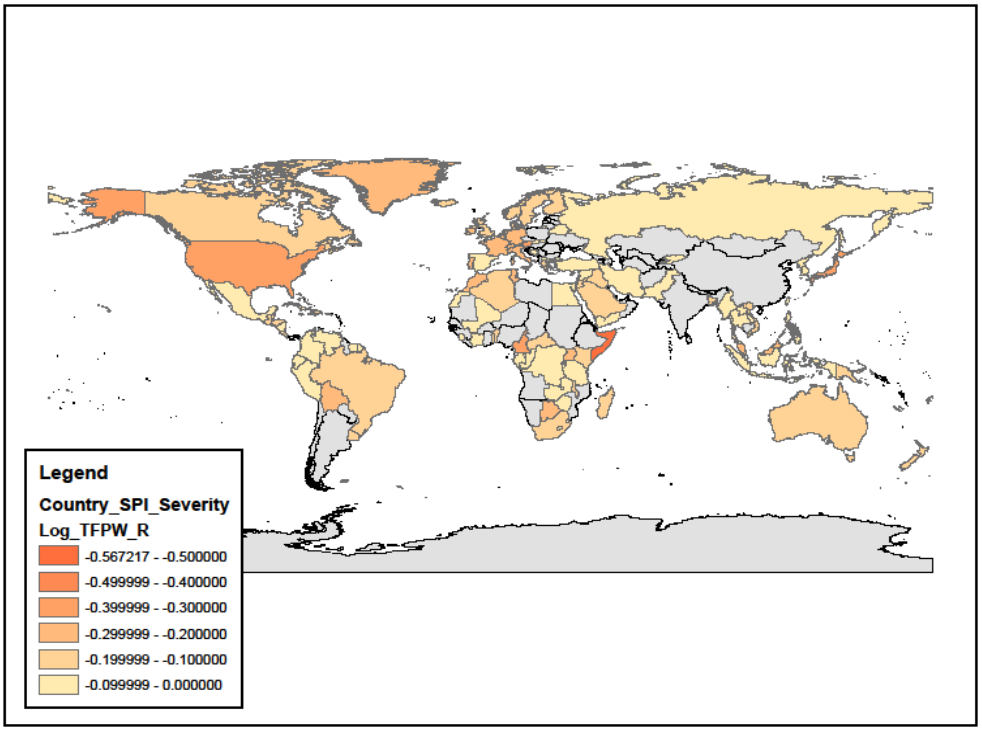

Figure 2-12: Global Log TFPW GDP SPI National Correlation Values

The last drought index tested was SPEI which was much more applicable to the eastern hemisphere than the previous drought indices. SPEI and Log-Quad detrending produced stronger correlations overall, seen in Figure 2-13. The strongest correlation for this method was -0.40 found in Burundi, with several other nations such as Bolivia, the Faroe Islands, and Mongolia producing correlations below -0.35. While the Log-Quad detrending produced stronger negative correlations on average, the correlations between SPEI and Log TFPW shown in Figure 2-14 gave more widespread negative correlations across the eastern hemisphere. A majority of the countries in Africa, Asia, and Eastern Europe produced a negative correlation using this method with the strongest negative correlation occurring in Slovenia at -0.48.

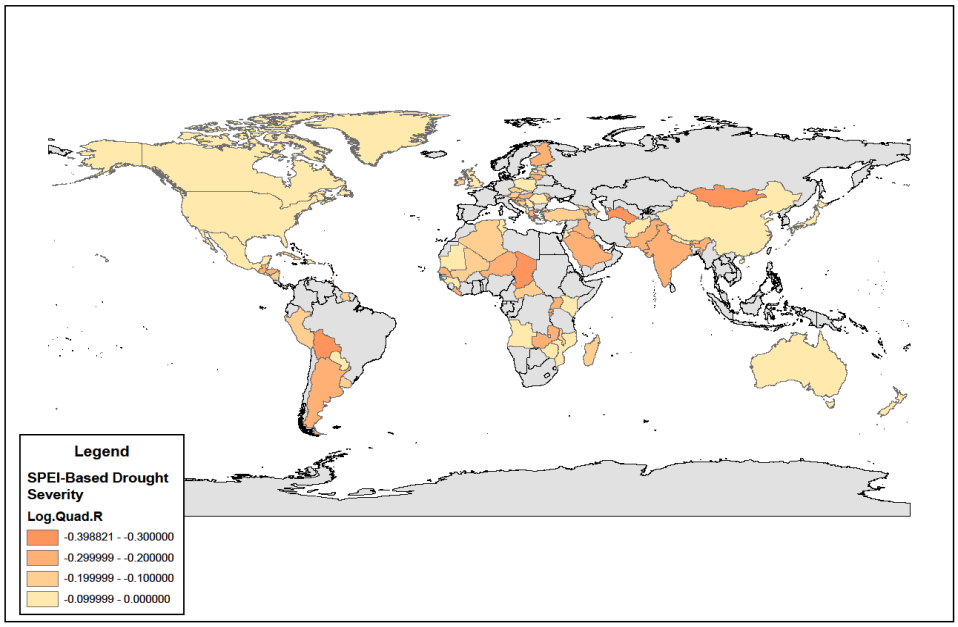

Figure 2-13: Global Log-Quad GDP SPEI National Correlation Values

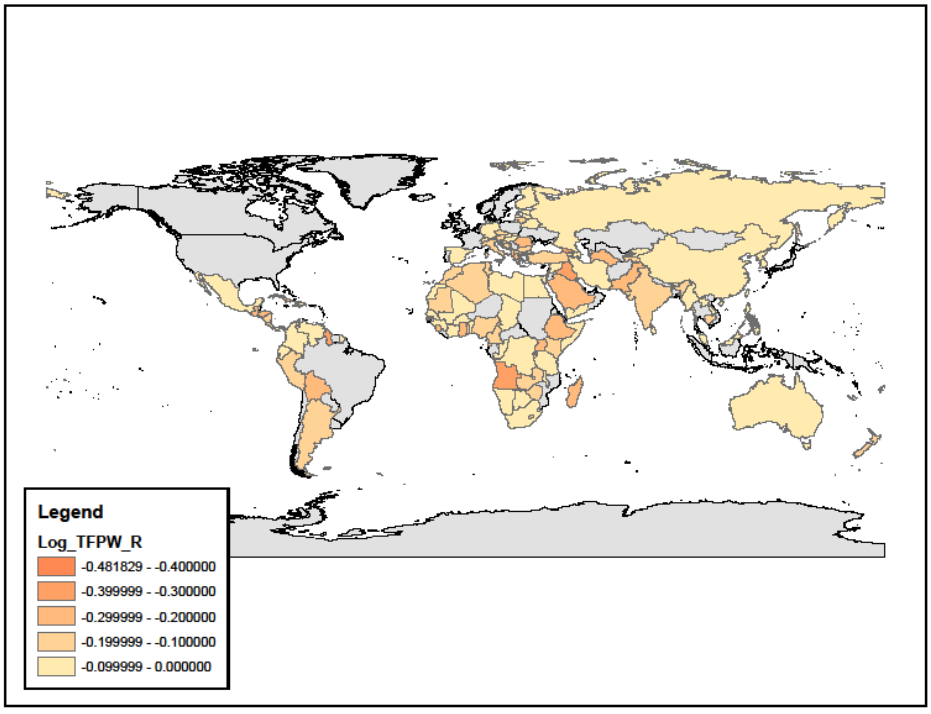

Figure 2-14: Global Log TFPW GDP SPEI National Correlation Values

In order to succinctly summarize the correlation processes that were performed, a statistical t-test of negative significance with 95% confidence was performed for both drought indices and both detrending methods on the correlation values for all the countries with data. This test helped reaffirm some of the statements made previously about the different drought indices and detrending methods. PDSI on average was found to be more widely applicable than any of the other indices to signify the negative impact between drought severity and national GDP. It also helped in eliminating some of the countries where negative trends could have been calculated by cross correlation with external variables. Among all of the countries of the world, the only country to have negative statistical significance in all drought indices no matter the detrending method was Madagascar. These statistical significance tests for PDSI, SPI, and SPEI are shown in Figures 2-15 – 2-17, respectively.

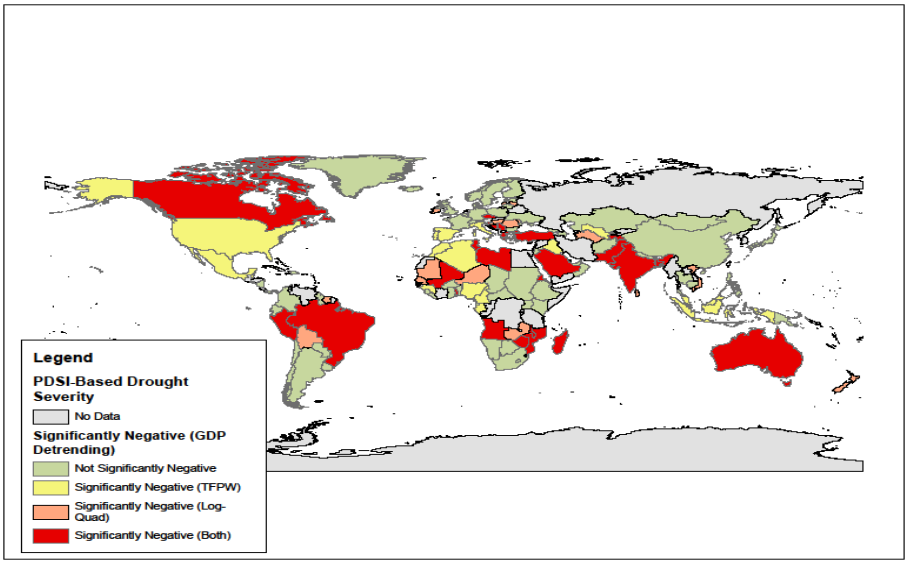

Figure 2-15: Global GDP PDSI National Significantly Negative Correlation Values

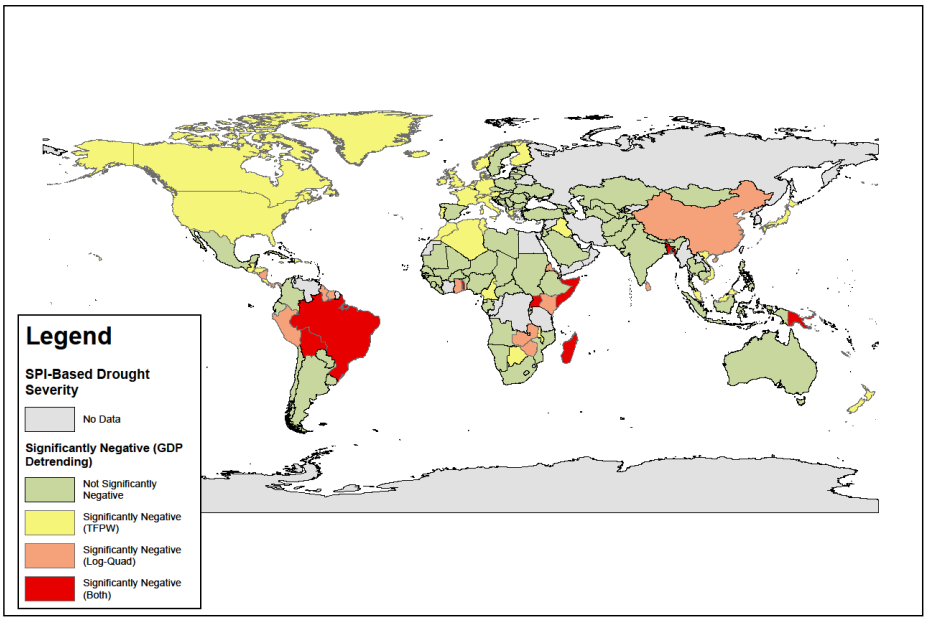

Figure 2-16: Global GDP SPI National Significantly Negative Correlation Values

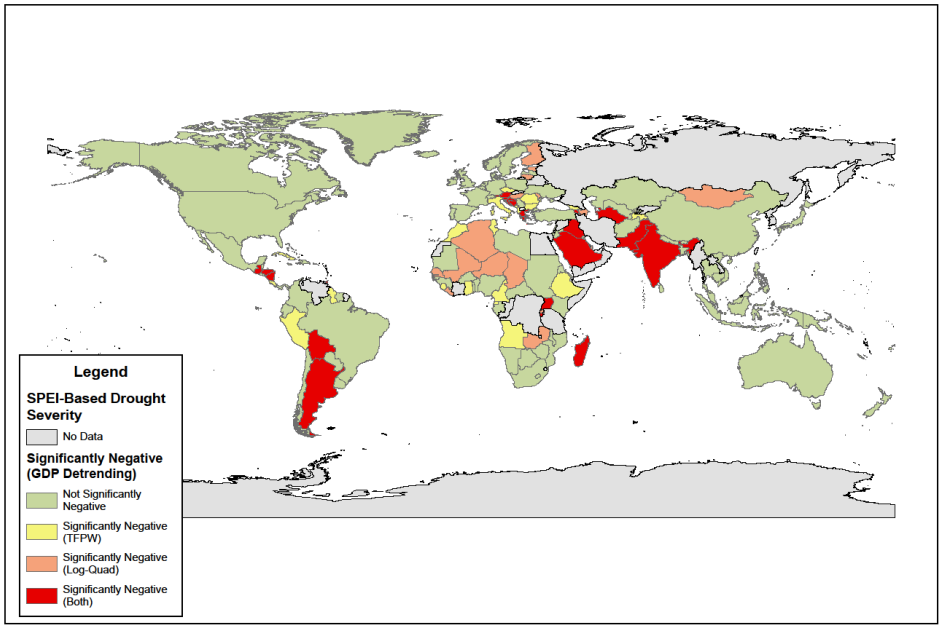

Figure 2-17: Global GDP SPEI National Significantly Negative Correlation Values

2.5 Conclusion

Through all of the methods and comparisons performed in this study, almost all countries that included both drought and GDP data had some combination of drought index and GDP detrending method that produced a significantly negative trend in GDP. This means that the effects that droughts have on a nation's economic strength, while not necessarily independent of the economic development of the country, is an international phenomenon. The particular combination of drought index and GDP detrending appears to be regional in its effectiveness as combinations that produced strong negative correlations for a single country usually produced similar correlations to those nations surrounding it. On an international scale, PDSI was the most effective drought index on average while Log TFPW seemed to be more effective on average than Log-Quad in terms of GDP detrending. It is important to note that the correlation values presented in this study do not reflect the magnitude of effect that drought has on economic strength but rather the consistency that the effects occur. A high negative correlation value between drought index and detrended GDP gives a high probability that the drought will affect GDP, but does not speak on how drastic that effect will be.

2.6 Resources

Ahmad, M.I., et al. “Log-Logistic Flood Frequency Analysis.” Journal of Hydrology, vol. 98, no. 3-4, 15 Apr. 1988, pp. 205–224., doi:10.1016/0022-1694(88)90015-7.

Amin, M. T., M. Rizwan and A. A. Alazba. "A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan" Open Life Sciences, 11.1 (2016): 432-440. Retrieved 20 Mar. 2020, from doi:10.1515/biol-2016-0057

Anderson, R. L. “Distribution of the Serial Correlation Coefficient.” The Annals of Mathematical Statistics, vol. 13, no. 1, 1942, pp. 1–13. JSTOR, www.jstor.org/stable/2236157. Accessed 20 Mar. 2020.

Beguería, Santiago, et al. “SPEIbase V.2.5.” DIGITAL.CSIC, Spanish National Research Council, 25 July 2017, digital.csic.es/handle/10261/153475.

Dai, A., K. E. Trenberth, and T. Qian, 2004: A global data set of Palmer Drought Severity Index for 1870-2002: Relationship with soil moisture and effects of surface warming. J. Hydrometeorology, 5, 1117-1130.

“Data: IRI Analyses SPI SPI-PRECL0p5_1-Month.” International Research Institute, Columbia University, 9 Mar. 2020, https://iridl.ldeo.columbia.edu/SOURCES/.IRI/.Analyses/.SPI/.SPI PRECL0p5_1-Month/.

Fouquet, R. Path dependence in energy systems and economic development. Nat Energy 1, 16098 (2016). https://doi.org/10.1038/nenergy.2016.98

Freire-González, Jaume, et al. “The Economic Impacts of Droughts: A Framework for Analysis.” Ecological Economics, vol. 132, Feb. 2017, pp. 196–204., doi:10.1016/j.ecolecon.2016.11.005.

“GDP (Current US$).” World Bank Data, World Bank Group, 2019, data.worldbank.org/indicator/NY.GDP.MKTP.CD

G. E. Boyle & T. G. McCarthy (1999) Simple measures of convergence in per capita GDP: a note on some further international evidence, Applied Economics Letters, 6:6, 343-347, DOI: 10.1080/135048599353041

Harou, J. J., Medellín-Azuara, J., Zhu, T., Tanaka, S. K., Lund, J. R., Stine, S., Olivares, M. A., and Jenkins, M. W. (2010), Economic consequences of optimized water management for a prolonged, severe drought in California, Water Resour. Res., 46, W05522, doi: 10.1029/2008WR007681.

