Characterizing Multi-Decadal Temperature Variability in the Southeastern United States

Marcus De'Andre Williams
CHARACTERIZING MULTI-DECADAL TEMPERATURE VARIABILITY IN THE SOUTHEASTERN UNITED STATES

By

MARCUS DE’ANDRE WILLIAMS

A Thesis submitted to the Department of Earth, Ocean & Atmospheric Science in partial fulfillment of the requirements for the degree of Master of Science

Degree Awarded:
Summer Semester, 2010

Copyright © 2010 Marcus De’Andre Williams All Rights Reserved
The members of the committee approve the thesis of Marcus D. Williams defended on July 2nd 2010.

__________________________________
Mark. A. Bourassa
Professor Directing Thesis

__________________________________
Eric P. Chassignet
Professor Directing Thesis

__________________________________
Philip Sura
Committee Member

__________________________________
David F. Zierden
Committee Member

Approved:

__________________________________
Paul H. Ruscher, Area Representative, Meteorology

__________________________________
Joseph Travis, Dean, Arts and Sciences

The Graduate School has verified and approved the above-named committee members.
ACKNOWLEDGEMENTS

This research was funded by NOAA and USDA. I would like to thank Dr. Chassignet and Dr. O’Brien for this research opportunity and Melissa Griffin for all of her help throughout this project. I would also like to thank the following individuals for their help: Dr. Bourassa, David Zierden and Dr. Sura for their time and guidance that they offered to me, Dmitry Dukhovskoy for his help, Precious Lewis for her suggestions, and Angela Williams for her support.
TABLE OF CONTENTS

LIST OF FIGURES ...................................................................................................................... v

ABSTRACT .................................................................................................................................... vii

INTRODUCTION ........................................................................................................................ 1

DATA ............................................................................................................................................. 4

DIFFERENCE IN MEAN TEMPERATURES .............................................................................. 6

PROBABILITY DISTRIBUTIONS .......................................................................................... 10

RANKED SUM TEST ................................................................................................................ 15

THRESHOLD OCCURRENCES ............................................................................................. 20

SPATIAL CORRELATION ...................................................................................................... 26

CONCLUSIONS ......................................................................................................................... 28

REFERENCES ............................................................................................................................ 30

BIOGRAPHICAL SKETCH ..................................................................................................... 32
LIST OF FIGURES

Fig 1.1: Annual Average temperature for Alabama sampled from the Oklahoma Climatological Survey. The x-axis is years, shaded by decades, and the y-axis is the temperature in °F. Each point represents the average for that year, and shaded blue and red regions represent the five year tendencies. ........................................................................................................................................ 2

Fig. 2.1: Station coverage map of the 104 stations used in the SE US encompassing the states of Alabama, Florida, Georgia, North Carolina, and South Carolina.................................................. 4

Fig. 3.1: Mean minimum temperature map. The WR period is subtracted from the CR period. If the difference is negative, a blue dot is assigned to the station, and positive differences are assigned a red dot. The size of the dot indicates the degree difference with the largest dots indicating a difference of four degrees or more............................................................................. 8

Fig. 3.2: Same as 3.1 except for maximum temperatures......................................................... 8

Fig. 4.1: Winter minimum temperature distribution for Camp Hill, AL. The x-axis represents the temperature in °F, and the y-axis represents the derivative of the probability. The red line represents the warm regime, the blue line represents the cold regime, and the black line represents the base period from 1920-2007.................................................................................................................. 11

Fig. 4.2: Same as fig. 4.1 except for De Funiak Springs, FL....................................................... 13

Fig. 4.3: Same as 4.1 except for fall maximum temperatures in Henderson, NC..................... 14

Fig. 5.1: Regional map showing minimum temperature confidence limits for the rejection of the null hypothesis. Red circles, triangles and squares represent confidence at or above the 99% significance level. Yellow circles represent stations that lie within the 95-98% level, green circles represent stations that lie within the 90-94% level, and purple circles represent stations that lie at or below the 89% level................................................................................................. 16

Fig. 5.2: Same as 5.1 except for maximum temperatures.......................................................... 17

Fig. 5.3: Same as fig. 4.1 except for Albemarle, NC................................................................. 18

Fig. 6.1: Same as 4.1 except for summer maximum temperatures in Talladega, AL.............. 21

Fig. 6.2: Showing the counts of occurrences for the temperature threshold set at 95°F in Talladega, AL. The x-axis represents the number of occurrences per year. The y-axis represents the months, centered on JJA (summer) for the 95°F threshold. The blue line represents the CR, the red line represents the WR, and the black line represents the base period from 1920-2007. . 22

Fig. 6.3: Same as fig. 6.2 except for 90°F or greater occurrences in Cullowhee, NC.............. 23

Fig. 6.4: Same as fig. 4.1 except for Lake City, FL................................................................. 24
Fig. 6.5: Same as fig. 6.2 except for 32°F or below occurrences in Lake City, FL and the x-axis represents the months center on DJF (winter).