Hauke, Jan and Tomasz Kossowski. "Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data" Quaestiones Geographicae, 30.2 (2011): 87-93. Retrieved 20 Mar. 2020, from doi: 10.2478/v10117-011-0021-1

Headey, D.D. and Hodge, A. (2009), The Effect of Population Growth on Economic Growth: A Meta-Regression Analysis of the Macroeconomic Literature. Population and Development Review, 35: 221-248. doi:10.1111/j.1728-4457.2009.00274.x

Heberger M. (2012) Australia's Millennium Drought: Impacts and Responses. In: Gleick P.H. (eds) The World's Water. The World's Water. Island Press, Washington, DC

Iceland, Charles. "A Global Tour of 7 Recent Droughts." World Resources Institute, 2 Aug. 2019, www.wri.org/blog/2015/06/global-tour-7-recent-droughts.

Julia Martin-Ortega, Mikel González-Eguino, Anil Markandya; The costs of drought: the 2007/2008 case of Barcelona. Water Policy 1 June 2012; 14 (3): 539–560. doi: https://doi.org/10.2166/wp.2011.121

Kirby, Mac, et al. "The Economic Impact of Water Reductions During the Millennium Drought in the Murray-Darling Basin." AARES 2012 Conference (56th), Australian Agricultural and Resource Economics Society, Feb. 2012, Fremantle, Australia. Conference Paper

Kuznets, Simon. "Economic Growth and the Contribution of Agriculture: Notes on Measurement." 1961 IAAE Conference, International Association of Agricultural Economists August 1961, Hotel Casino de la Selva, Cuernavaca, Morelos, Mexico. Conference Presentation.

Landefeld, J. Steven, Eugene P. Seskin, and Barbara M. Fraumeni. 2008. "Taking the Pulse of the Economy: Measuring GDP." Journal of Economic Perspectives, 22 (2): 193-216.

Lund, Jay, et al. “Lessons from California’s 2012–2016 Drought.” Journal of Water Resources Planning and Management, vol. 144, no. 10, Oct. 2018, https://doi.org/10.1061/(ASCE)WR.1943-5452.0000984.

March, H., Domènech, L. & Saurí, D. Water conservation campaigns and citizen perceptions: the drought of 2007–2008 in the Metropolitan Area of Barcelona. Nat Hazards 65, 1951–1966 (2013). https://doi.org/10.1007/s11069-012-0456-2

Mendoza, Enrique G. “Real Business Cycles in a Small Open Economy.” The American Economic Review, vol. 81, no. 4, 1991, pp. 797–818. JSTOR, www.jstor.org/stable/2006643. Accessed 20 Mar. 2020.

Modis, Theodore. “Long-Term GDP Forecasts and the Prospects for Growth.” Technological Forecasting and Social Change, vol. 80, no. 8, Oct. 2013, pp. 1557–1562., doi:10.1016/j.techfore.2013.02.010.

OECD (2020), Gross national income (indicator). doi: 10.1787/8a36773a-en (Accessed on 12 March 2020)

Sayemuzzaman, Mohammad, and Manoj K. Jha. “Seasonal and Annual Precipitation Time Series Trend Analysis in North Carolina, United States.” Atmospheric Research, vol. 137, Feb. 2014, pp. 183–194., doi:10.1016/j.atmosres.2013.10.012.

Şen, Zekâi. “Innovative Trend Analysis Methodology.” Journal of Hydrologic Engineering, vol. 17, no. 9, Sept. 2012, pp. 1042–1046., doi:10.1061/(asce)he.1943-5584.0000556

Tortajada, Cecilia, et al. “The California Drought: Coping Responses and Resilience Building.” Environmental Science & Policy, vol. 78, Dec. 2017, pp. 97–113., doi:10.1016/j.envsci.2017.09.012.

Uribe, Martin, and Stephanie Schmitt-Grohe. Open Economy Macroeconomics. Princeton University Press, 2017.

van den Bergh, Jeroen; Antal, Miklós (2014) : Evaluating Alternatives to GDP as Measures of Social Welfare/Progress, WWWforEurope Working Paper, No. 56, WWWforEurope, Vienna

Van Dijk, Albert I. J. M, et al. "The Millennium Drought in Southeast Australia (2001-2009): Natural and Human Causes and Implications for Water Resources, Ecosystems, Economy, and Society." Water Resources Research, vol. 49, no. 2, 25 Feb. 2013, pp. 1040–1057., doi:10.1002/wrcr.20123.

Yue, S., and Wang, C. Y., Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test, Water Resour. Res., 38(6), doi:10.1029/2001WR000861, 2002.

CHAPTER 3: DROUGHT FORECASTING AND REGIONAL ANALYSIS

3.1 Introduction

3.1.1 US Climate Regions

The National Climatic Data Center (NCDC) under the National Oceanic and Atmospheric Administration (NOAA) has divided the United States into nine climatically similar regions for spatial comparison of climatic data (Enloe, 2020). These regions were derived by a 1984 study by Karl & Koss according to similar average temperature values. These national climatic regions are shown in Figure 3-1. One of the purposes of this study is to determine if these NOAA climatic regions are appropriate to apply for drought comparison.

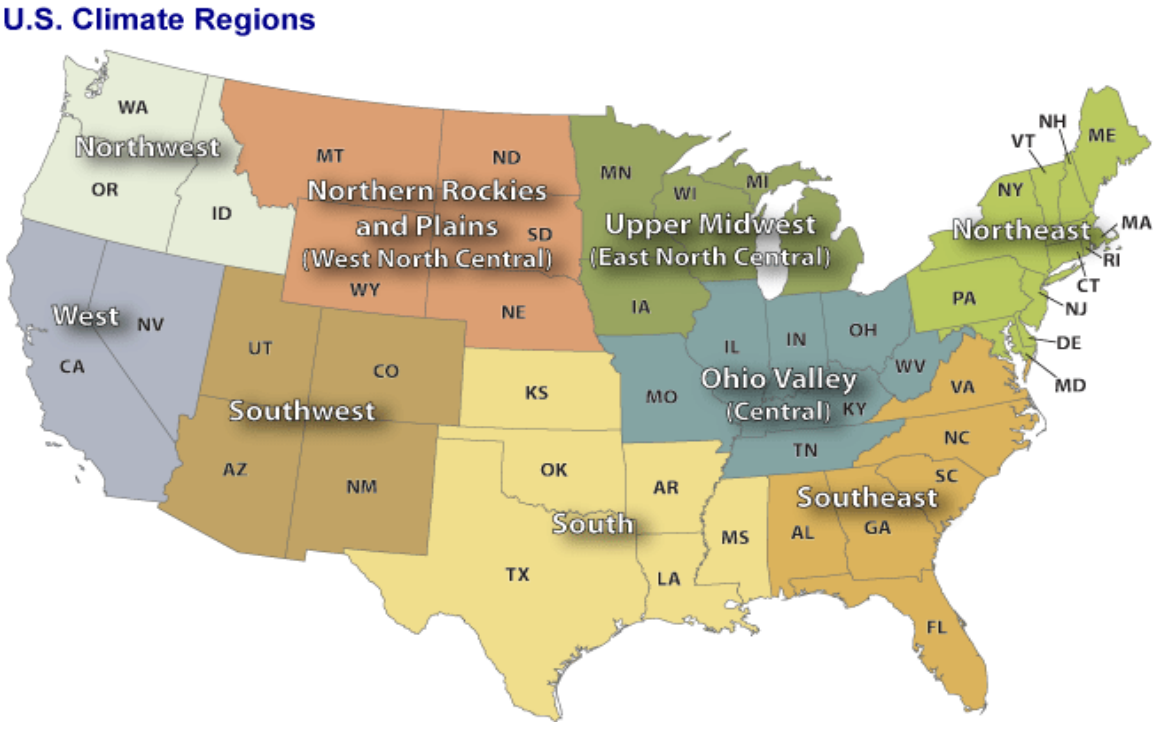

Figure 3-1: NOAA US Climatic Regions
(Enloe, Sanchez-Lugo. "US Climatic Regions." NOAA, 6 March 2020.)

3.1.2 Time Series Modeling

It is useful to consider a series of drought index values in a certain area as a time series in order to perform key statistical analyses. A temporal component is important to consider as droughts, unlike other natural disasters, occur over a certain duration of time instead of a relatively instantaneous event (Mishra & Singh, 2010). Of the various time statistical analyses used in time series studies, stochastic modeling is used often to suggest that the semi randomized behavior of weather patterns and droughts in particular can be described in a mathematical model (Katz & Parlange, 1998). An important assumption of the use of any stochastic modeling technique is that the data is assumed stationary and that no underlying trend would have influence over any specific point or set of points. The specific use of an Auto Regressive Integrated Moving Average (ARIMA) model in this study was from of the use of the Palmer Drought Severity Index (PDSI) because its derivation includes the influence of the value from the previous time step which would encourage auto regressive behavior (Jacobi et al., 2013). Another reason for the use of the ARIMA model was the popularity in the use of ARIMA models for forecasting precipitation and drought data (Mishra & Desai, 2005; Mishra & Singh, 2011). This study aimed to determine if the use of ARIMA modeling for forecasting purposes was appropriate for the climatic regions given.

3.1.3 Spatio-Temporal Analysis

Another component of droughts that is important to consider is the scale of their spatial influence. Mishra & Desai (2010) stated that "a regional drought is assumed when a significant fraction of the total area of the region is under drought conditions." This means that in order to fully understand if the climatic regions presented by the NOAA are applicable to droughts, they need to be studied holistically. It is also important to determine if some areas within these regions are more affected by droughts than others as droughts are not always homogenous in their effects (Mishra & Singh, 2010). A good method for summarizing the effects of drought on an entire region are through the use of Severity Area Frequency (SAF) curves. This technique has been used in multiple areas but studies that apply this analysis on this large of spatial scale are few (Mishra & Desai, 2005; Amirataee et al. 2018). The previous studies applying this technique also often focused on a singular drought index rather than the comparison of multiple indices. This study can not only determine if the climatic regions are useful for spatio-temporal analyses and comparisons but also which drought index would be most applicable.

3.2 Methodology

3.2.1 Stochastic Modeling

For the stochastic modeling portion of the statistical analysis, a single state within each of the national climatic regions were chosen to represent the region. For this study, the Southeast region was represented by South Carolina (SC), South region by Texas

(TX), Southwest region by Colorado (CO), West region by California (CA), Northwest region by Washington (WA), Northern Rockies and Plains region by South Dakota (SD), Upper Midwest region by Wisconsin (WI), Ohio Valley region by Illinois (IL), and Northeast region by New York (NY). Once each state was selected, a Palmer Drought Severity Index (PDSI) time series was found on the Climate Data Online (CDO) Divisional Select database run by the National Climatic Data Center (NCDC) through the National Oceanic and Atmospheric Administration (NOAA). The CDO database gave a single PDSI time series as a state-wide summary for each state selected (Baldwin, 2020). Each time series was split into a training period for the stochastic models of 1895-1989 and a testing period of 1990-2019.

Once all of the PDSI time series data was accumulated and processed, the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) was run on each of the state training period time series at multiple lag values. When the ACF, which determines the correlation value of a variable with itself at a particular time lag and all other previous lags, is run at multiple lags and plotted in a correlogram, the significant time lags used in the Autoregressive (AR) portion of the ARIMA stochastic models are clearly shown (Anderson, 1942). If the correlogram of the ACF shows exponential decay with no specific lags breaking this trend, then there are no significant lags suggested for use in an AR model. The process for calculating the autocorrelation value $\hat{\rho}$ of a variable x with n observations at a time lag L is shown in Equations 3-A – 3-B.

$$\hat{\gamma}(L) = \frac{1}{n}\sum_{i=1}^{n-L}\left((x_{i+L} - \bar{x}) * (x_i - \bar{x})\right) \quad \text{(3-A)}$$

$$\hat{\rho}(L) = \frac{\hat{\gamma}(L)}{\hat{\gamma}(0)} \tag{3-B}$$

Similarly, when the PACF, which determines the correlation value of a variable with itself only at one particular time lag, is run at multiple lags and plotted in a correlogram, the significant time lags used in the Moving Average (MA) portion of the ARIMA stochastic models are clearly shown. If the correlogram of the PACF shows exponential decay with no specific lags breaking this trend, then there are no significant lags suggested for use in a MA model. Since the process of finding the true partial autocorrelation value involves taking a partial derivative of the entire population, an estimate of the PACF for discrete samples has been developed from the best approximation, through mean square error, of the sample using an AR model $\hat{x}$ (Dürre & Liboschik, 2015). The function for estimating the partial autocorrelation value $\hat{\pi}$ of a variable *x* with *n* observations at a time lag *L* is shown in Equation 3-C.

$$\hat{\pi}(L) = 2 * \frac{\sum_{i=L+1}^{n}((x_i - \hat{x}_i) * (x_{i-L} - \hat{x}_{i-L}))}{\sum_{i=L+1}^{n}((x_i - \hat{x}_i)^2 + (x_{i-L} - \hat{x}_{i-L})^2)} \tag{3-C}$$

Once the correlogram of both the ACF and PACF were plotted and the significant lags of the AR and MA portions of the candidate ARIMA models were determined, each model was performed on the training period data. The best candidate model was then selected for each state by a method based on similar concepts to Akaike's Information Criterion (AIC) (Akaike, 1973; Ozaki, 1977). The approximate AIC of each model with *n* observations and residuals *E* was quantified by Equation 3-D. The candidate model with

the lowest approximate AIC was then selected and applied to the testing period of the PDSI data. The accuracy of each model to the testing period was also calculated.

$$AIC_{approx} = n * \ln\left(\sigma^2(E)\right) + 2 * (\#\ of\ parameters) \quad \text{(3-D)}$$

3.2.2 Severity Area Frequency (SAF) Curves

Unlike the stochastic modeling, the spatio-temporal SAF curve analysis used the entire climate region defined by NOAA instead of using a single state to represent the region. The analysis was performed using both PDSI values and 1-month aggregated Standardized Precipitation Index (SPI) data. The gridded PDSI data used was provided by the Dai Dataset from the National Center for Atmospheric Research (NCAR) (Dai & Qian, 2004). The dataset gave monthly time series values from 1850-2010 for each grid point on a 2.5° x 2.5° grid covering all of the global land coverage. The data points located outside of the United States were ignored for this analysis.

The SPI data was found using the International Research Institute for Climate and Society (IRI) online database, organized by Columbia University, and provided by NCDC through NOAA ("NOAA NCDC CIRS NClimDiv v1 sp01: 1-Month Standardized Precipitation Index Data," 2017). This dataset gave index values from 1895-2017 at station locations throughout the United States. Since the SPI data is not equally spaced like the PDSI values, it was necessary to develop a method to determine area weight. In order to determine the region that each station location affected, an Inverse Distance Weight (IDW) analysis was performed on the station locations

throughout the US using the climate region borders as break lines (Lu & Wong, 2008). Since the data points were discrete and not in a continuous raster format, there was no need for interpolation between points when using the IDW analysis so the search radius for all areas was set to only be the nearest point. This means that when considering any given point within the entire area of study, the distance to each station was calculated and inversed so that the station with the lowest distance would hold the most or in this study all of the weight of influence over that point. This split all of the regions into separated areas where the SPI values at the stations applied homogenously to the surrounding area. A map of the regions with the station points is shown in Figure 3-2.

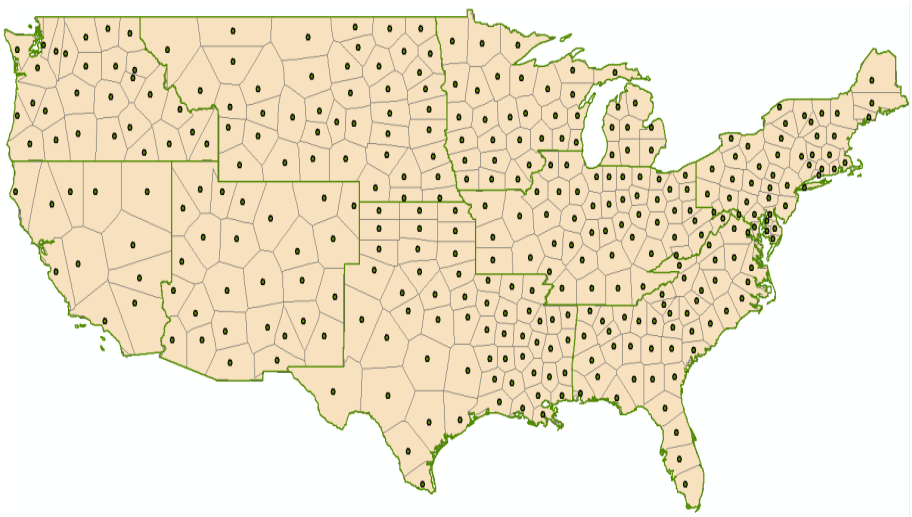

Figure 3-2: SPI Data Stations and Associated Regions of Influence

After each drought index dataset was obtained and processed, a gamma distribution was fit to the time series of both drought indices at each data point. The gamma distribution, whose Probability Density Function (PDF) is shown in Equation 3-E, is based on the mean μ and variance σ^2 of the data (Thom, 1958). The shape α and

scale β parameters for the distribution are calculated by Equations 3-F – 3-G. Once these parameters were calculated for each data point, the 5, 10, 30, 50, and 100 year frequency drought severities were calculated for each drought index time series.