Fig. 7.1: Winter temperature spatial correlation map for 104 stations used in study. The x-axis represents the value of the correlation for the perspective distance, and the y-axis represents the distance in kilometers. The red line represents the winter maximum temperatures, and the blue line represents the winter minimum temperatures.
ABSTRACT

Prior studies of the long-term temperature record in the Southeastern United States (SE US) mostly discuss the long-term cooling trend, and the inter-annual variability produced by the region’s strong ties to El Niño Southern Oscillation (ENSO). An examination of long-term temperature records in the SE US show clear multi-decadal variations in temperature, with relative warm periods in the 1920’s through the mid 1950’s and a cool period in the late 1950’s through the late 1990’s. This substantial shift in multi-decadal variability is not well understood and has not been fully investigated. It appears to account for the long-term downward trend in temperatures. An accurate characterization of this variability could lead to improved interannual and long-term forecasts, which would be useful for agricultural planning, drought mitigation, water management, and preparation for extreme temperature events. Statistical methods are employed to determine the spatial coherence of the observed variability on seasonal time scales. The goal of this study is to characterize the nature of this variability through the analysis of National Weather Service Cooperative Observer Program (COOP) station data in Florida, Georgia, Alabama, North Carolina, and South Carolina. One finding is a shift in the temperature Probability Distribution Function (PDF) between warm regimes and cool regimes.
CHAPTER 1 INTRODUCTION

To predict future variability, and to interpret this variability with appropriate confidence, it is highly desirable to have an understanding of past variability and climate cycles. With climate change and its impacts becoming a hot topic globally the need to understand types of long-term variability are becoming more pressing. In the Southeastern United States (SE US) the El Nino-Southern Oscillation (ENSO) signal is widely recognized as the dominate inter-annual mode of climate variability, particularly for temperature and precipitation patterns (Ropelewski and Halpert 1986). Work done by the Intergovernmental Panel on Climate Change fourth assessment report (IPCC AR4 2007) and others discusses climate variability on a global domain ranging from inter-annual to multi-decadal time scales. Much of the analysis done by the IPCC AR4 was in the realm of trend analysis. Concerns about how multidecadal variability might influence the interpretation of these results have been recognized: “although trend analysis has been frequently used to evaluate climate variation over the United States and North America during the twentieth century, low-frequency cyclic variation in the data can cause the significance or even the sign of trends to depend on the period over which trends are fitted” (Mauget 2003). This study examines the Southeastern United States for such longer term variability within a 106 year record.

The primary domain encompassed in the study is the states of Alabama, Florida, Georgia, North Carolina, and South Carolina. The analysis presented in this research identifies a mode of variability for the SE US that has much larger magnitude than the trend, and is influential on a multi-decadal timescale. Many studies have identified the SE US as an outlier from the global pattern of recent warming in the temperature records (Karl et al. 1996, IPCC AR4). One of the more recognized studies done by Easterling et al. (1997) highlights these recent findings. His analysis shows that the SE US and parts of Asia are the only parts of the globe that show a cooling trend. Easterling’s results exhibit a need to fully characterize the long-term variability that has been displayed in the SE US.

Historical warm and cool periods in the SE US have been documented by several studies. (Dickson and Namias 1976; Van Loon and Williams 1976; Diaz and Quayle 1980; Yonetani and McCabe 1994; Dittinger et al. 1995). Many of these studies investigated monthly mean
temperatures across the 344 climate divisions across the United States. Many of these studies have found that changes in temperature are tied to changes in atmospheric circulation from a zonal regime to a meridional regime. Other studies have alluded to multidecadal patterns in the SE US, with ties to teleconnection patterns such as the Pacific North American (PNA) pattern (Kalnicky 1987; Lathers and Palecki 1991, 1992). These studies focused on atmospheric circulation patterns and their effects on United States surface temperatures. We have identified a multidecadal regime in the long-term temperature records that shifts from periods of warmer than normal (Warm regime-WR) to periods of colder than normal (Cold regime-CR). The period for the WR is from 1920-1957 and the CR is during the period from 1958-1998 (fig. 1.1).