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad \text{(3-E)}$$

$$\mu = \alpha\beta \quad \text{(3-F)}$$

$$\sigma^2 = \alpha\beta^2 \quad \text{(3-G)}$$

With these drought severity frequency values associated to each data point throughout the area of the United States, all of the data points were separated by their associated climatic region. Once separated, the data points in each region were sorted from in descending order by the magnitude of their drought severity at each of the calculated frequency values. Once the severities were sorted at each frequency in each region, the cumulative area percentage that they represented was calculated to the scale of the highest severity representing the frequency value for the entire area (100%) and the lowest severity representing the frequency value for none of the area (0%). Since the PDSI data was equally spaced throughout the regions, the cumulative area percentage was calculated based on the rank *k* of the severity magnitude out of *n* observations shown in Equation 3-H. The SPI drought severities at the specified frequencies had associated area values a_i so the cumulative area percentage was scaled using these areas as well as the rank *k* and number of region observations *n*, shown in Equation 3-I. These cumulative area percentage values of both of the drought indices were plotted against their respective drought severities to make the SAF curves necessary for the analysis.

$$A_{PDSI} = 1 - \frac{k-1}{n-1} \quad \text{(3-H)}$$

$$A_{SPI} = 1 - \frac{a_n / \sum a_i}{n-1} * (k-1) - \frac{\sum_{i=1}^{k-1} a_i}{\sum a_i} \quad \text{(3-I)}$$

3.3 Regional Applicability of Stochastic Models on PDSI

In order to determine the appropriate candidate models for the PDSI forecasting, the correlograms for the ACF and PACF of the training period of PDSI time series were plotted. While only the ACF and PACF correlograms of the Southeast region are shown in Figures 3-3 – 3-4 respectively, the correlograms of all the regions were extremely similar. The PACF correlogram for all of the regions showed exponential decay with no significant drop to indicate possible MA applicability. The possible candidate models are shown in Table 3-i.

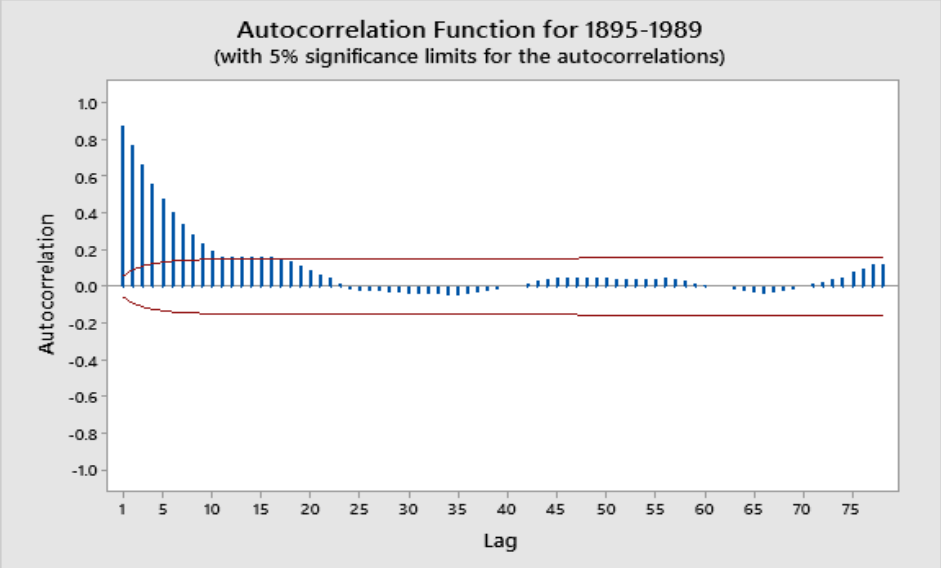

Figure 3-3: ACF Correlogram for Southeast region (SC) PDSI from 1895-1989

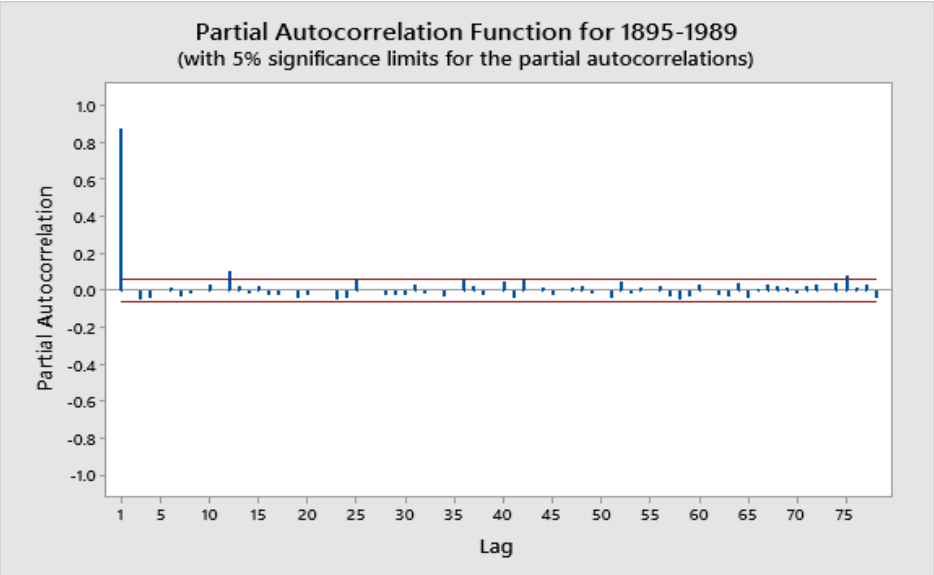

Figure 3-4: PACF Correlogram for Southeast region (SC)
PDSI from 1895-1989

Table 3-i: Candidate Models for All Regions

Region	Candidate Models
Southeast (SC)	AR(1)
South (TX)	AR(1)
	AR(2)
Southwest (CO)	AR(1)
West (CA)	AR(1)
Northwest (WA)	AR(1)
Northern Rockies & Plains (SD)	AR(1)
Upper Midwest (WI)	AR(1)
Ohio Valley (IL)	AR(1)
	AR(2)
Northeast (NY)	AR(1)

With the possible candidate models calculated through the correlograms of the training periods of PDSI data, the best possible model for each region was chosen. The method for choosing the appropriate model for each region followed the minimum AIC method (mAIC) also developed by Akaike in 1973 stating that the lowest approximate

AIC calculated when applying each of the models to the testing period gave the most appropriate model. As shown in Table 3-ii, most of the states followed an AR(1) model with only the South region (TX) following an AR(2) model. The chosen best models for each of the regions are shown in Figures 3-5 – 3-13.

Table 3-ii: Best Fit Models for All Regions

Region	Model	AIC	Precipitation	Temperature
Southeast (SC)	AR(1)	451.39	3.98	62.59
South (TX)	AR(2)	568.24	2.27	64.81
Southwest (CO)	AR(1)	377.32	1.50	44.84
West (CA)	AR(1)	442.34	1.85	57.63
Northwest (WA)	AR(1)	431.56	3.52	46.28
Northern Rockies & Plains (SD)	AR(1)	369.51	1.61	44.68
Upper Midwest (WI)	AR(1)	295.73	2.64	42.61
Ohio Valley (IL)	AR(1)	370.10	3.17	51.78
Northeast (NY)	AR(1)	401.51	3.41	44.77

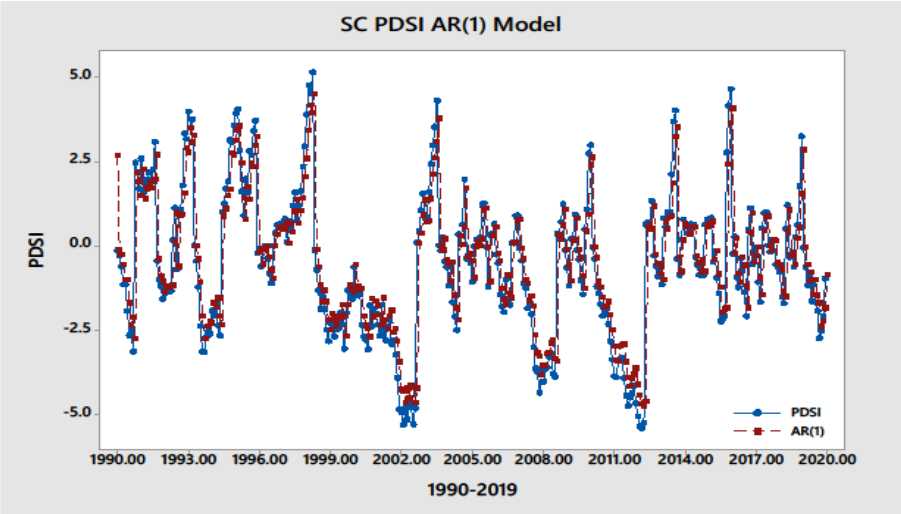

Figure 3-5: Southeast Region (SC) AR(1) PDSI Model In Testing Period of 1990-2019

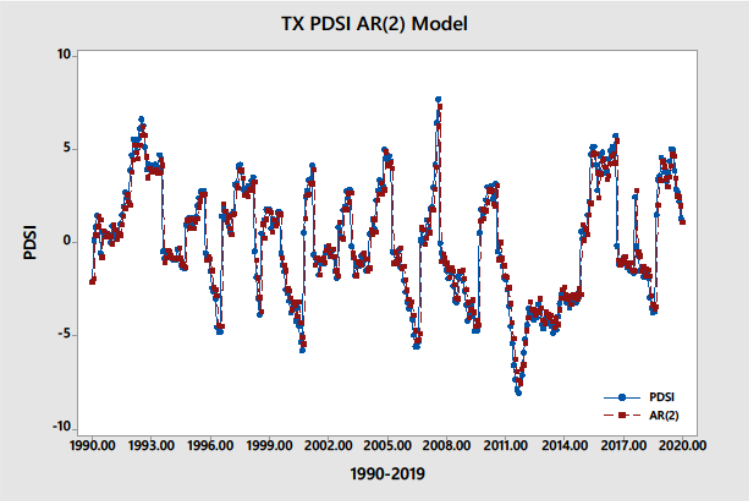

Figure 3-6: South Region (TX) AR(2) PDSI Model In Testing Period of 1990-2019

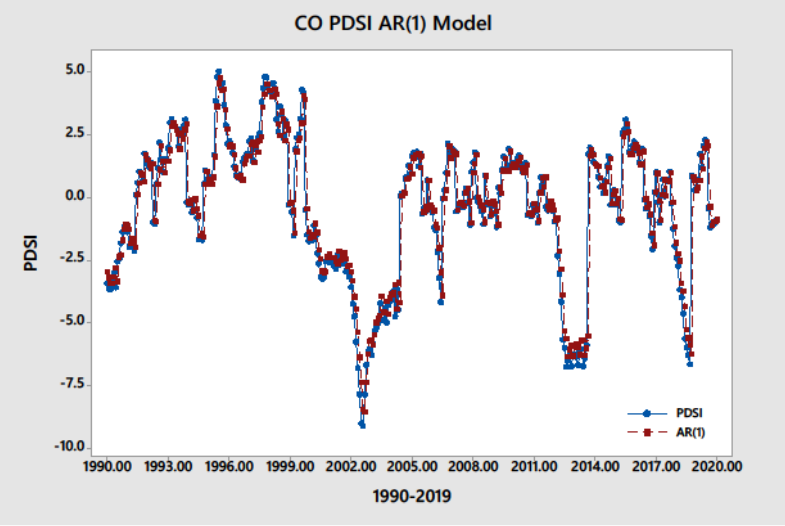

Figure 3-7: Southwest Region (CO) AR(1) PDSI Model In Testing Period of 1990-2019

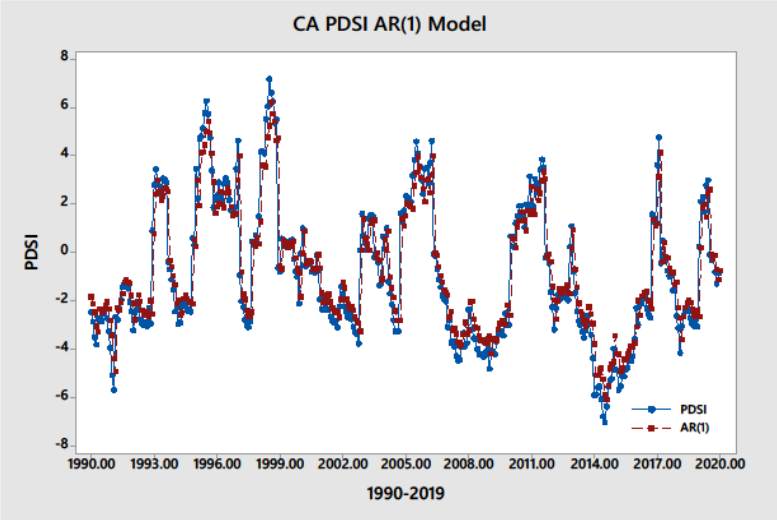

Figure 3-8: West Region (CA) AR(1) PDSI Model In Testing Period of 1990-2019

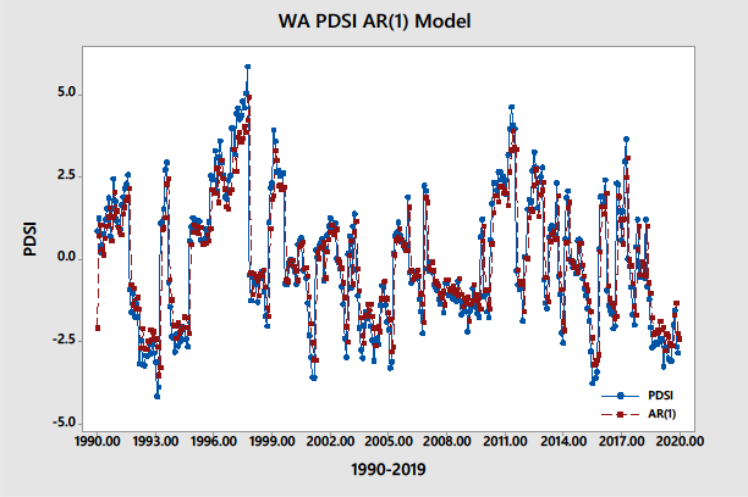

Figure 3-9: Northwest Region (WA) AR(1) PDSI Model In Testing Period of 1990-2019

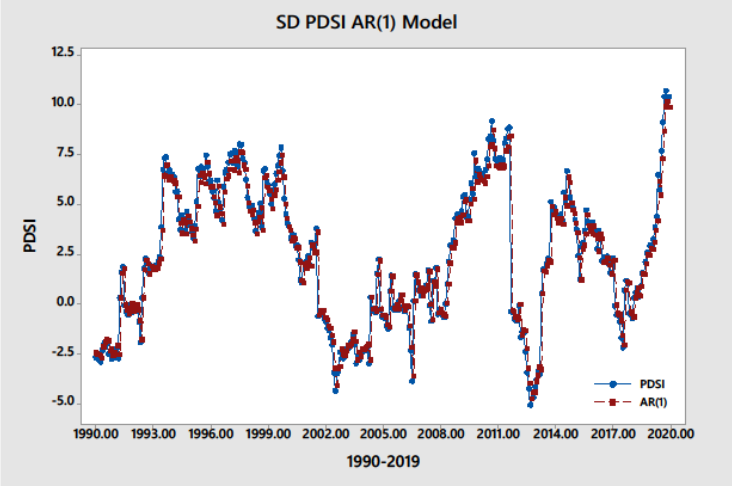

Figure 3-10: Northern Rockies & Plains Region (SD) AR(1) PDSI Model In Testing Period of 1990-2019