http://climate.mesonet.org/climate_trends.html

Fig 1.1: Annual Average temperature for Alabama sampled from the Oklahoma Climatological Survey. The x-axis is years, shaded by decades, and the y-axis is the temperature in °F. Each point represents the average for that year, and shaded blue and red regions represent the five year running mean.
As mentioned previously, around 1920-1957 has been referenced by many studies as being warmer than normal. Yonetani and McCabe noted the abrupt nature of temperature regime changes in the United States, and 1957 has been identified as a year of an anomalous “break” in mid-tropospheric circulation patterns by Lathers and Palecki. The time periods found for this multidecadal regime are in agreement with the timeframe mentioned in previous studies. The analysis presented in this research goes into greater detailed characterization of temperature changes in the SE US.

Several methods of analysis are used to characterize the variability of the long-term temperature record for 104 National Weather Service Cooperative Observation Program stations in the Southeastern United States. The first method was to create temperature probability distribution function (PDF) plots for all stations and for all four seasons. Analysis of the PDF’s yielded a shift in the distribution between the WR and the CR. To assess the statistical significance of the shift between the two regimes the Wilcoxon-Mann-Whitney ranked sum test was utilized. The results of the test proved that the temperatures during the WR and CR were from statistically different populations. The spatial pattern of the regime shift is shown to be largely coherent in the SE US, and present in all seasons.
CHAPTER 2 DATA

The data sets used for this study are a daily summary of the day data set (DS 3200) and a digital version of DS 3206. Both data sets were provided by the National Weather Service Cooperative Observation Program (COOP), which has provided climate and weather data for over 100 years. Each data set consists of 8,000 active observing stations recording daily values of maximum temperature, minimum temperature, and precipitation (NCDC 2008). For this study, 104 COOP stations (fig. 2.1) were selected from the states of Alabama, Florida, Georgia, North Carolina and South Carolina.

Fig. 2.1: Station coverage map of the 104 stations used in the SE US encompassing the states of Alabama, Florida, Georgia, North Carolina, and South Carolina.
The selected stations all have long observations records, dating back to at least 1920, and all meet the criteria set forth by Smith (2006), who stated that a station cannot have more than five consecutive years of missing data.

As with any data set, there are limitations contained in the COOP station data that make it difficult to draw conclusions about the variability being investigated. Biases can be introduced into the data set by changes in observation time, station location, the local environment, and instrumentation. The National Climatic Data Center (NCDC) fashioned the United States Historical Climate Network (USHCN) data set to correct many of the biases mentioned above, and several of the 104 stations we used are part of the USHCN. However, for the purpose of this study the daily, unadjusted observations are used in order to stay as close to original data set as possible and to avoid losing the strength of the variability through bias corrections made to the data set. In some cases the large warming signature of urban stations can eliminate the cooling trends observed at some stations in the SE US. Also at the time of this study, the adjusted observations are only available on the monthly time scale as monthly averages. The preferred time scale for this study is the daily observations, with the minimum and maximum values averages separately. Studies conducted by Easterling et al (1996) and Christy (2002) mentioned the concerns and limitations of the homogenized USHCN data set. In particular Easterling et al (1996) mention the differences between the USHCN and unadjusted data sets are larger for fewer stations.

Missing data are addressed using the methodology developed by Smith (2006). Stations that meet the initial requirements undergo a multiple linear regression method to replace any missing data with data from two to five surrounding stations that lie within a 50-mile radius. Correlations between the reference stations and the surrounding stations are done once the data are detrended and the seasonal cycle is removed. If the correlation is less than 0.6 then that particular nearby station will not be used. If one or more surrounding stations meet the 0.6 correlation threshold, then the data from the reference stations are used to calculate a multiple linear regression. The results of the multiple linear regressions are used to replace the missing data points.
CHAPTER 3 DIFFERENCE IN MEAN TEMPERATURES

The difference between WR and CR mean temperatures is determined and tested for statistical significance. The importance of producing the difference in mean maps is to assess the regional characteristics of the regime change. The magnitude of the regime shift will be shown to be stronger in any particular season, with seasonal changes in the spatial coherence of the signal.

To conduct the analysis, the mean temperature for the WR (1920-1957) is subtracted from the mean temperature for the CR (1958-1998). This is done for each station and done on the maximum and minimum temperatures for each of the four standard seasons (DJF, MAM, JJA, and SON). Since the expectations are that the WR mean temperatures are warmer than the CR mean temperatures, the differences are expected to be negative. The stations with negative differences are assigned a blue dot, and stations with positive differences are assigned a red dot. Stations with red dots imply that the CR means are warmer than the WR means for that particular station. The sizes of the dot assigned to each station are indicative of the degree difference between the two means. The range of dot sizes are from 0-1°F (smallest) to greater than 4°F (largest).