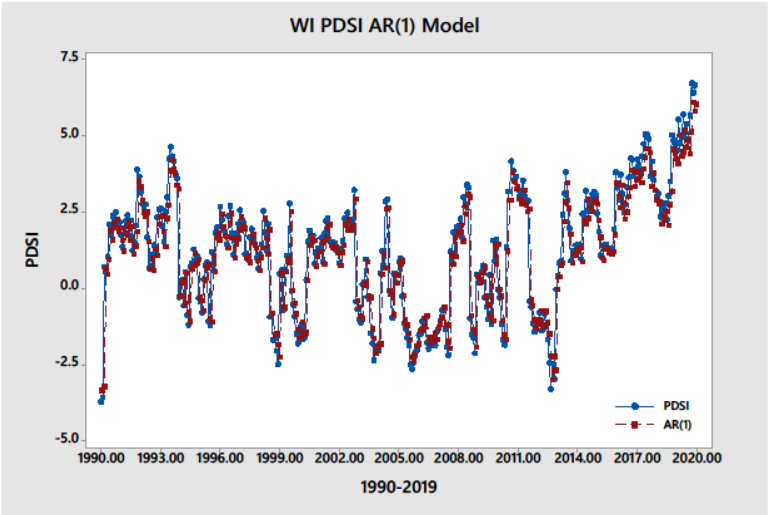

Figure 3-11: Upper Midwest Region (WI) AR(1) PDSI Model In Testing Period of 1990-2019

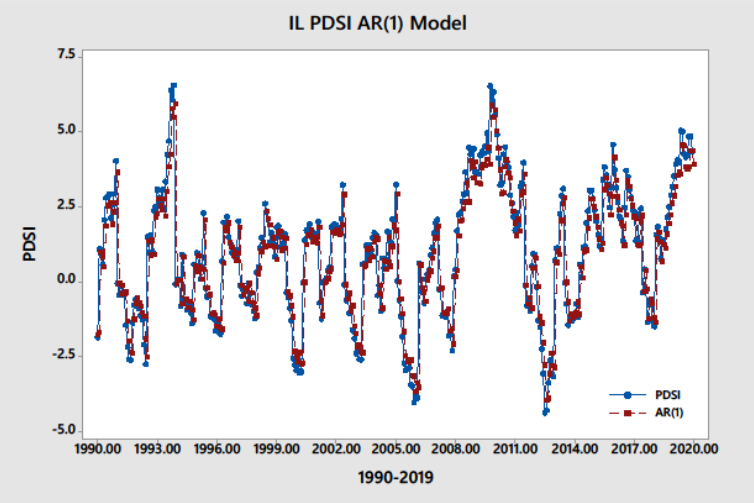

Figure 3-12: Ohio Valley Region (IL) AR(1) PDSI Model In Testing Period of 1990-2019

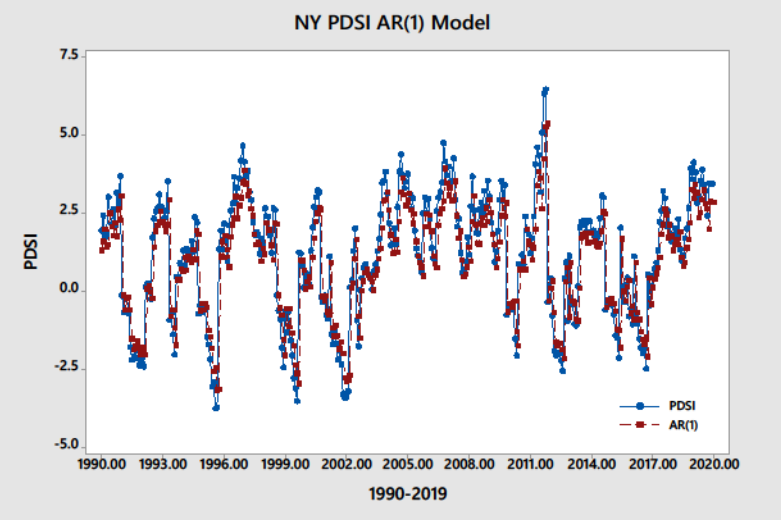

Figure 3-13: Northeast Region (NY) AR(1) PDSI Model In Testing Period of 1990-2019

With the same methods used for each region, the approximate AIC derived for determining the models used for each region could be used to compare suitability of applying stochastic models for forecasting between each of the regions. According to this accuracy comparison, the Upper Midwest region represented by Wisconsin had the lowest approximate AIC of 295.73 showing that stochastic modeling was most applicable in this region. Contrastingly, the region with the highest approximate AIC representing low stochastic model suitability was the South region represented by Texas with a score of 568.24. Since PDSI is mainly based on precipitation and temperature in the area, the average monthly precipitation and temperature of the different regions is also included in Table 3-ii in order to determine which of these factors, if any, have an effect on the accuracy of stochastic modeling for forecasting PDSI (Jacobi et al., 2013). A visual summary was made for these statistics in Figures 3-14 – 3-16 While Figures 3-14 – 3-15 compare the accuracy to precipitation and temperature directly, Figure 3-16 compares accuracy to both temperature and precipitation by considering how the region ranks among all regions in the specified category. All of these figures support the idea that temperature affects stochastic modeling applicability more than precipitation. This claim would need to be furthered researched in future studies.

In addition to checking the applicability of stochastic models for each region through an approximate AIC, an introductory study was performed on the stationarity assumption necessary to apply stochastic modeling. This was done by splitting the entire PDSI time series into two sixty year halves from 1900-1959 and 1960-2019. The maximum drought severity, intensity, and percentage of period in drought is shown in

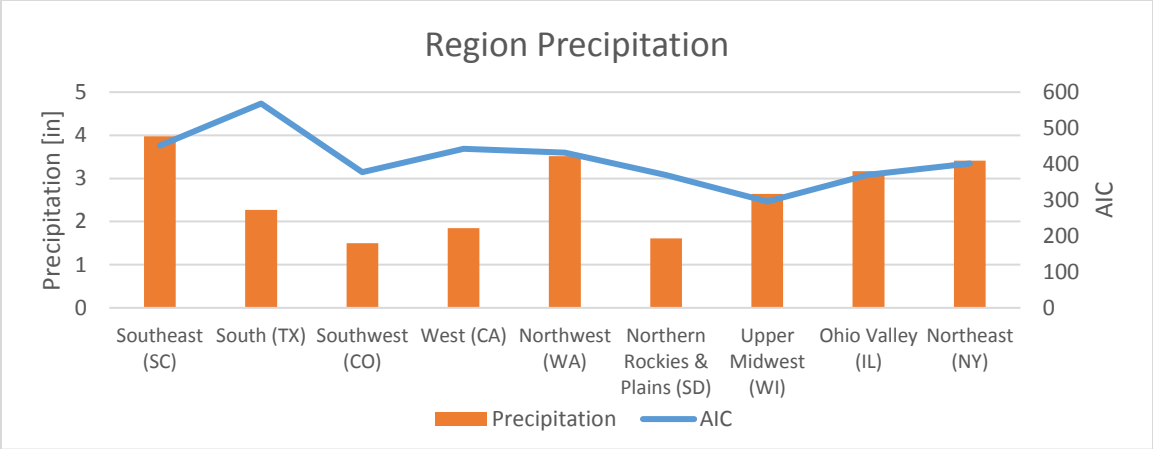

Figure 3-14: Regional Average Monthly Precipitation and Stochastic Model Accuracy

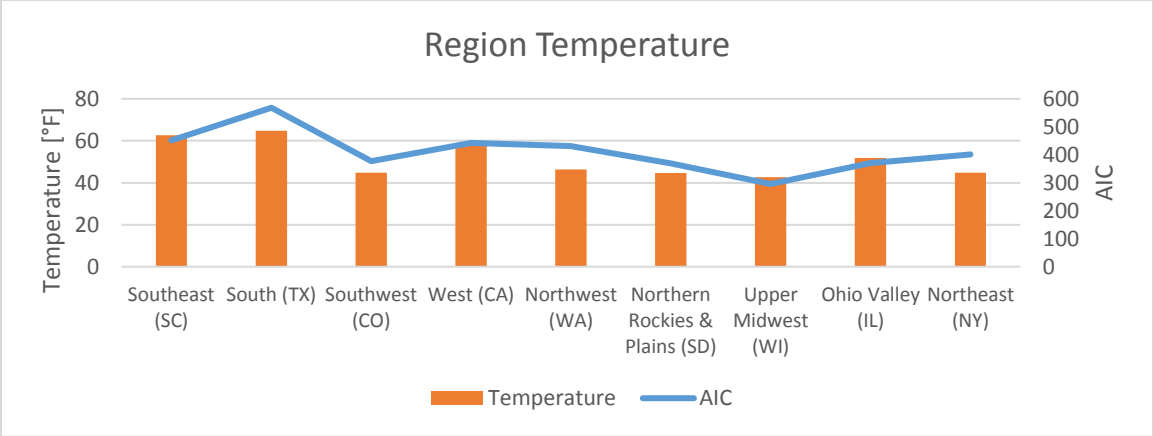

Figure 3-15: Regional Average Monthly Temperature and Stochastic Model Accuracy

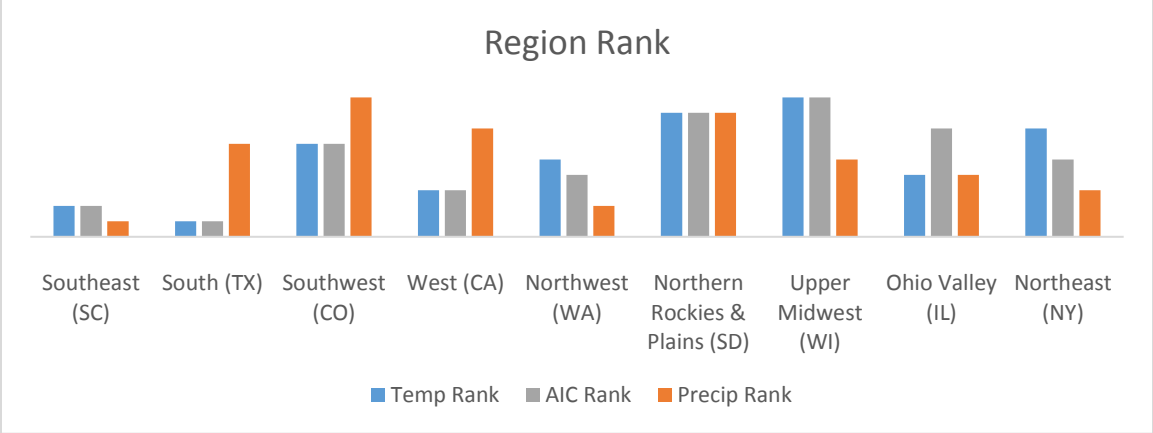

Figure 3-16: Regional Average Monthly Precipitation/Temperature and Stochastic Model Accuracy

Table 3-iii. These statistics are also represented visually for severity, intensity, and drought percentage in Figures 3-17 – 3-19. A drought event in the Northern Rockies & Plains region (SD) during the first half of data gave it a large difference from the first half of data to the second in both drought severity and intensity but the drought percentage, which is less effected by individual drought events, for this region was relatively equal among the two time periods. In the drought percentage statistic the largest difference shown were in the South and Upper Midwest regions, represented by Texas and Wisconsin respectively, with both areas having about 13% more drought months in the first sixty years compared to the second. The region that seemed to follow the stationarity assumption the closest would be the Northwest region shown through Washington as the values for maximum drought intensity and severity as well as percentage of months that were considered in drought were relatively equal among both time periods.

Table 3-iii: Comparing Maximum Drought Severity and Intensity and Percentage of Drought between 1900-1959 & 1960-2019

Region	1900-1959			1960-2019		
	Max. Severity	Max. Intensity	Drought Percentage	Max. Severity	Max. Intensity	Drought Percentage
Southeast (SC)	199.57	5.46	58.47	136.78	5.41	49.58
South (TX)	341.62	7.77	65.00	199.52	8.06	52.08
Southwest (CO)	196.70	6.43	48.19	221.17	9.09	47.78
West (CA)	85.59	6.03	51.67	181.47	7.01	61.67
Northwest (WA)	125.67	5.10	51.67	111.60	4.78	52.64
Northern Rockies & Plains (SD)	423.33	8.31	46.53	127.04	5.10	39.17
Upper Midwest (WI)	193.34	6.41	50.28	75.73	7.88	37.64
Ohio Valley (IL)	155.77	7.05	52.64	92.20	5.17	40.97
Northeast (NY)	101.56	4.76	54.31	204.31	5.89	42.22

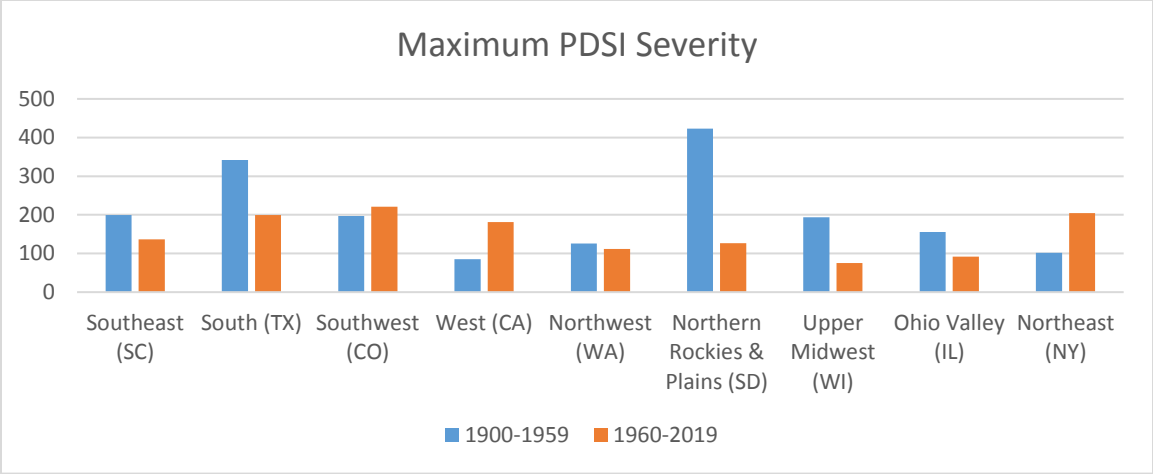

Figure 3-17: Maximum Drought Severity between 1900-1959 & 1960-2019

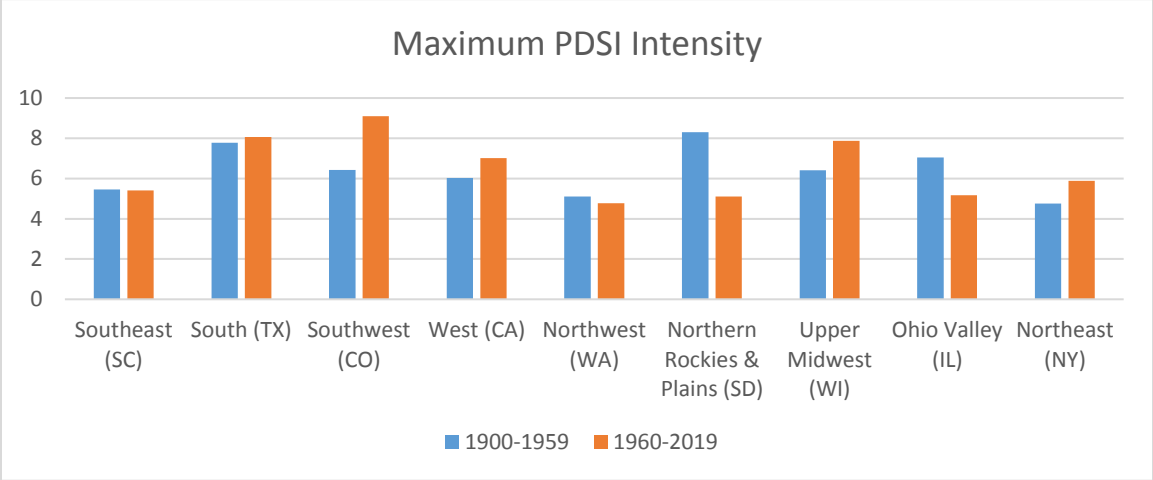

Figure 3-18: Maximum Drought Intensity between 1900-1959 & 1960-2019

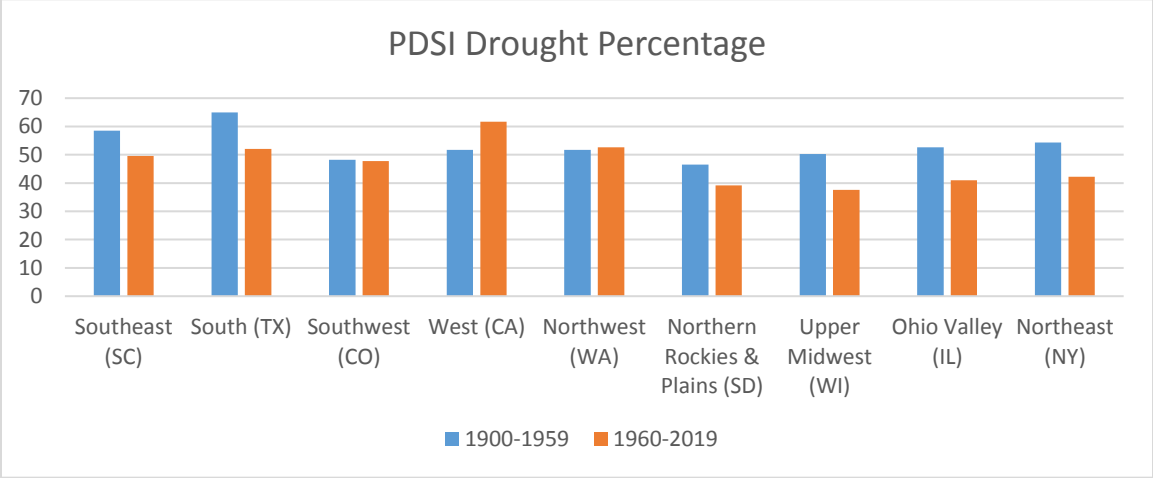

Figure 3-19: Maximum Percentage of Drought between 1900-1959 & 1960-2019

3.4 Spatio-Temporal Regional Analysis

After calculating the specified drought severity frequencies for each point and determining the associated area values, the severity and area values were plotted together for each frequency to form the SAF curves for each region. The PDSI SAF curves are shown in Figures 3-20 – 3-28. For areas such as the Southeast, Ohio Valley, and Northeast regions in Figures 3-20, 3-27, and 3-28 the flatter curves show a homogenous region where drought severity applies equally across the region. For areas with SAF curves that have a large jump in severity at low cumulative area values such as the Southwest, West, Northern Rockies & Plains, and Upper Midwest in Figures 3-22 – 3-23 and 3-25 – 3-26 show that there are specific points within the region that are particularly prone to drought. The curves shown in the Northwest region in Figure 3-24 have two clearly separate homogenous regions. This may show that this region would need to be further divided into sub regions when analyzing PDSI based drought frequency. The curves shown in Figure 3-21 for the South region show a constant linear trend in all of the curves which means that the areas where droughts are particularly severe diffuse throughout the region to create a continuity of drought severities.