The student t-test (Wilks 1995) is employed in order to deduce if the means of the WR and the CR are significantly different. To calculate the t score the difference between group means are divided by the combined variability of the groups. The equation for calculating the t-score is as follows,

\[
 t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\text{var}_1}{n_1} + \frac{\text{var}_2}{n_2}}} \tag{3.1}
\]

where the var is the variance for the group, and \(n\) is the number of independent observations. \(N\) is divided by the decorrelation timescale if the data fails the independence test at the particular station. The WR is used as the first mean and the CR is used as the second mean. If the t score is positive then the first mean is larger than the second mean, and if negative the second is larger than the first. On average the t scores were very large and positive. Only on a rare occasion was a positive t score below the 99% confidence level.
The shift in temperatures between the two regimes is clearly illustrated by the difference in mean minimum temperature map (Fig. 3.1). The difference in means between the two regimes has been proven to be significant for all seasons in the minimum and maximum temperature. The t scores calculated were on average large and positive, which suggest significance on at least the 99% confidence level. Rarely was a positive t score below 2.5. The differences in means between the two regimes are largest during the winter season, followed by summer, spring and fall. Spring and fall displayed the most negative t scores and those stations that have the negative scores can be seen in difference in mean plots. Central and southern Florida are outliers to this pattern during the fall season. This region of the SE US is normally influenced by a different weather regime during this time of year. Stations that are in close proximity to bodies of water and mountainous areas also experience a dampening in the difference between the means. Some stations could also experience a dampening and sometimes reversal of the signal in the minimum temperatures due to land use changes and anthropogenic influences. This statement needs to be further investigated to validate this suggestion. The next Section investigates the shift in temperatures of the two regimes over a range of temperatures broader than just the mean.

Plots are created for each calendar season. Each station is assigned a dot indicating the sign of the difference and the magnitude of the difference as noted above. Analyses of the maps show a clear, coherent signal that is present in each season for the entire SE US sector. The signal is clearly the strongest in the winter season followed by summer, spring, and fall (fig. 3.1-3.2).
Fig. 3.1: Mean minimum temperature map. The WR period is subtracted from the CR period. If the difference is negative, a blue dot is assigned to the station, and positive differences are assigned a red dot. The size of the dot indicates the degree difference with the largest dots indicating a difference of four degrees or more.

Fig. 3.2: Same as 3.1 except for maximum temperatures.
Within each season the maximum and minimum temperatures are investigated separately. The minimum temperatures display more of a difference than the maximum temperatures with summer mean temperatures being the exception. Minimum mean temperatures showing more of a change could be due to the fact that the boundary layers are more conducive to mixing during the afternoon maximum temperatures more so than the overnight minimum temperatures. The magnitude of the difference in means appears to be greater for spring and fall, than summer. Given the high variability of the spring and fall temperature distributions, the difference in the means are expected to be greater. Although the difference in the means for spring and fall are significant, more confidence can be assessed in the difference in the means for the summer season.
CHAPTER 4 PROBABILITY DISTRIBUTIONS

Probability distribution plots illustrate substantial shifts in temperatures between the warm regime (WR; 1920-1957) and cool regime (CR; 1958-1998). The plots are created for each station, by making Probability Distribution Functions (PDFs) of seasonally averaged minimum and maximum temperatures. These plots allow for an examination of the shift in temperatures over the whole range of temperatures, including both means and extremes. Section three illustrated the difference in the means between the two regimes and verified that the differences are statistically significant. The occurrence of extremes and the statistical significance of the shifts in temperature between the two regimes are highlighted in Sections five and six of this study.