In order to compare each region, the 50 year and 100 year SA curves for all regions were combined in Figures 3-29 and 3-30 respectively. Comparing these regions shows that the eastern regions, specifically the Southeast, Northeast, Ohio Valley and South regions, are especially comparable and could be merged in future analyses. While not as comparable, similar claims could be made for the Upper Midwest and Northern Rockies & Plains regions.

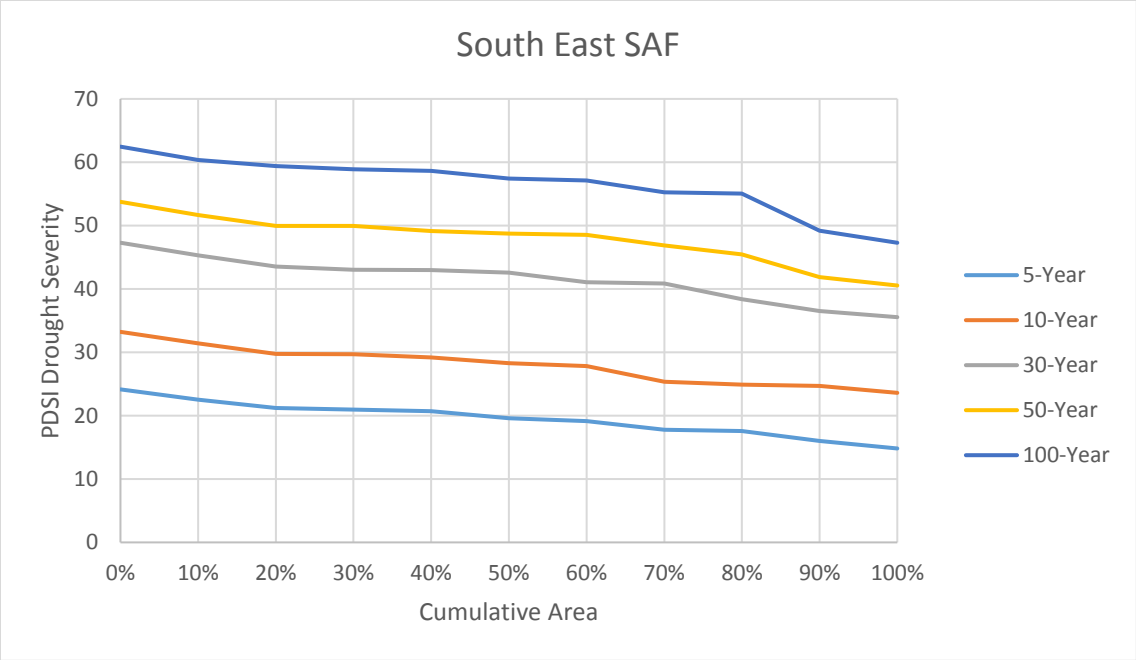

Figure 3-20: Southeast PDSI SAF Curves

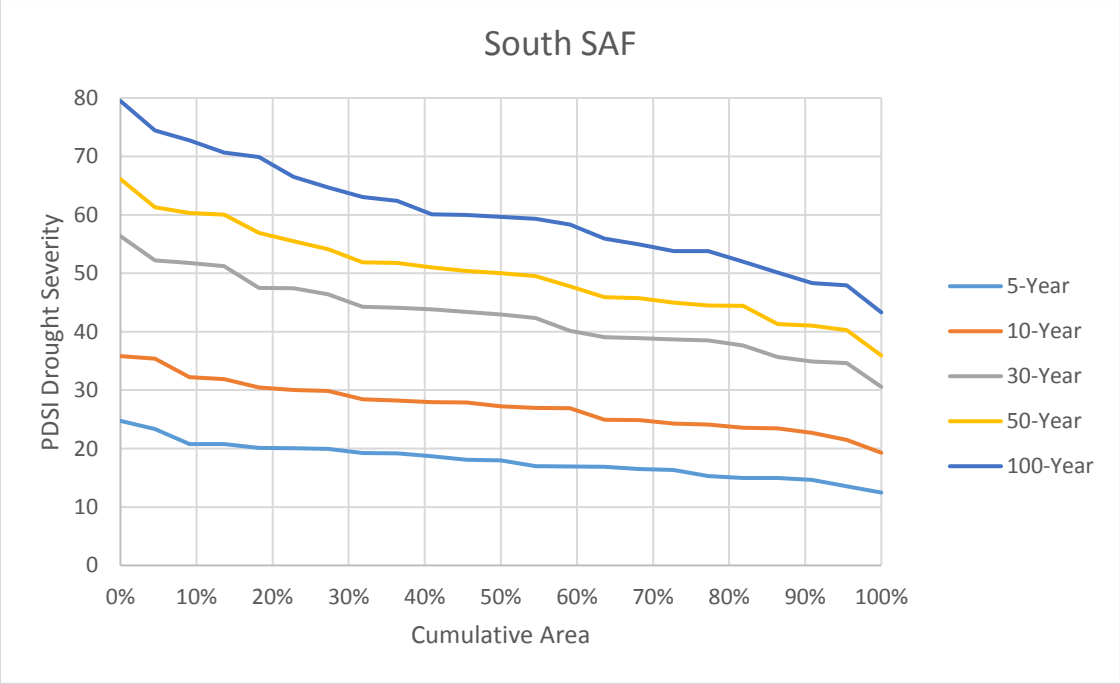

Figure 3-21: South PDSI SAF Curves

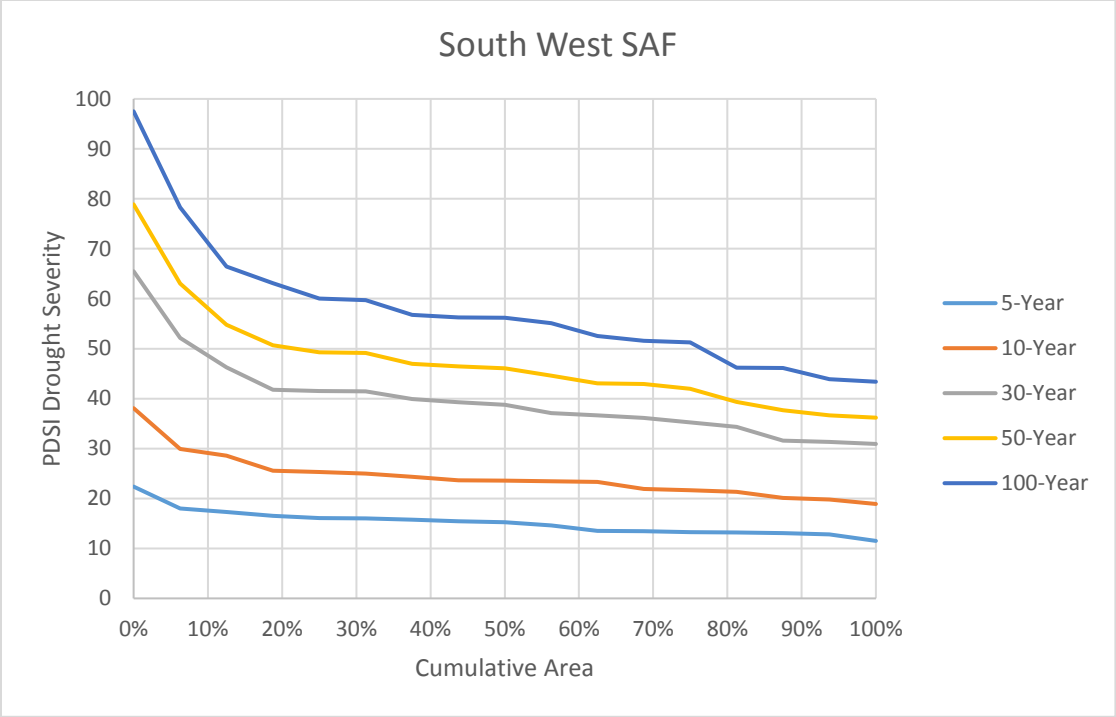

Figure 3-22: Southwest PDSI SAF Curves

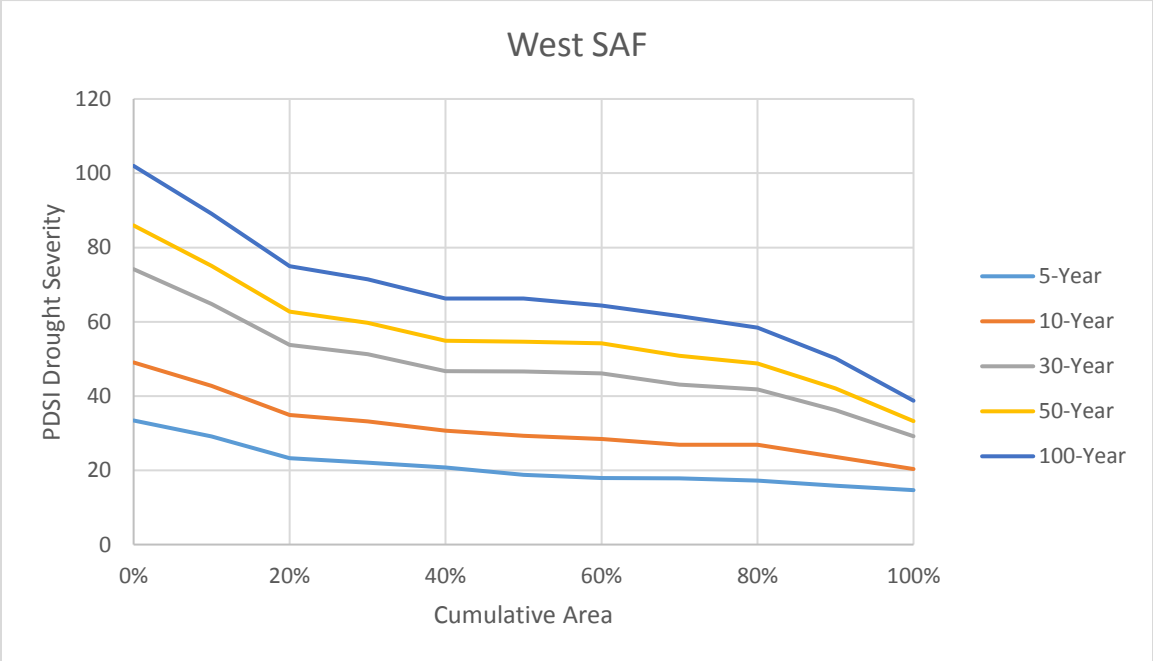

Figure 3-23: West PDSI SAF Curves

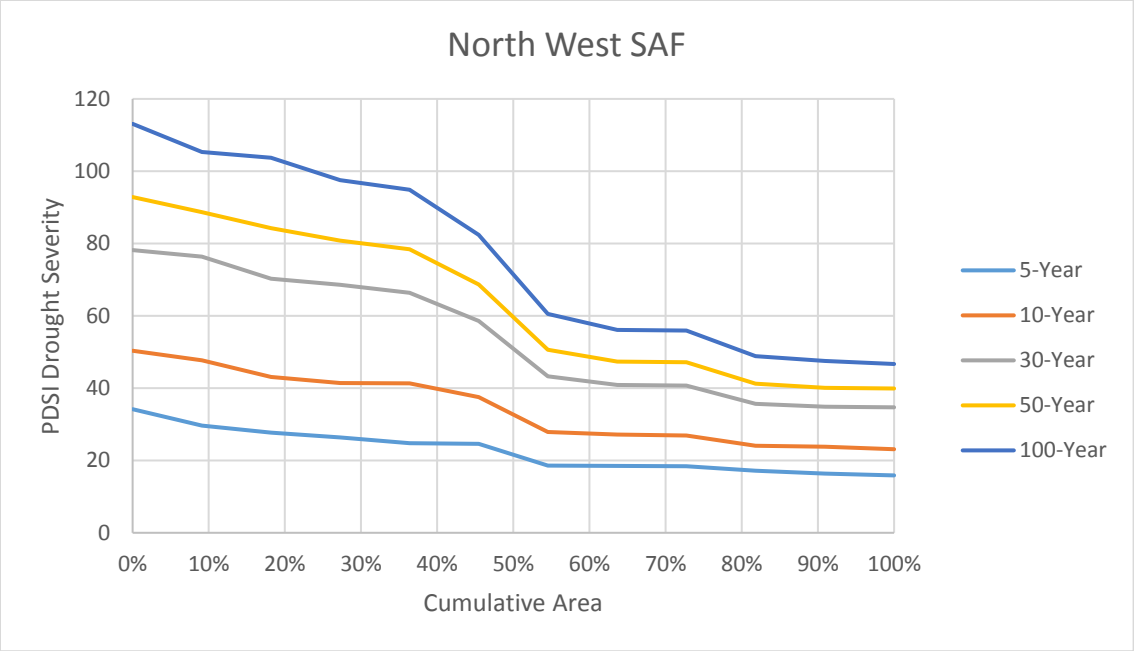

Figure 3-24: Northwest PDSI SAF Curves

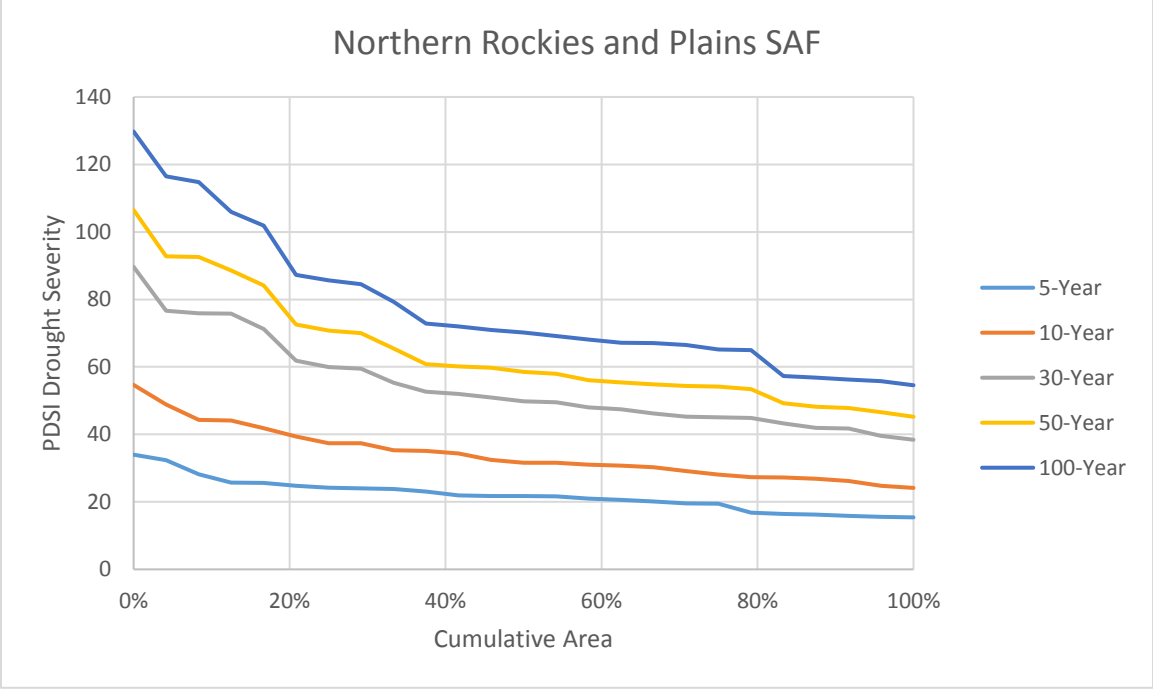

Figure 3-25: Northern Rockies & Plains PDSI SAF Curves

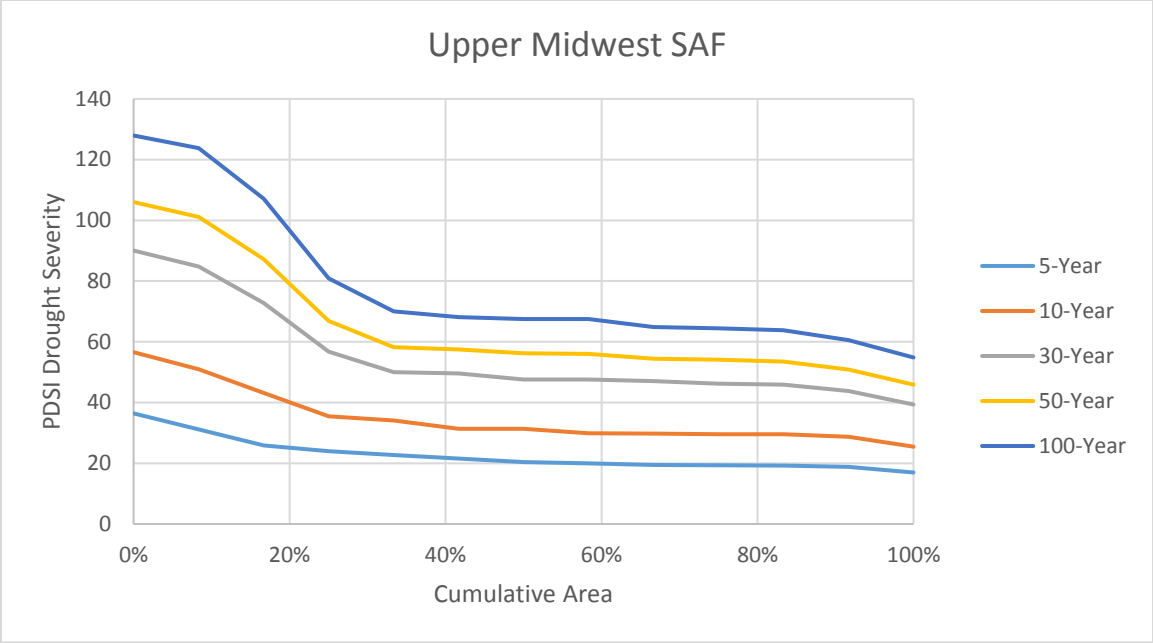

Figure 3-26: Upper Midwest PDSI SAF Curves

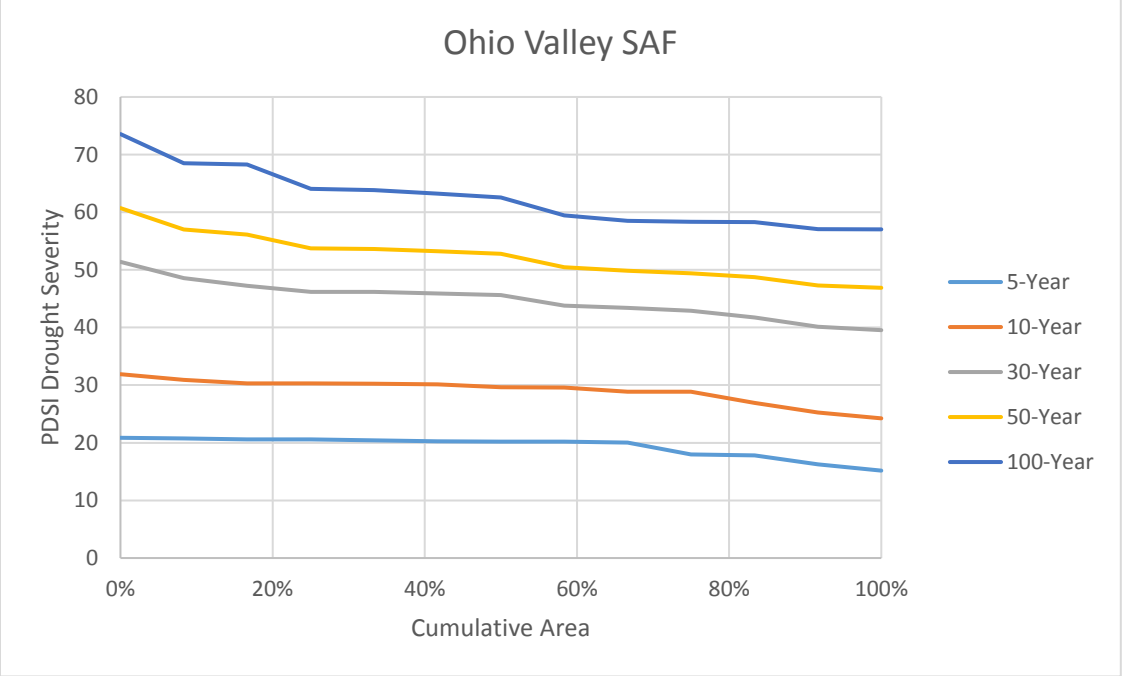

Figure 3-27: Ohio Valley PDSI SAF Curves

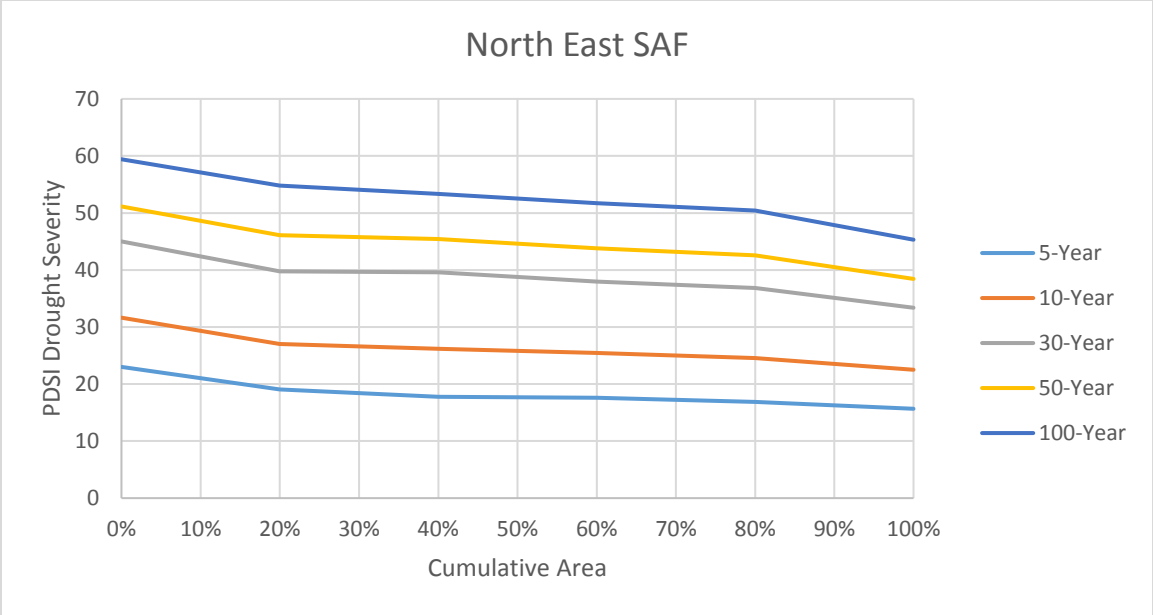

Figure 3-28: Northeast PDSI SAF Curves

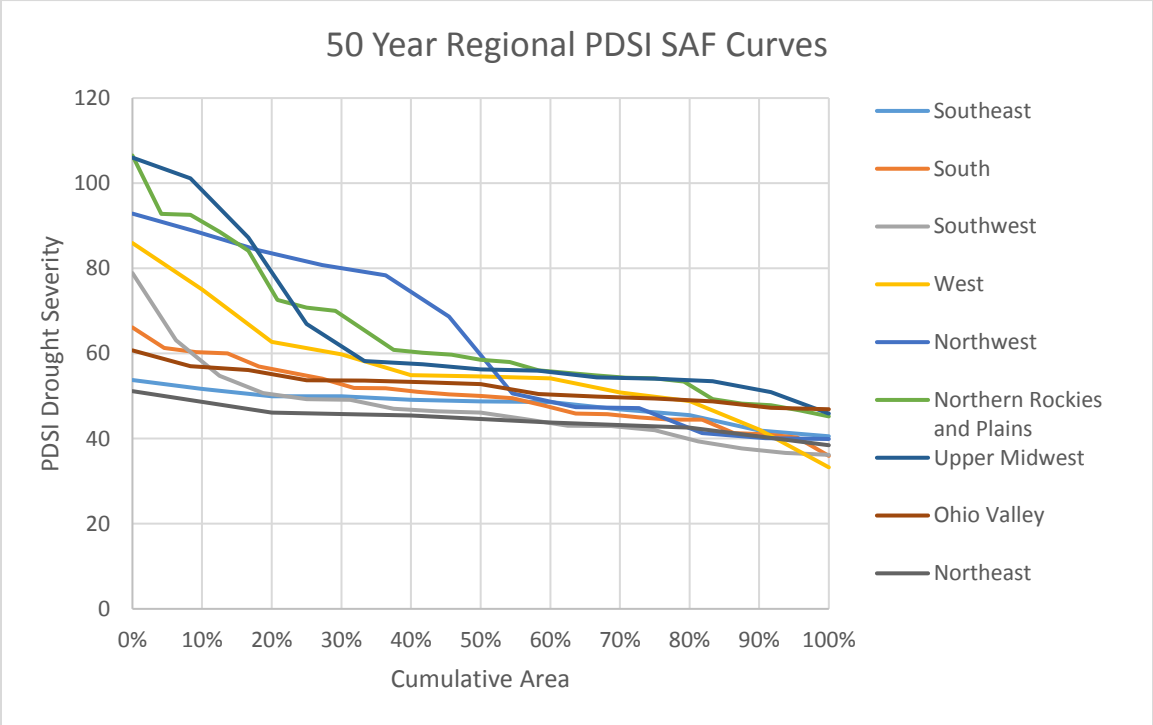

Figure 3-29: Climate Regions 50 Year PDSI SA Curves

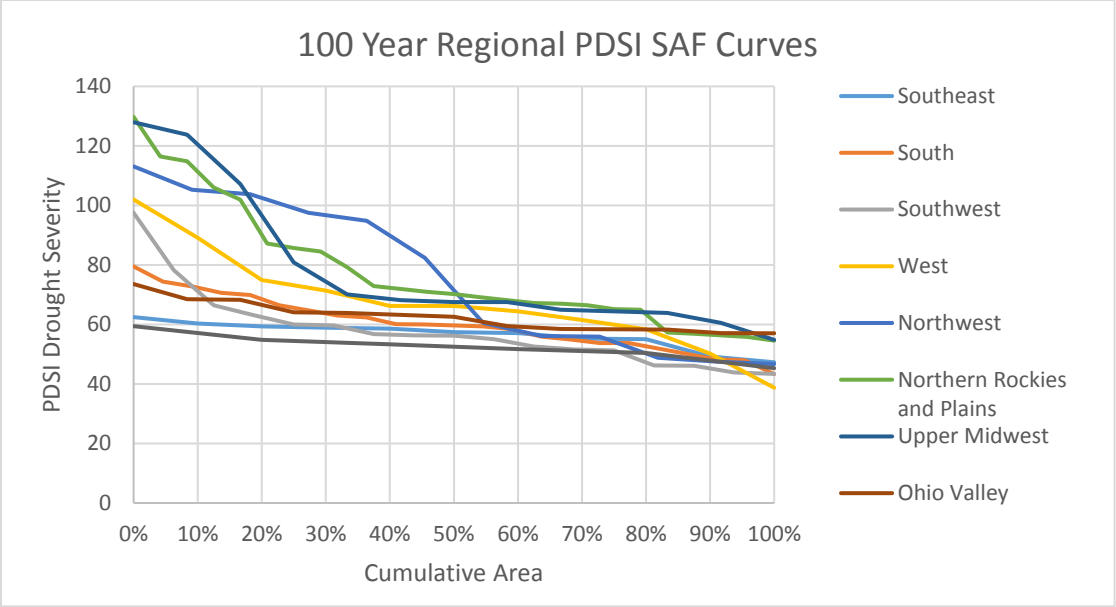

Figure 3-30: Climate Regions 100 Year PDSI SA Curves

Using a similar procedure for the SPI based droughts, the SAF curves for these index values are shown in Figures 3-31 – 3-39. The curves derived from these index severities show much more homogenous regions with slightly linear trends indicating a continuous drought severity affecting the entire region. The lower severities in the upper end of the cumulative area of the Southwest and West regions in Figures 3-33 and 3-34 respectively show that there is a small region that is particularly resilient against droughts in the southwestern area. This is also seen in the comparison graphs of the 50 year and 100 year frequency SA curves in Figures 3-40 and 3-41 as all regions seem to be comparable except for the upper cumulative areas of the West and Southwest regions. The only other exception in this comparison was the lower cumulative area of the South region showing an area in this region that is especially affected by droughts compared to the rest of the nation.