To create the PDF for each station, data from the WR and CR are placed into either 5°F (winter) bins, 3°F (spring, summer) bins or 8°F (fall) bins. This allows for the creation of histograms that contain counts of occurrences of a given temperature for the season of analysis. Since the number of years in the two regimes differs, to make a fair comparison of the regimes it is necessary to normalize the histogram: the product of the total number of observations and the bin width divides the histogram values. Figure 4.1 shows a clear shift in the distributions of the PDF between the two regimes.
As an example the winter minimum distribution for the COOP station located in Camp Hill, AL is shown. The winter temperature distribution (Fig.4.1) has warmer temperatures for the WR and colder temperatures for the CR. The winter temperature distribution for the CR has a much sharper peak than the distribution for the WR, with the peak for the CR occurring between the 21-26°F temperature ranges while the peak for the WR occur around the 31-36°F temperature range. The peaks for the distributions can be representative of the median values for each regime. The distribution for the 41-year period defined as the cold regime is positively skewed, while the 37-year period identified as the warm regime has a negative skewness. The negative skewness for the WR distributions aligns the temperatures in the distribution towards the higher ranges of the winter minimum temperatures. The opposite holds true for the positive skewness of the CR. As highlighted earlier in Section three there is a clear difference in the mean value between both regimes, which is 5.60°F for Camp Hill, Alabama (fig. 4.1). Analysis of the behavior of the distribution over all temperatures shows that the tails of the distributions are
affected, but there is very little change in the most extreme values at both ends of the distributions. Section six will further investigate the behavior of the tails of the distribution through threshold occurrence analysis.

All seasons display evidence of a shift between the two regimes for minimum and maximum temperatures. Most minimum temperature distributions generally follow the pattern displayed by the Camp Hill, AL COOP station and De Funiak Springs, FL (fig 4.2). Winter is the season that typically has the largest shift between the two regimes, which is in agreement with the results obtained from the difference in mean analysis. The minimum temperature distributions have a more prominent shift when compared to the maximum temperature distributions. Summer distributions have a greater shift than the spring and fall distributions. The magnitude of the shift in the summer distributions is the same as for minimum and maximum temperatures. Spring distributions behave sporadically, with some stations showing moderate shifts and other stations showing little to no change. An interesting finding resides in the fall maximum temperature distributions where the shift in the distribution occurs primarily at the warmer ranges of temperatures. For Henderson, NC (fig. 4.3), the chance of having a temperature around 89°F greatly increases during the WR. This suggests the extension of summer like temperatures well into fall during the WR. These general results are found throughout the SE US; the spatial consistency is examined in Section seven.
Fig. 4.2: Same as fig. 4.1 except for De Funiak Springs, FL.
Fig. 4.3: Same as 4.1 except for fall maximum temperatures in Henderson, NC.
CHAPTER 5 RANKED SUM TEST

In Sections three and four a shift in temperatures between the WR and CR are illustrated, and the differences in the means are proven to be statistically and physically significant. Probability distribution function plots were analyzed (Section 4) to determine the behavior of the regimes over a broad range of temperatures. To prove the differences between the distributions of the two regimes are statistically significant the Mann-Whitney-Wilcoxon rank sum test is utilized (Wilcoxon 1945). This is an important step to verify that the changes between the regimes are valid, and not a result of random noise or a systematic error. If the distributions are proven to be statistically different, the likelihood of finding the underlying physical mechanism of causation is greatly improved.

The Mann-Whitney-Wilcoxon rank sum test applies to two sets of samples, with the null hypothesis being that the two data samples are drawn from the same distribution. One sided and two sided hypothesizes are possible with the test. The rank sum test statistic is not a function of the data values, but a function of the ranks within the pooled samples. The Mann-Whitney U-statistic is calculated by \( U_1 = R_1 - 0.5 \, n_1(n_1+1) \) and \( U_2 = R_2 - 0.5 \, n_2(n_2+1) \) where \( R_1 \) is the sum of ranks (in the pooled collection) of the members of the first sample, \( R_2 \) is the sum of ranks (in the pooled collection) of the members of the second sample, \( n_1 \) is the number of members of the first sample, and \( n_2 \) is the number of members of the second sample.

The Mann-Whitney U-statistic is calculated for each regime by season and for minimum and maximum temperature. The \( U_1 \) test statistic is assigned to the CR, and \( U_2 \) is assigned to the WR. To ensure the independence of the data within each regime, the decorrelation timescale (\( \tau \)) is calculated for each station. The decorrelation time scale is calculated by integrating the autocorrelation function from zero time lag to at the maximum lag time, which is the number of years in the period minus ten. A threshold criterion was set at .9 for \( \tau \). If a station fails to meet the threshold criteria set, the station is sub-sampled every \( \tau \times 1.5 \) years (a conservative approach). Z-scores are obtained for each station along with the probability that the null hypothesis is rejected. The formula for calculating the z-score is as follows:

\[
z = \frac{U - m_U}{\sigma_U} \tag{5.1}
\]
Confidence limits are set on a station by station basis for the rejection of the null hypothesis. Each station is assigned a colored symbol corresponding to a range of confidence limits. Stations with confidence limits $\geq 99\%$ are coded red, stations that lie within the 95-98% confidence interval are coded yellow, stations that lie within the 90-94% confidence intervals are coded green, and stations $<90\%$ confidence limit are coded purple. As seen in fig. 5.1 and fig5.2, during the winter and summer seasons roughly one fifth of the stations fall below the 90% confidence limit. The fraction of stations that fail to meet the 90% or greater confidence intervals increases during the spring and fall to about one third of all stations. Many stations have confidence limits that are extremely high (red squares and triangles).