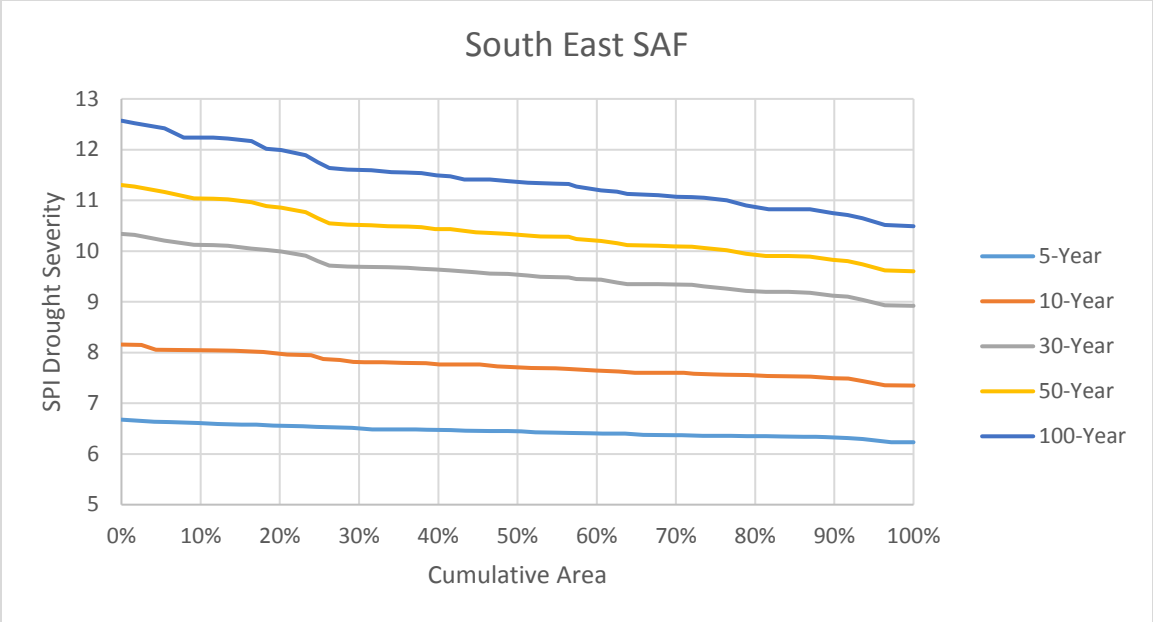

Figure 3-31: Southeast SPI SAF Curves

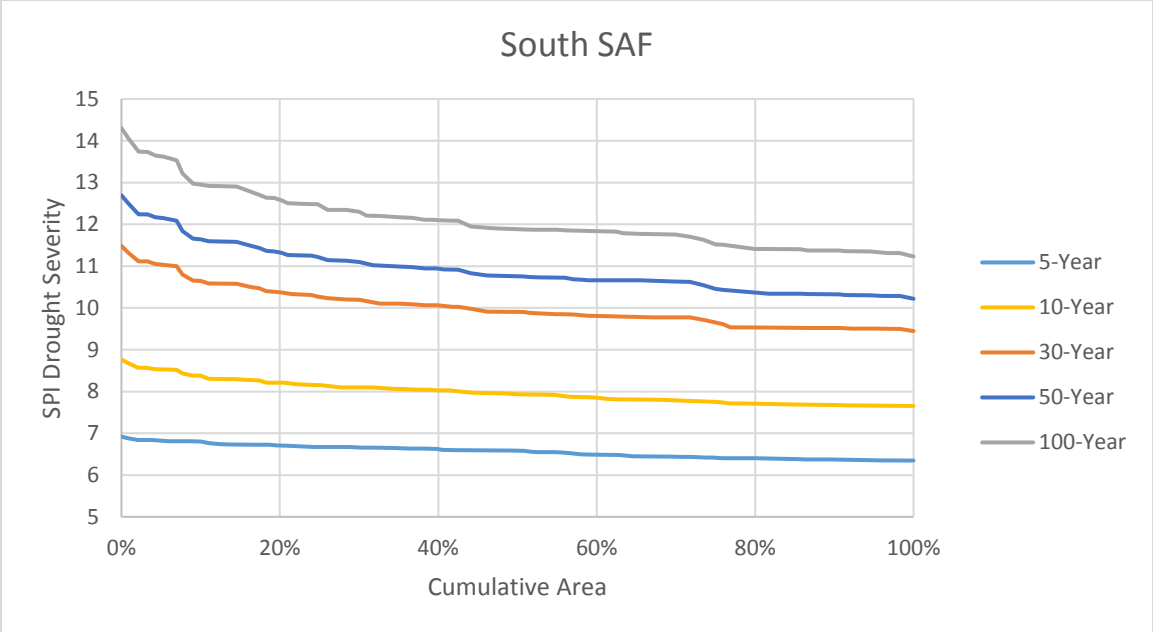

Figure 3-32: South SPI SAF Curves

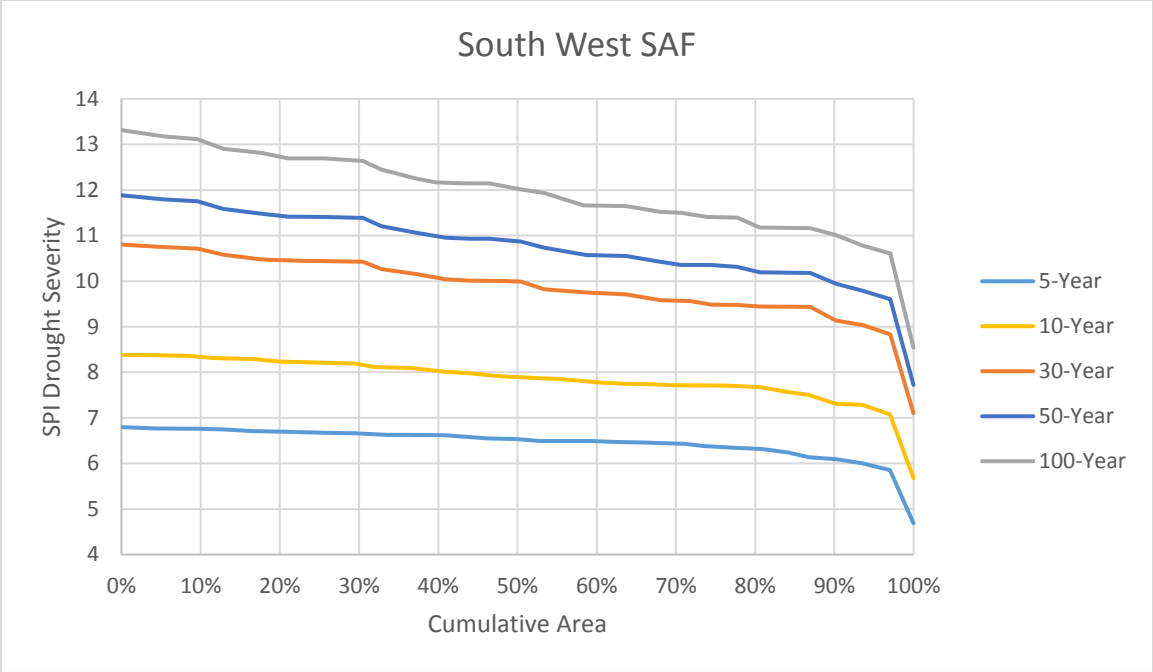

Figure 3-33: Southwest SPI SAF Curves

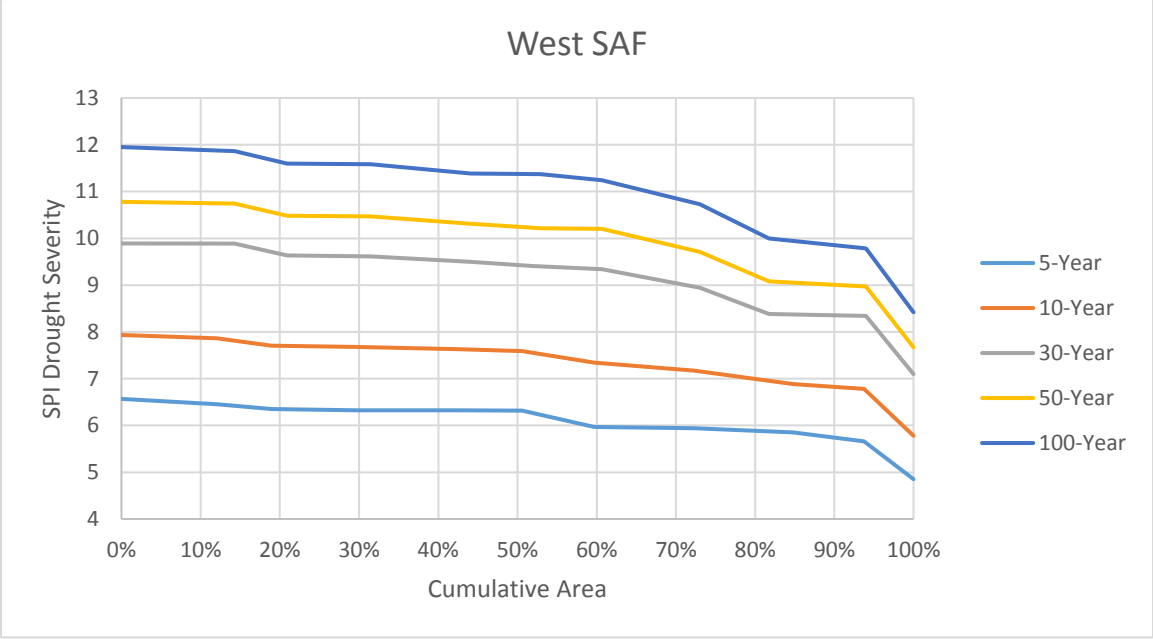

Figure 3-34: West SPI SAF Curves

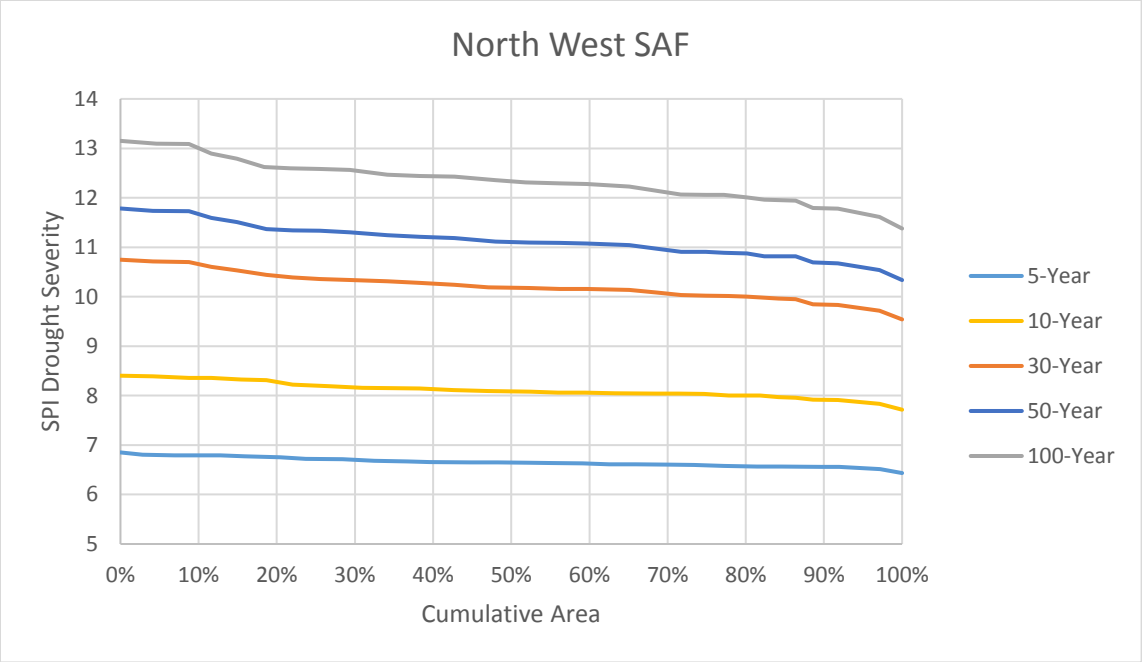

Figure 3-35: Northwest SPI SAF Curves

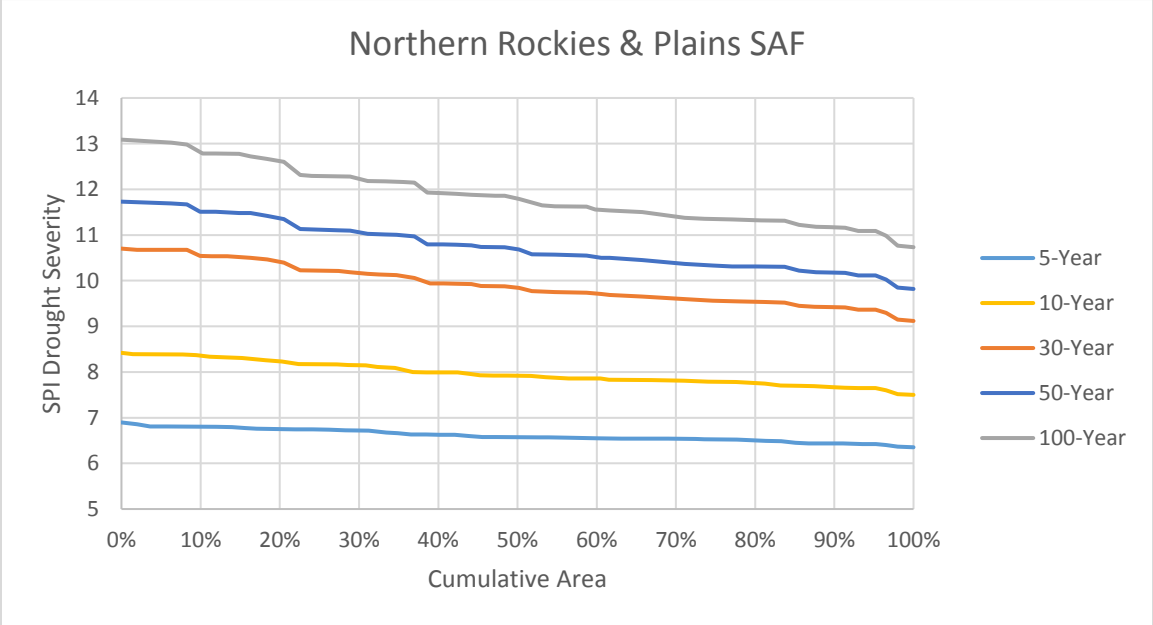

Figure 3-36: Northern Rockies & Plains SPI SAF Curves

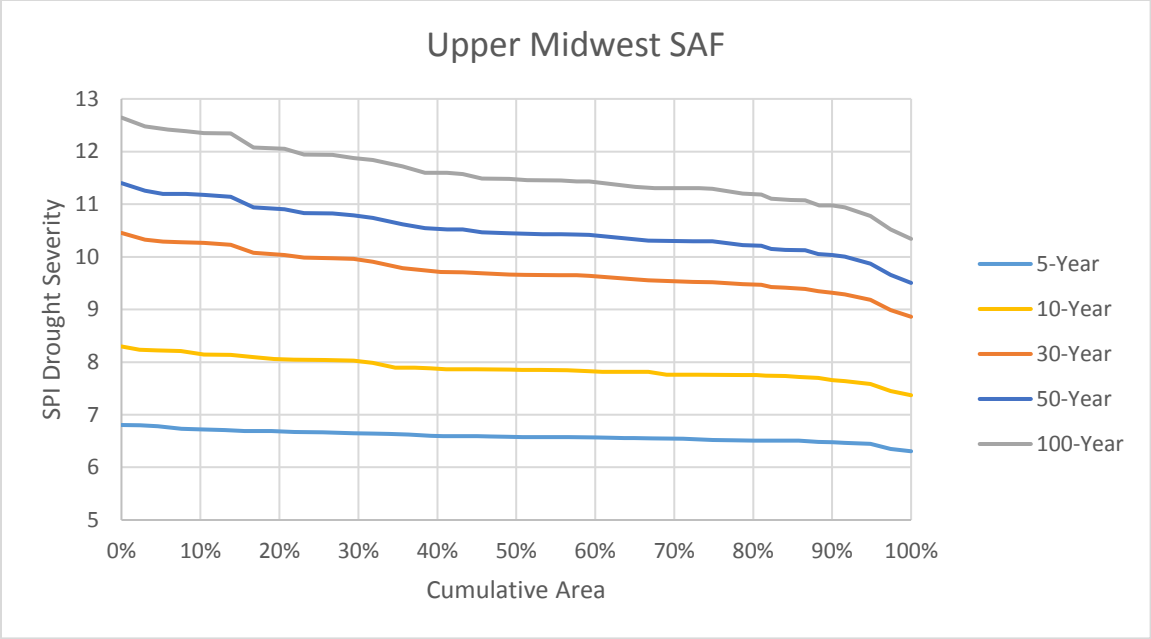

Figure 3-37: Upper Midwest SPI SAF Curves

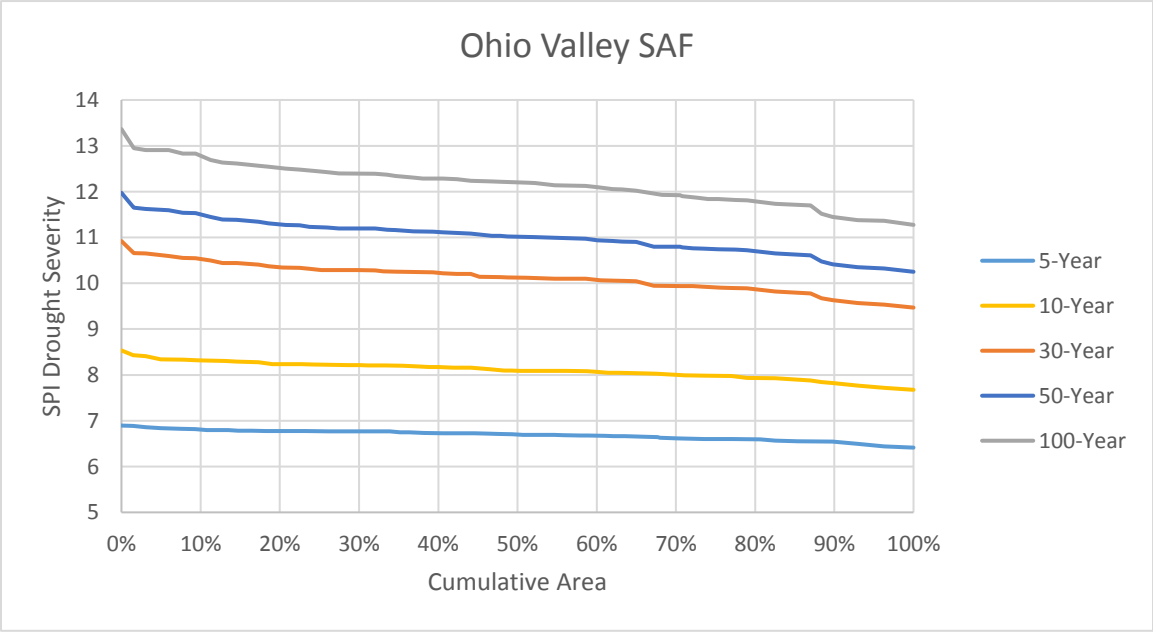

Figure 3-38: Ohio Valley SPI SAF Curves

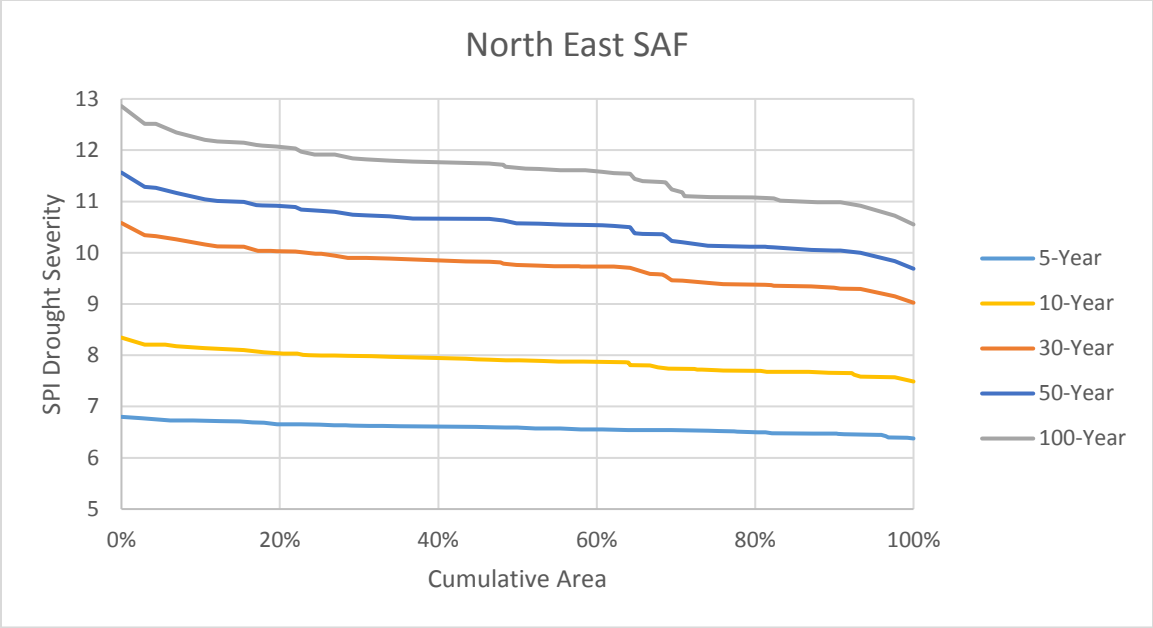

Figure 3-39: Northeast SPI SAF Curves

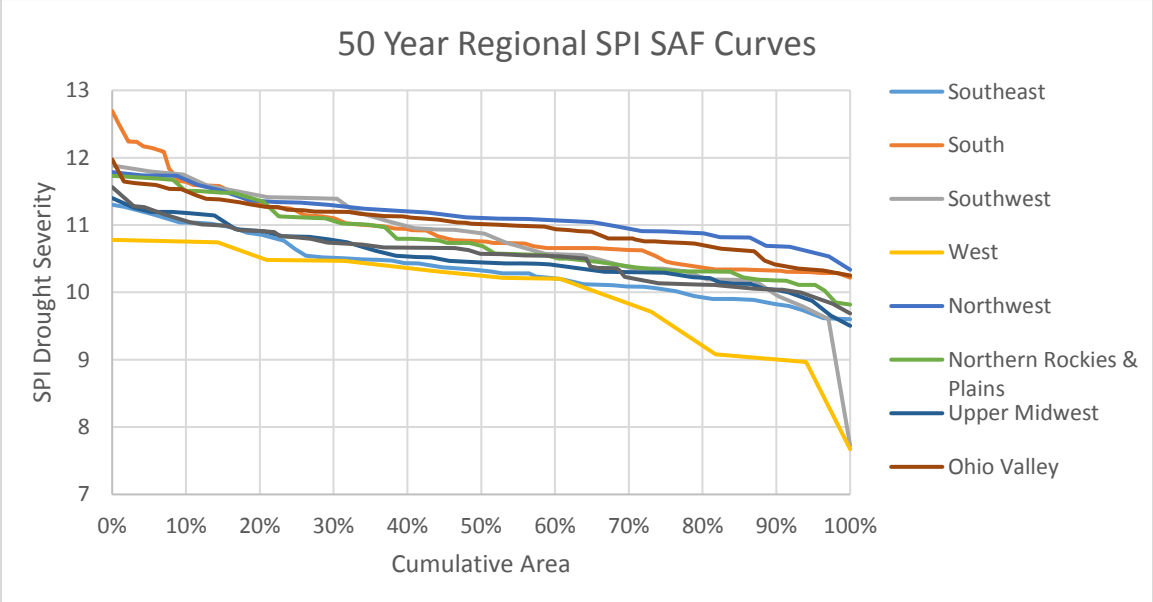

Figure 3-40: Climate Regions 50 Year SPI SA Curves

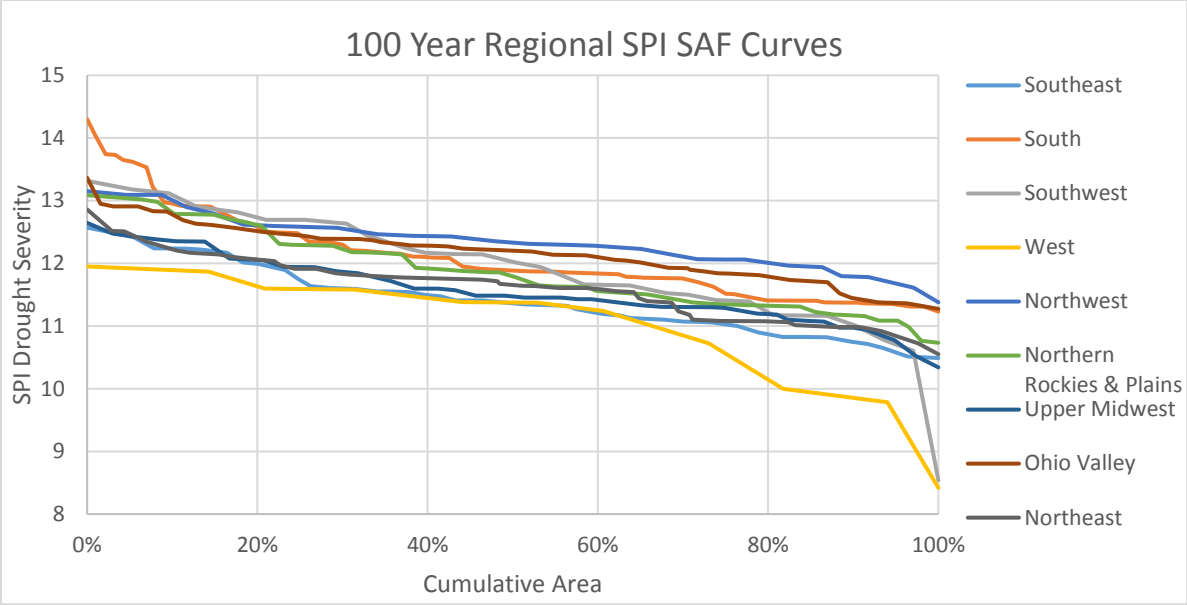

Figure 3-41: Climate Regions 100 Year SPI SA Curves

3.5 Conclusion

Through both ARIMA modeling and SAF curve analyses, some of the climatic regions produced by the NCDC are more applicable to drought study than others. Through the stochastic modeling accuracy, the regions with lower extreme temperatures such as the Northern Rockies & Plains, Upper Midwest, and Ohio Valley regions proved to have more accurate forecasting of PDSI values. The Southwest region also gave a fairly accurate ARIMA model despite commonly having higher temperatures in the region but this could be due to the state of Colorado, which represented the region, experiencing lower temperatures in some regions of the state compared to the rest of the region. The accuracy of drought index forecasting means that these regions can be more prepared for droughts and their expected impacts. The stationarity assumption of stochastic modeling was also tested and while the assumption was not proved false, the

differences of various drought qualities between the first half of PDSI data compared to the second suggests that the assumption should be tested further.

When taking the spatial aspect of droughts into account in the SAF curve analysis, the entirety of the climatic regions can be more properly assessed for drought applicability. Between the two indices tested, PDSI seemed to show more variability than SPI and could indicate a more detailed assessment of the regions. The lack of variability in SPI does indicate that precipitation affects most of the regions homogenously meaning that purely meteorological studies performed using these regions would be appropriate. The regions that lacked variability in the PDSI SAF curves, including the Southeast, Northeast, and Ohio Valley regions, indicate that these regions are appropriate for drought study. Most of the other regions have particular areas where droughts are more likely to occur, the Northwest region had two distinct areas with differing drought severities in each frequency which indicates a necessity for possible sub-regions when performing drought studies in the area.

3.6 Resources

A.K. Mishra PhD Scholar & V.R. Desai (2005) Spatial and temporal drought analysis in the Kansabati river basin, India, International Journal of River Basin Management, 3:1, 31-41, DOI: 10.1080/15715124.2005.9635243

Akaike, Htrotugu. Maximum likelihood identification of Gaussian autoregressive moving average models, Biometrika, Volume 60, Issue 2, August 1973, Pages 255–265, https://doi.org/10.1093/biomet/60.2.255

Amirataee, Babak, et al. "Regional Analysis and Derivation of Copula-Based Drought Severity-Area-Frequency Curve in Lake Urmia Basin, Iran." Journal of Environmental Management, vol. 206, 15 Jan. 2018, pp. 134–144., doi:10.1016/j.jenvman.2017.10.027.

Anderson, R. L. "Distribution of the Serial Correlation Coefficient." The Annals of Mathematical Statistics, vol. 13, no. 1, 1942, pp. 1–13. JSTOR, www.jstor.org/stable/2236157. Accessed 20 Mar. 2020.

Baldwin, Rich. Divisional Data Select, National Oceanic and Atmospheric Association, 26 Mar. 2020, https://www7.ncdc.noaa.gov/CDO/CDODivisionalSelect.jsp#.

Dai, A., K. E. Trenberth, and T. Qian, 2004: A global data set of Palmer Drought Severity Index for 1870-2002: Relationship with soil moisture and effects of surface warming. J. Hydrometeorology, 5, 1117-1130.

Dürre, A., Fried, R. and Liboschik, T. (2015), Robust estimation of (partial) autocorrelation. WIREs Comput Stat, 7: 205-222. doi:10.1002/wics.1351

Enloe, Sanchez-Lugo. "U.S. Climate Regions." National Climatic Data Center, National Oceanic and Atmospheric Administration, 6 Mar. 2020, www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php.

Jacobi, J., Perrone, D., Duncan, L. L., and Hornberger, G. (2013), A tool for calculating the Palmer drought indices, Water Resour. Res., 49, 6086– 6089, doi:10.1002/wrcr.20342.

Katz, Richard W., and Marc B. Parlange. “Overdispersion Phenomenon in Stochastic Modeling of Precipitation.” Journal of Climate, vol. 11, no. 4, Apr. 1998, pp. 591–601., doi:10.1175/1520-0442(1998)011<0591:opismo>2.0.co;2.

Lu, George Y., and David W. Wong. “An Adaptive Inverse-Distance Weighting Spatial Interpolation Technique.” Computers & Geosciences, vol. 34, no. 9, 2008, pp. 1044–1055., doi:10.1016/j.cageo.2007.07.010.

Mishra, A.K., Desai, V.R. Drought forecasting using stochastic models. Stoch. Environ. Res. Ris. Assess. 19, 326–339 (2005). https://doi.org/10.1007/s00477-005-0238-4

Mishra, Ashok K., and Vijay P. Singh. “A Review of Drought Concepts.” Journal of Hydrology, vol. 391, no. 1-2, 9 July 2010, pp. 202–216., doi:10.1016/j.jhydrol.2010.07.012.

Mishra, Ashok K., and Vijay P. Singh. “Drought Modeling – A Review.” Journal of Hydrology, vol. 403, no. 1-2, June 2011, pp. 157–175., doi:10.1016/j.jhydrol.2011.03.049.

Mishra, A. K., and V. P. Singh (2010), Changes in extreme precipitation in Texas, J. Geophys. Res., 115, D14106, doi:10.1029/2009JD013398.

Mishra, Ashok K., et al. “Anatomy of a Local-Scale Drought: Application of Assimilated Remote Sensing Products, Crop Model, and Statistical Methods to an Agricultural Drought Study.” Journal of Hydrology, vol. 526, July 2015, pp. 15–29., doi:10.1016/j.jhydrol.2014.10.038.

“NOAA NCDC CIRS NClimDiv v1 sp01: 1-Month Standardized Precipitation Index Data.” International Research Institute, Columbia University, 10 Apr. 2017, https://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.CIRS/.nClimDiv/.v1/sp01/.

Ozaki, T. (1977), On the Order Determination of Arima Models. Journal of the Royal Statistical Society: Series C (Applied Statistics), 26: 290-301. doi:10.2307/2346970

Robeson, S. M. (2015), Revisiting the recent California drought as an extreme value, Geophys. Res. Lett., 42, 6771– 6779, doi:10.1002/2015GL064593.

Thom, H. C. S. “A Note On The Gamma Distribution.” Monthly Weather Review, vol. 86, no. 4, Apr. 1958, pp. 117–122., doi:10.1175/1520-0493(1958)086<0117:anotgd>2.0.co;2

Thomas R. Karl and Walter James Koss, 1984: "Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895-1983." Historical Climatology Series 4-3, National Climatic Data Center, Asheville, NC, 38 pp.

CHAPTER 4: CONCLUSION

4.1 Summary of Results

The objective of this study was to understand the severity which droughts affect different sectors of society and gain further depth on expectations of drought effects throughout the United States. The second chapter of this study focused on drought correlation with the international economic sector in order to quantify the impact that droughts have on different regions of the world. The initial global correlation boxplots showed the necessity of detrending GDP data to remove the exponential influence of economic trends and gain a more accurate understanding of drought impact on the regional and international economy. It was also shown through each of the international correlation maps that each country or region has a different combination of GDP detrending and drought index that produces the most significant negative correlation. Overall, the combination of Log TFPW GDP detrending and PDSI drought index seemed to give the most complete picture of negative correlation between drought and GDP. Each nation's economists could look at which combination of drought index and detrending method was most productive for their country in order to determine the impact that droughts have on their individual national economy.

Once the importance that droughts have on the economic sector was established, the third chapter focused on the application of the NOAA US national climatic divisions to drought studies to understand if these regions were appropriate for drought summaries. The stochastic ARIMA modeling in this chapter were performed on the PDSI values of

state representatives of each region because of the strong negative correlation that this index was shown to have on the American economy shown in the previous chapter. The forecasting results were compared for accuracy to determine which regions could have more predictability with droughts and plan for future resiliency techniques more reliably. The northern regions of the Upper Midwest, Northern Rockies & Plains, and Ohio Valley were shown to have the most accurate ARIMA models while the southern regions of the South and Southeast were found to be the least accurate. This was then supported by the comparison of model accuracy to individual factors influencing PDSI where it was shown that states with higher temperatures generally gave less accurate models. The test for the stationarity assumption in stochastic modeling became inconclusive as there were some differences in the halves of PDSI data but not a significant enough trend to debunk the assumption.

Knowing the affects that precipitation and temperature had on the different drought modeling schemes, the PDSI and SPI indices for each region were analyzed through SAF curves to determine if droughts were affecting each region homogenously. This tested the original objective to determine the applicability of the climatic regions for drought studies. The analysis for SPI did not give much variability in the regions suggesting that the divisions were appropriate from a meteorological perspective. The PDSI results gave much variability in the regions and showed that some regions should be applied to drought studies with caution. The Northwest region specifically seemed to show two subregions under this analysis and suggested that further division seemed necessary for appropriate drought study.

4.2 Study Limitations and Future Recommendations

There were many limitations in this study and methods that could be improved in recommended studies to be performed in the future. The first of these limitations is the exclusive nature of the second chapter focusing solely on GDP as an indicator for drought effects on different sectors of society. While many sectors have monetary value attributed to them and would be captured by the use of GDP, there are also more specific indicators that could be used to not only reflect the economic sector but other sectors of society as well. Future studies should find correlations and impacts between droughts and other indicators such as crop yields, renewable and non-renewable energy production, or mortality rates. A summary or combination of multiple indices would give a more accurate and complete representation of the severity that drought impact has on each country and the international community.

Secondly, while the ARIMA and SAF analyses did give a summary of different drought indices of the regions of the United States, other statistical analyses could have been performed in these regions to give a further understanding of drought compatibility with these climatic regions. Other forecasting models such as Markov chain or neural network modeling could have been performed to support or contest the findings of which region could be most prepared for the prediction of droughts. Other temporal analyses could have been performed in the form of decision tree or probability analysis in order to determine the severity of various drought frequencies which would reflect on how large of an impact droughts have on each climatic region. All of these analyses could have been performed on a wider range of drought indices such as the Palmer Hydrological

Drought Index (PHDI), Palmer Z-score, Crop Moisture Index (CMI), or the Surface Water Supply Index (SWSI). Other temporal scales could have also been applied to the applicable indices that were used in the study of SPI and SPEI instead of just the 1-month time scales.

The last and probably most important limitation is the test of the stationarity assumption when using stochastic modeling for drought forecasting. The procedure to test this assumption performed in this study was introductory and as the results were inconclusive, the recommendation for future studies would be to test this assumption using more in depth analyses. Due to the changing nature of recent climate trends, stationarity analyses such as the Man-Kendall trend, Hurst exponent, or Thiel-Sen Slope analysis could be performed to see if any non-stationarity found in the time series data was significant. If the stationarity of this data was found to be rejected, the use of stochastic models to determine the predictability of droughts in each region would be inappropriate and the use of other modeling methods that did not require stationarity would be recommended.