$$m_U = \frac{n_1 n_2}{2} \quad (5.2)$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} \quad (5.3)$$

Fig. 5.1: Regional map showing minimum temperature confidence limits for the rejection of the null hypothesis. Red circles, triangles and squares represent confidence at or above
the 99% significance level. Yellow circles represent stations that lie within the 95-98% level, green circles represent stations that lie within the 90-94% level, and purple circles represent stations that lie at or below the 89% level.

Fig. 5.2: Same as 5.1 except for maximum temperatures.

Using Albemarle, NC, as an example of a failing station, fig. 5 shows that there is no visual difference in the winter minimum temperature distribution of a shift between the two regimes. One would expect there to be no statistical evidence for the two data samples to be pulled from different distributions. One possible cause for the failure of the Albemarle location is cold air draining. Albemarle is located in a river valley that funnels southward to a low-lying area that is around 400 ft lower in elevation than the surrounding area. Winter temperatures are
dominated by frontal systems for this region, and the cold from the north funnels through Albemarle.

Fig. 5.3: Same as fig. 4.1 except for Albemarle, NC.

Several additional stations that fell below the 90% confidence limits also had seasonal temperature distributions that displayed no visual shift, and there are general characteristics that appear to cause stations to fail to meet the confidence limit threshold set. For winter minimum temperatures, two of the most glaring causes are proximity to mountains (likely causing a great deal of noise, and possibly also reducing the signal) and proximity to bodies of water (likely reducing the signal). Of the 19 stations that failed for winter minimum temperatures, seven of the stations are located in the Blue Ridge Mountain region in North Carolina. The elevations of these stations severely dampen the shift in temperatures between the WR and CR. In general, close proximity to water, land cover change, and urbanization seem to dampen the strength of the signal. Although these relationships are outside the scope of this research, it appears that some
sites have microclimates that either enhance or minimize the effects of a given synoptic regime. If the land cover was altered at a particular station, it would likely have a greater affect on the winter minimum temperatures than the summer maximum temperatures.

In summary, visible shifts in the seasonal temperature distributions are detected in several stations in the SE US. In order to quantify if the shifts in the temperature distributions are statistically significant, the Mann-Whitney-Wilcoxon rank sum test is employed. The purpose of the Mann-Whitney-Wilcoxon rank sum test is to determine if the two data samples being tested are pulled from the same distribution. The null hypothesis is that the data are pulled from the same distribution, and if the null hypothesis is rejected it is safe to assume that shifts in the distribution are statistically significant and physically meaningful. Many stations had extremely high confidence limits (>99.999%). For winter and summer temperature distributions, roughly one fifth of the 104 stations analyzed fell below the 90% confidence limit. For spring and fall temperature distributions, roughly one third of the stations fell below the 90% confidence limit. The results for spring and fall are expected because of their high variability. Very few of the stations had around confidence intervals from 90% to <99%, particularly for winter minimum temperatures (5); more stations fell into this range during fall maximum temperatures (20). The strength of the signal or the amount of noise appear to be modified by elevation, proximity to water, land cover change, and urbanization, but these statements need to be further investigated for their validity.
CHAPTER 6 THRESHOLD OCCURRENCES

The shift in the seasonal temperature PDFs implies there is an increase in occurrences of temperature found at one tail of the PDF distributions. Knowledge of this variability minimizes adverse impacts on the agricultural industry, water resources, and energy planning. Thresholds are set at 32°F and 95°F and used to determine the annual average count of occurrences outside these thresholds. This analysis explores the behavior of the two regimes at these temperature ranges. The 32°F is chosen because it signifies the freezing level, which is relevant for agricultural and many other applications that are of importance to the region. The 95°F threshold is chosen because temperature values around 90°F are fairly common in most of the SE US region during the summer months and doesn’t provide an appropriately extreme value to compare the two regimes. High temperature values around the 95°F range can implicate heat stress, heat related mortalities (Kalkstein and Davis 1989). Talladega, AL and Lake City, FL are used to illustrate the change in threshold occurrences between the warm regime and cold regime.

Talladega, AL provides an example for the change in the number of occurrences of 95°F or higher (fig. 6.1), demonstrating the apparent shift in the summer maximum temperature distribution around the 95°F temperature range.
Section three has proven that the differences in mean temperatures are significant at this location, and Section five has proven that the shifts in the distributions are significant for this regime. Fig 6.2 shows that the actual counts of occurrences are in good agreement with the likelihood of occurrence (fig. 6.1). The number of times per year the temperature in Talladega reached 95°F (fig. 6.2) for the CR (~4) doubled for the WR (~10). It is important to note the large disparities in late spring and early fall between the two regimes. The CR records for May shows around 2 occurrences per year of temperatures ≥95°F, but that nearly triples for the WR to approximately 8 occurrences per year. This same pattern holds for late fall where there are around three times as many occurrences during the warm regime. This implies that it gets hotter faster, and stays hotter much longer during the WR versus the CR. The extension of summer like temperatures into early fall was briefly mentioned in Section four, where it was noted that the greatest disparity between the fall minimum and maximum PDF’s for the cold and warm regime happened around the warmer range of temperatures. Fig. 6.2 demonstrates that the finding that summer like temperatures are extended into early fall. On a yearly basis the warm regime records 31 days at or above 95°F, while the CR records 11 days per year. If the pattern of the regime change is able to be modeled and predicted, this result could have great impacts on energy usage and planning.
during the summer months. Most stations followed this general pattern shown by the Talladega station; however, there are some regional differences. Some stations display little or no difference in occurrences between the two regimes, and in some cases there are a greater number of occurrences in July for the CR. No pattern has been determined in stations that display these atypical patterns, but land use or anthropogenic influences can be speculated as responsible for these atypical results. Some stations have not recorded any temperatures greater than 95°F: For the stations that average less than one occurrence of temperatures at or above 95°F per year (fig. 6.3) there is a shift in the number of occurrences ≥90°F. Some stations in Florida display similar numbers of ≥95°F days between the two regimes, but there is still evidence that summer like temperatures are extended into middle to late October.

Fig. 6.2: Showing the counts of occurrences for the temperature threshold set at 95°F in Talladega, AL. The x-axis represents the number of occurrences per year. The y-axis represents the months, centered on JJA (summer) for the 95°F threshold. The blue line represents the CR, the red line represents the WR, and the black line represents the base period from 1920-2007.
Winter 32°F threshold occurrences follow a similar format for the CR as the maximum temperature occurrences for WR, although there is no clear cut evidence that there is an extension of winter like temperatures into late fall and early spring. Lake City, Florida is used as a general example, and regional differences are discussed. The winter minimum temperature distribution for Lake City (Fig. 6.4) demonstrates a shift in the PDF around the 32°F temperature range, which changes the number of occurrences per year for the two regimes (Fig. 6.5). The peak value for the CR takes place in January, where on average there are approximately seven occurrences in January for the CR, compared to the peak value of four per year for the WR. Comparing the yearly totals for each regime, the CR experienced 18 occurrences of 32°F per year while the WR experienced 11 per year. Comparing the standard calendar winter for Lake City, the average CR winter 16 days at or below 32°F, while the WR only totaled 9 days for the average winter. As one would expect, latitudinal differences contribute to the number of days the 32°F threshold is reached. Florida recorded the least amount of occurrences, while Northern
Alabama and North Carolina recorded the greatest amount of occurrences. Most non-Florida stations experienced an increase in occurrences of 10 to 20 per years during the CR.

Fig. 6.4: Same as fig. 4.1 except for Lake City, FL.
Fig. 6.5: Same as fig. 6.2 except for 32°F or below occurrences in Lake City, FL and the x-axis represents the months center on DJF (winter).
CHAPTER 7 SPATIAL CORRELATION

The shift in temperatures between the two regimes is shown to be spatially coherent throughout the SE US in the difference in mean maps. This result is expected because the temperatures at the stations in this region are well correlated spatially. Verification of the spatial correlation of the stations is an important step. It confirms that the behavior of the regime shift is a phenomenon that takes place throughout the region.

The spatial correlation is calculated by determining the correlation between each pair of stations, and then averaging the correlations as a function of distance between the stations. Stations are grouped by distance, and binned into 16 groups.

The temperatures at the stations are almost all well correlated with each other for all seasons and for minimum and maximum temperatures. Warm periods in Florida coincide with warm periods in North Carolina, South Carolina, Georgia, and Alabama. This pattern is upheld to a greater degree for periods of cooler temperatures. Fig. 7.1 illustrates this pattern for winter average minimum and maximum temperatures. Correlations for winter temperatures in the region remain above 0.6. Winter and spring temperatures show the best correlation spatially, while summer and fall temperatures show lower correlations at greater distances. This result is expected because of the disparity in differential heating that would take place between a station located in North Carolina, and south Florida. There are some stations in North Carolina that have not recorded a temperature above 95°F during their entire period of observation. Also as previously mentioned, if this variability is driven by continental air masses from upstream, those types of weather systems generally aren’t present in central and southern Florida early in the fall season. The relatively poor correlation between coastal and mountain stations (particularly south Florida stations) and other stations causes the correlations to drop from unity at zero distance to approximately 0.8 in the next distance bin. For larger distances the change in correlation is very small, supporting the conclusion of very high spatial correlation for most regions.
Fig. 7.1: Winter temperature spatial correlation map for 104 stations used in study. The x-axis represents the value of the correlation for the perspective distance, and the y-axis represents the distance in kilometers. The red line represents the winter maximum temperatures, and the blue line represents the winter minimum temperatures.
CHAPTER 8 CONCLUSIONS

As the earth’s climate continues to evolve, scientists are conducting studies to see how the climate will evolve in the future, how the evolution of the climate will affect human life on earth, and what role humans have played in this evolution. While scientists have found many answers about the natural variability of the earth’s climate and anthropogenic induced variability, not all of the pieces of the puzzle are in place. This places an importance to identify and characterize past climate variability, and to state with reasonable bounds of confidence what might take place in the future. The SE US has long been identified as an outlier from the global trend. This study characterizes the long-term variability observed in the SE US, which is much greater than the variability associated with a linear trend.

Daily maximum and minimum temperatures are analyzed over the time period from 1920-2007, and a clear shift in the daily minimum and maximum temperatures between 1920-1957(WR) and 1958-1998(CR) is identified. This regime shift is characterized and compared through analysis of mean temperatures, occurrences of extreme values, and shifts of the temperature over the entire range of the distribution. Analysis is broken into seasons to distinguish if the phenomenon takes place in a particular season, or if the regime shift is prevalent in all seasons.

It is shown that there is a significant shift in temperatures between the two regimes, as large as 4-6°F in some locations. Section five of this study proved that the observed shifts in temperature between the two regimes are overwhelmingly significant, especially for the winter and summer seasons. The signal of the regime shift is nearly consistent in all seasons and in all of the States of analysis (AL, FL, GA, NC, SC), with winter minimum and maximum temperatures displaying the strongest shift. The signal of the regime shift is strongest in Alabama, and harder to distinguish in some coastal stations, some stations in close proximity to water, and some stations in mountainous regions. Inland regions of the SE US display the strongest signal, suggesting that the regime change is the result of the influence of continental air masses passing through the region. There is a large shift in the mean temperatures as noted in Section three of this study, and these differences are proven to be statistically significant as well. The change in seasonal mean temperatures between the WR and CR illustrates the spatial coherence of regime change, showing that the abrupt change in temperatures is a phenomenon
taking place throughout the region, and not at just one station, or one State. This result is also verified by the fact that the temperatures at most of the analyzed stations are well correlated with each other for all seasons for minimum and maximum temperatures. As expected, the shift in temperatures between the two regimes affected the occurrences of temperature at the $\geq 95^\circ F$ threshold as well as temperatures at the $\leq 32^\circ F$ threshold. The number of occurrences of $\geq 95^\circ F$ increased during the WR, and the number of days $\leq 32^\circ F$ increased during the CR. The disparity of occurrences between the two regimes is much larger for the $\leq 32^\circ F$ threshold. Knowledge and understanding of this variability is of importance to the agricultural industry, water resources, and energy planning. Questions still need to be answered about what caused this regime change, and if a similar phenomenon present in other parts of the United States or the World.
REFERENCES

Christy, J.R. (2002). When was the Hottest summer? a state climatologist struggles for an answer. BAMS, 723-734.


BIOGRAPHICAL SKETCH

Marcus De’Andre Williams was born on April 21, 1982 in Tallahassee, Florida. He attended Florida State University from 2000-2006 and graduated with his Bachelor of Science with a major in Meteorology and a minor in Mathematics and Physics. After graduation in 2006, he began working at the Center for Atmospheric Prediction Studies (COAPS) in the Florida Climate Center working closely with David Zierden and Melissa Griffin. In August of 2008 he began working as a graduate student for Dr. Mark A. Bourassa and Eric P. Chassignet at COAPS. He will continue working in the area of climate research at the University of Georgia.